A Method for Assessing Multivariate Measurement Systems

Michele Scagliarini[1],

[1] *Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy, michele.scagliarini@unibo.it*

Abstract-Multivariate measurement systems analysis is usually performed designing suitable gauge R&R experiments ignoring available data generated by the measurement system while used for inspection or process control. This work proposes an approach that, using the data that are routinely available from the regular activity of the instrument, offers the possibility of assessing multivariate measurement systems without the necessity of performing a multivariate gauge study. It can be carried out more frequently than a multivariate gauge R&R experiment, since can be implemented at almost no additional cost. Therefore the synergic use of the proposed approach and the traditional multivariate gauge R&R studies can be a useful strategy for improving the overall quality of multivariate measurement systems and is effective for reducing the costs of multivariate measurement systems analysis performed with a certain frequency.

I. Introduction

In a manufacturing environment, critical decisions about process and product quality depend on the quality of the measurement systems. Measurement systems analysis (MSA) is a set of statistical techniques used to quantify the uncertainty of the measurement instruments. [1] provided a review of gauge repeatability and reproducibility (R&R) methods for assessing the precision of measurement systems. In the case of univariate measurement systems, several MSA-approval metrics are commonly used. For an overview on this topic we suggest [2]. Multiple characteristics processes are by now so common that studies concerning measurement systems analysis cannot be restricted to the univariate framework. In the case of multiple quality characteristics, [3] proposed multivariate extensions of three commonly used MSA-approval criteria using a multivariate analysis of variance (MANOVA) method of estimating the covariance matrices for one-factor and two-factor gauge studies. In order to ensure constant flows of reliable data, manufacturers should periodically assess their measurement systems and the costs involved in maintaining well performing measurement systems are normally relevant. This issue motivates the present work. Multivariate measurement systems analysis is usually performed designing suitable gauge R&R experiments ignoring available data generated by the measurement system while used for inspection or process control. In recent literature, the use of these measurements from regular use of the instrument has been suggested for univariate MSA studies (see e.g, [4]). Here we propose the following approach. Once assessed, in the initial set up, the multivariate measurement instrument as adequate, its performances are assumed as benchmark. Therefore, using the data from the regular activity of the instrument, the periodic assessments of the measurement device are performed by comparing the present precision with the benchmark through a statistical test. Since the proposed method does not require to perform a multivariate gauge study, our proposal can be a useful tool for reducing the costs of multivariate MSA carried out with a certain frequency. The paper is organized as follows. Section II describes the multivariate MSA-approval criteria proposed in recent literature. Section III develops the test for assessing multivariate measurement instruments. Section IV studies the performances of the proposed method. Finally, section V contains the conclusions.

II. Multivariate measurement systems analysis

Let $\mathbf{X}' = [X_1, X_2, ..., X_m]$ represent the vector of m quality characteristics with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ positive definite. We assume that the multivariate process data are from a multivariate normal distribution. Let us denote with $\mathbf{LSL}' = [LSL_1, LSL_2, ..., LSL_m]$, $\mathbf{USL}' = [USL_1, USL_2, ..., USL_m]$ and $\mathbf{T}' = [T_1, T_2, ..., T_m]$ the m-vectors values of the lower specification limits, upper specification limits and target values respectively. MSA methodology assumes the model **Y**=**X**+**e** where **Y** is the vector of the observable quality characteristics, which is usually obtained from some physical measurements, **X** is the true quality characteristics vector that we are interested in monitoring and **e** is the multivariate measurement error. It is assumed that $\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e)$ with $\boldsymbol{\Sigma}_e$ positive definite and that **X** and **e** are independent. As a result, $\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_y)$ where $\boldsymbol{\Sigma}_y = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_e$. Let us denote with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$ the eigenvalues of $\boldsymbol{\Sigma}$, with $\lambda_{e1} \geq \lambda_{e2} \geq \ldots \geq \lambda_{em}$ the eigenvalues of $\boldsymbol{\Sigma}_e$ and with $\lambda_{y1} \geq \lambda_{y2} \geq \ldots \geq \lambda_{ym}$ the eigenvalues of $\boldsymbol{\Sigma}_y$. In the multivariate framework, [3] proposed the **multivariate precision-to-tolerance ratio**, which according to the specification of the tolerance region simplifies in $P/T_{1m} = \left(\left(\prod_{i=1}^{m} \sqrt{\chi^2_{\alpha,m} \lambda_{ei}} \right) \pi^{m/2} \Big/ \left(\prod_{I=1}^{m} TOL_i \right) \Gamma(1+m/2) \right)^{1/m}$ when a hypercube-shaped tolerance region is used, and $P/T_{2m} = \left(\prod_{i=1}^{m} \frac{2\sqrt{\chi^2_{\alpha,m}\lambda_{ei}}}{TOL_i} \right)^{1/m}$ for the case of a hyperellipsoid-shaped tolerance region. In the above equations $\Gamma(\cdot)$ is the gamma function, $TOL_i = USL_i - LSL_i$ and $\chi^2_{\alpha,m}$ is the 100(1-α)*th* percentile of the χ^2 distribution with m degrees of freedom with (1-α)100% usually fixed at 99%. The other two criteria are the **multivariate percent R&R ratio** $\%R\&R_m = \left(\prod_{i=1}^{m} \sqrt{\lambda_{ei}/\lambda_{yi}} \right)^{1/m} 100$ and the **multivariate signal-to-noise ratio** $SNR_m = \sqrt{2} \left(\prod_{i=1}^{m} \sqrt{\lambda_i/\lambda_{ei}} \right)^{1/m}$. The author also gave some guidelines for gauge acceptance. Approval values for P/T_{1m} and P/T_{2m} range from 0 to 0.3, $\%R\&R_m$ should be ≤30%, while based on SNR_m a measurement system is adequate when $SNR_m \geq 5$. Since the covariance matrices $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_y$ are usually unknown, [3] also proposes a multivariate analysis of variance (MANOVA) method of estimating the covariance matrices for one-factor and two-factor gauge studies. For brevity, we refer to the work of [3] for complete details about the MANOVA methodology.

III. A test for multivariate measurement systems

Let us suppose that at the beginning of the manufacturing activity, that for notation purpose we denote as time T=0, a multivariate MSA is performed and that the measurement instrument is assessed as adequate. We denote with $\boldsymbol{\Sigma}_{e0}$ the precision of the measurement instrument, with $\boldsymbol{\Sigma}_0$ the covariance matrix of the true quality characteristic and with $\boldsymbol{\Sigma}_{y0} = \boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_{e0}$ the covariance matrix of the measurements, at time T=0. After the initial set up, the measurement device is usually used for inspection or process control generating a lot of data at no additional costs. Let us consider a time interval in which the instrument has been routinely used. At time $T=t$ the measurement instrument is characterized by a precision $\boldsymbol{\Sigma}_{et}$ and, assuming the stability of the process i.e. $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_0$, the variability of **Y** is $\boldsymbol{\Sigma}_{yt} = \boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_{et}$. In this framework, we are interested in the detection of a worsening in the measurement instrument precision and differences in the variability of the observed measures are only caused by changes in the precision since $\boldsymbol{\Sigma}_{yt} = \boldsymbol{\Sigma}_{y0}$ if and only if $\boldsymbol{\Sigma}_{et} = \boldsymbol{\Sigma}_{e0}$. Let us consider the hypothesis $H_0 : \boldsymbol{\Sigma}_{et} = \boldsymbol{\Sigma}_{e0}$ (instrument precision at instant 0 is equal to the precision at instant t) and the alternative $H_1 : \boldsymbol{\Sigma}_{et} - \boldsymbol{\Sigma}_{e0}$ is positive definite (precision at instant t is worse than precision at instant 0). Let **S** be the sample covariance matrix of a random sample of size n from **Y** at time t. If H_0 holds, then $(n-1)\mathbf{S}\boldsymbol{\Sigma}_{y0}^{-1} \sim W(\mathbf{I}, n-1)$ where $W(\mathbf{I}, n-1)$ denotes a Wishart distribution with parameters **I** and $n-1$. If we denote with $\hat{\lambda}_{\Omega 1}$ the largest eigenvalue of the matrix $\boldsymbol{\Omega} = \mathbf{S}\boldsymbol{\Sigma}_{y0}^{-1}$, then H_0 is not rejected if and only if $(n-1)\hat{\lambda}_{\Omega 1} < u_1$, where u_1 is the upper α percentage point of the largest characteristic root of a Wishart matrix. The advantage of this method is that at time $T=t$ the measurement instrument is assessed by comparing the current precision with the precision at time T=0, without the necessity of performing a multivariate gauge study (MANOVA) since the sample covariance matrix **S** can be estimated using the data available by the routine use of the measurement device at no additional costs.

IV. Case Studies

To discuss the ability of the test for detecting worsening in the measurement instrument performances, we consider as the situation at time T=0 (the benchmark) the case discussed by [3], then we examine a variety of worsening-precision scenarios at time $T=t$. For each of the proposed scenarios, we compute the multivariate

MSA approval metrics presented in Section II and we design suitable simulation experiments for studying the performances of the proposed test. Let us therefore consider the case discussed in [3] where the data come from an automotive body panel gauge-study involving a four dimensional (m=4) process (see Table 1 of [3] for the original data). Using a two-factor MANOVA method the matrices estimates are

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} 0.01811 & 0.01600 & -0.02180 & -0.00763 \\ 0.01600 & 0.25163 & -0.15732 & 0.35463 \\ -0.02180 & -0.15732 & 0.20856 & -0.39249 \\ -0.00763 & 0.35463 & -0.39249 & 0.98631 \end{bmatrix}, \quad \hat{\mathbf{\Sigma}}_e = \begin{bmatrix} 0.00094 & 0.00168 & -0.00141 & 0.00189 \\ 0.00168 & 0.00632 & -0.00475 & 0.00702 \\ -0.00141 & -0.00475 & 0.00486 & -0.00581 \\ 0.00189 & 0.00702 & -0.00581 & 0.00852 \end{bmatrix}$$

and

$$\hat{\mathbf{\Sigma}}_y = \begin{bmatrix} 0.01905 & 0.01768 & -0.02321 & -0.00574 \\ 0.01768 & 0.25795 & -0.16207 & 0.36165 \\ -0.02321 & -0.16207 & 0.21342 & -0.39830 \\ -0.00574 & 0.36165 & -0.38830 & 0.99483 \end{bmatrix}$$

The eigenvalues of the covariance matrices $\hat{\mathbf{\Sigma}}$, $\hat{\mathbf{\Sigma}}_e$, and $\hat{\mathbf{\Sigma}}_y$ are reported in Table 1 (Panel A) and the results concerning the multivariate MSA metrics are summarized in Table 1 (Panel B), where in order to make our analysis as general as possible, we computed the precision-to-tolerance ratio for both shapes of the tolerance regions. The results reported in Table 1 show that the measurement instrument is assessed as acceptable by all the multivariate criteria, therefore it can be used in the manufacturing process and for our purposes we can assume $\hat{\mathbf{\Sigma}}_e = \mathbf{\Sigma}_{e0}$, $\hat{\mathbf{\Sigma}} = \mathbf{\Sigma}_0$ and $\hat{\mathbf{\Sigma}}_y = \mathbf{\Sigma}_{y0}$.

Panel A: Eigenvalues				Panel B: Multivariate MSA metrics	
λ_1	λ_2	λ_3	λ_4	P/T_{1m}	P/T_{2m}
1.29428	0.11185	0.05438	0.00410	0.25031	0.18654
λ_{e1}	λ_{e2}	λ_{e3}	λ_{e4}	$\%R\&R_m$	SNR_m
0.01908	0.00081	0.00050	0.00025	12.26061	11.30385
λ_{y1}	λ_{y2}	λ_{y3}	λ_{y4}		
1.311189	0.11392	0.05557	0.00457		

Table 1. Eigenvalues of the estimated covariance matrices and multivariate MSA metrics

Now we examine several scenarios where are considered realistic worsening of the measurement instrument after a period of use. We base this discussion on the spectral decomposition of $\mathbf{\Sigma}_{e0}$: $\mathbf{\Sigma}_{e0} = \mathbf{U}_{e0}\mathbf{D}_{e0}\mathbf{U}'_{e0}$, where $\mathbf{U}_{e0}=(\mathbf{u}_{e01},\mathbf{u}_{e02},...,\mathbf{u}_{e0m})$ is the matrix of eigenvectors with columns $\mathbf{u}_{e0i}$ (i=1,2,…,m), and $\mathbf{D}_{e0}=diag(\lambda_{e01},\lambda_{e02},...,\lambda_{e0m})$ is the diagonal matrix of the eigenvalues. The diagonal matrix $\mathbf{D}_{e0}$ is the covariance matrix of the latent independent factors that represent the primary independent sources of variability introduced by instrument at time T=0. The instrument after a period of use is characterised by a measurement error covariance matrix $\mathbf{\Sigma}_{et} = \mathbf{U}_{et}\mathbf{D}_{et}\mathbf{U}'_{et}$ where $\mathbf{U}_{et}=(\mathbf{u}_{et1},\mathbf{u}_{et2},...,\mathbf{u}_{etm})$ with columns $\mathbf{u}_{eti}$ (i=1,2,…,m) and $\mathbf{D}_{et}=diag(\lambda_{et1},\lambda_{et2},...,\lambda_{etm})$. Many alternative cases for $\mathbf{\Sigma}_{et}$ worse than $\mathbf{\Sigma}_{e0}$ can be considered, however $\mathbf{\Sigma}_{et}$ cannot be attained by changing the elements of $\mathbf{\Sigma}_{e0}$ arbitrarily. It is realistic to examine for $\mathbf{\Sigma}_{et}$ cases where changes in the variability are due to changes in variance of the latent independent factors. In other words, the factors that cause the variability in the instrument at instant t remain the same as for instant 0, but with larger variance. This means that cases for $\mathbf{\Sigma}_{et}$ with practical meaning should be those involving the eigenvalues, $\lambda_{eti} \geq \lambda_{e0i}$ for (i=1,2,…,m), but maintaining unchanged the eigenvectors ($\mathbf{U}_{e0}=\mathbf{U}_{et}$). A change in the eigenvectors can be interpreted as the presence of serious problem in the instrument such that the independent sources of variability become dependent. In what follow, we consider two cases for $\mathbf{\Sigma}_{et}$ where the eigenvectors remain unchanged. Furthermore, for the sake of completeness, we will also consider the case of a change in the eigenvectors. Before discussing the details of the case studies it is useful to spend a few words reminding that the multivariate MSA-metrics are designed thinking at different ways for assessing the measurement precision. The

P/T_{1m} and P/T_{2m} criteria compare the multivariate instrument variability with the multivariate tolerance region (hypercube or hyperellipsoid). The remaining metrics do not involve the tolerances: $\%R\&R_m$ expresses the relative widths of the multivariate distributions of the errors **e** and the measured values **Y**; SNR_m compares the width of the multivariate distribution of the true quality characteristics **X** with the corresponding volume of the multivariate errors **e**. Since the test in question does not involve the tolerance regions, we expect a test behavior similar to those of $\%R\&R_m$ and SNR_m. Now we examine the case where the eigenvalues of $\mathbf{\Sigma}_{e0}$ are increased by the same additive term δ : $\mathbf{D}_{et}=\mathbf{D}_{e0}+\delta\mathbf{I}$. Note that this case is equivalent to add the diagonal matrix $\delta\mathbf{I}$ directly to $\mathbf{\Sigma}_{e0}$: $\mathbf{\Sigma}_{et}=\mathbf{\Sigma}_{e0}+\delta\mathbf{I}$ since from the spectral decomposition of $\mathbf{\Sigma}_{e0}$ we get the expression $\mathbf{U}'_{e0}\left(\mathbf{\Sigma}_{e0}+\delta\mathbf{I}\right)\mathbf{U}_{e0}=\mathbf{U}'_{e0}\mathbf{\Sigma}_{e0}\mathbf{U}_{e0}+\delta\mathbf{U}'_{e0}\mathbf{U}_{e0}=\mathbf{D}_{e0}+\delta\mathbf{I}=\mathbf{D}_{et}$. Thus, the eigenvectors remain unchanged, $\mathbf{U}_{e0}=\mathbf{U}_{et}$, and the eigenvalues can be expressed as $\lambda_{eti}=\lambda_{e0i}+\delta$ for i=1,2,…,m. Within this case we consider scenarios where the worsening term δ ranges from 0 to 0.006 (0.0001). Therefore, for each value of δ: a) we compute the multivariate MSA approval criteria and the results are shown in Figure 1a; b) using the R-software we generate 10^4 samples (n=50,75,100,150) from $\mathbf{Y}\sim N\left(\boldsymbol{\mu},\mathbf{\Sigma}_0+\mathbf{\Sigma}_{et}\right)$ where $\mathbf{\Sigma}_{et}=\mathbf{\Sigma}_{e0}+\delta\mathbf{I}$. For each sample we compute the statistic $\left(n-1\right)\hat{\lambda}_{\Delta 1}$, where $\hat{\lambda}_{\Delta 1}$ is the largest eigenvalue of the matrix $\mathbf{S}\mathbf{\Sigma}_{y0}^{-1}$ and **S** is the sample covariance matrix estimated from the sample. Therefore, we evaluate the power of the test computing for each value of δ and n the proportion of rejections of H_0. Note that, fixed $\alpha=0.05$, we used the R-package **RMTstat** ([5]) for computing the critical values of the test. The simulation results are summarized in Figure 1b where are reported, for each sample size, the rejection rates of hypothesis H_0 as a function of δ.

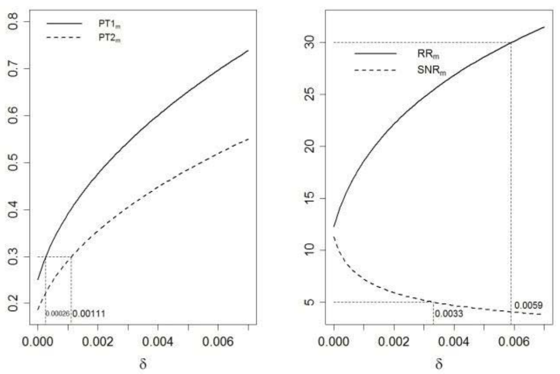

Figure 1a. Multivariate MSA-approval criteria as a function of δ for Case 1

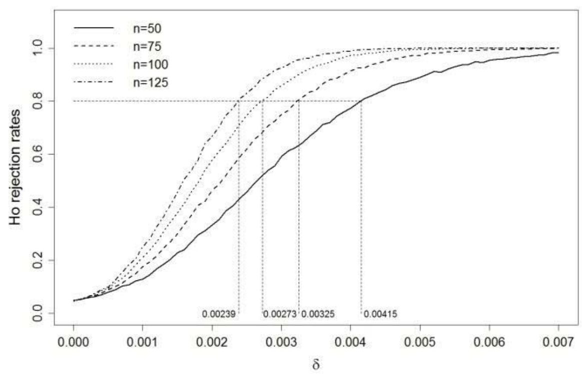

Figure 1b. H_0 rejection rates versus δ for Case 1

As expected, Figure 1a shows that the increase of δ lead to a reduction of the instrument performances and that P/T_{1m} and P/T_{2m} react differently to worsening changes in the measurement instrument with respect to $\%R\&R_m$ and SNR_m. Examining the results we note that SNR_m assesses instrument as unacceptable for $\delta \geq 0.0033$, while the test concludes, with a power greater than 80%, that instrument at time T=t is worse than instrument at time T=0 for $\delta \geq 0.00325$, when the sample size is n=75, and for $\delta \geq 0.00415$ when n=50. Therefore, pointing out that $\%R\&R_m$ evaluates as inadequate instrument for $\delta \geq 0.0059$, we can conclude that in this case, for moderate sample sizes (n=50 and n=75), the performances of the test are among the outcomes of metrics SNR_m and $\%R\&R_m$.

Next, we consider the case where the eigenvalues at time T=t are proportional to those at time T=0: $\mathbf{D}_{et}=\delta\mathbf{D}_{e0}$. Note that this case is equivalent to consider the instrument precision at instant t proportional to the precision at time T=0, *i.e.* $\mathbf{\Sigma}_{et}=\delta\mathbf{\Sigma}_{e0}$, since we can write $\mathbf{U}'_{e0}\left(\mathbf{D}_{et}\right)\mathbf{U}_{e0}=\mathbf{U}'_{e0}\left(\delta\mathbf{D}_{e0}\right)\mathbf{U}_{e0}=\delta\mathbf{U}'_{e0}\mathbf{D}_{e0}\mathbf{U}_{e0}=\delta\mathbf{\Sigma}_{e0}=\mathbf{\Sigma}_{et}$. Also in this case, the eigenvectors do not change, and the eigenvalues are expressed as $\lambda_{eti}=\delta\lambda_{e0i}$, for i=1,2,…,m. We perform our study by considering scenarios where the worsening factor δ ranges from 1 to 10 with a step of 0.1 and for each value of δ we proceed as for case 1. The results are summarised in Figures 2a and 2b. In this case $\%R\&R_m$ assesses the measurement instrument as inadequate for $\delta \geq 7.4$. The test concludes that instrument at instant t is worse than instrument at instant 0, with a power greater than 80%, when $\delta \geq 7.2$, $\delta \geq 6.2$, and $\delta \geq 5.6$ for n=75, n=100 and n=125 respectively. Therefore, considering that using SNR_m the instrument is unacceptable for $\delta \geq 5.2$, we conclude that for sample sizes ranging from n=75 to n=125 the results of the test are among the outcomes of metrics SNR_m and $\%R\&R_m$. Finally, let us examine the situation where the variations involve also the eigenvectors that can be interpreted as the presence of a serious problem in the measurement instrument. We examine the case where the first two diagonal elements of elements of $\mathbf{\Sigma}_{et}$ increase by a factor δ, with δ ranging from 1 to 8 with a step of 0.1, while the other matrix elements are equal to the corresponding elements of $\mathbf{\Sigma}_{e0}$. Figure 3a shows the multivariate MSA metrics computed for each value of δ. The Monte Carlo experiment has been performed as in the previous cases and the results are displayed in Figure 3b.

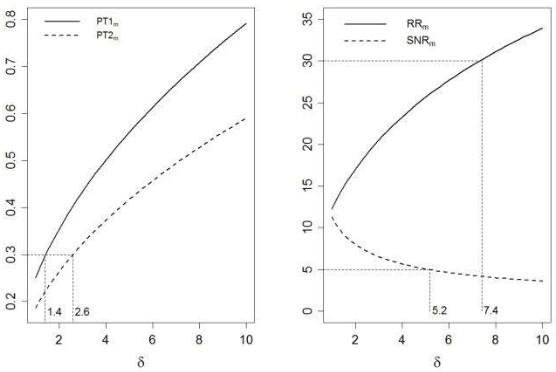

Figure 2a. Multivariate MSA-approval criteria as a function of δ for Case 2

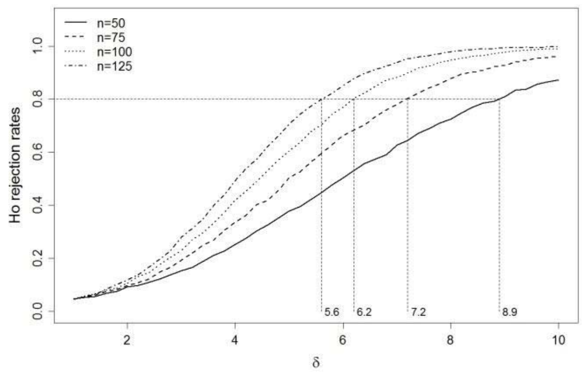

Figure 2b. H_0 rejection rates versus δ for Case 2

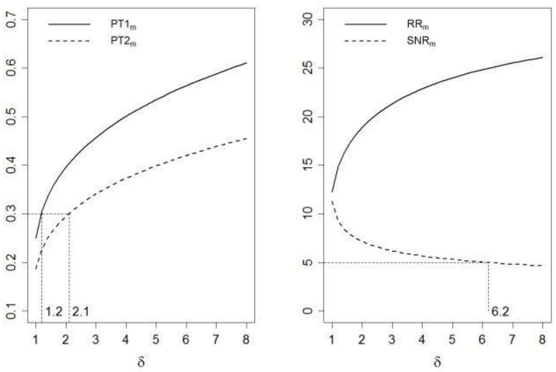

Figure 3a. Multivariate MSA-approval criteria as a function of δ for Case 3

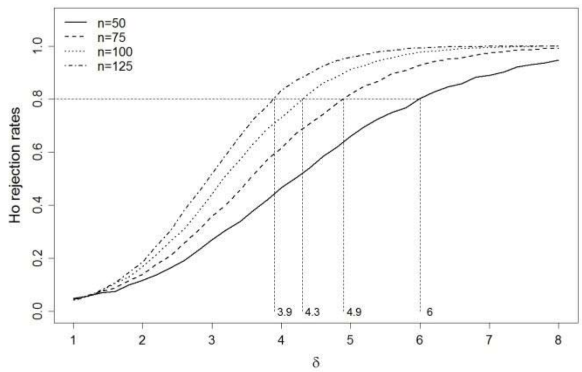

Figure 3b. H_0 rejection rates versus δ for Case 3

The results show that the test tends to be more sensitive to the increasing of δ than $\%R\&R_m$ and SNR_m. For instance, for n=50 the power of the test is greater than 80% for $\delta \geq 6$, while SNR_m evaluates the instrument as inadequate for $\delta \geq 6.2$. In this case the multivariate metric $\%R\&R_m$, although detects the worsening in the measurement instrument, it assesses the instrument as acceptable for all the values considered of δ. Summarizing, in the cases examined we aimed to study the performances of the test in realistic worsening scenarios of the measurement instrument after a period of use. The results show that the test provides outcomes with an appreciable level of agreement with the issues of $\%R\&R_m$ and SNR_m.

V. Concluding remarks

As any activity involving personnel, materials, tools and equipment, MSA usually requires a non-negligible financial support. Furthermore, the fact that these systems measure more than a single quality characteristic and that periodic assessments of measurement system performance are often required, engages manufacturers in important challenges. In this work, we have proposed a method which can be an additional tool for assessing the performances a multivariate measurement system. The method makes use of the data that are routinely available from the regular activity of the instrument and offers the possibility of assessing multivariate measurement systems without the necessity of performing a multivariate gauge study. Since the illustrated strategy can be implemented at almost no additional costs it may carried out more frequently than a MANOVA gauge study. Therefore, the synergic use of the proposed approach and the traditional multivariate gauge R&R studies can be: effective for reducing the costs of a multivariate MSA performed with a certain frequency; a useful strategy for improving the overall quality of multivariate measurement systems.

References

[1] Burdick, R.K., Borror, C.M., Montgomery, D.C. “A review of methods for measurements systems capability analysis”. *Journal of Quality Technology*, vol. 35, pp. 342-354, 2003.
[2] Automotive Industry Action Group (AIAG). *Measurement Systems Analysis*, 4th ed. Southfield, MI, 2010.
[3] Majeske, K.D. “Approval criteria for multivariate measurement systems”, *Journal of Quality Technology*, vol. 40, pp. 140-153, 2008.
[4] Danila, O., Steiner S.H., MacKay R.J. “Assessment of a Binary Measurement System in Current Use”. *Journal of Quality Technology*, vol. 42, pp.152-164, 2010.
[5] Johnstone, I.M., Ma, Z., Perry, P.O. Shahram, M. “Package RMTstat: Distributions, Statistics and Tests derived from Random Matrix Theory”. *http://cran.rproject. org/web/packages/RMTstat/*, (2009).