

A method for assessing multivariate measurement systems

Michele Scagliarini

Department of Statistical Sciences, University of Bologna, Via Belle Arti 41, 40126 Bologna, Italy

ABSTRACT

Multivariate measurement systems analysis is usually performed by designing suitable gauge R&R experiments ignoring available data generated by the measurement system while used for inspection or process control. This article 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 it 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 a multivariate MSA performed with a certain frequency.

Section: RESEARCH PAPER

Keywords: Covariance Matrices; Eigenvalues; Gauge Study; MANOVA; Wishart Distribution

Citation: Michele Scagliarini, A method for assessing multivariate measurement systems, Acta IMEKO, vol. 3, no. 4, article 8, December 2014, identifier: IMEKO-ACTA-03 (2014)-04-08

Editor: Paolo Carbone, University of Perugia

Received December 16th, 2013; **In final form** December 16th, 2013; **Published** December 2014

Copyright: © 2014 IMEKO. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Funding: This work was supported by grant from University of Bologna

Corresponding author: Michele Scagliarini, e-mail: michele.scagliarini@unibo.it

1. 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] and [2] 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 [3] and [4].

In current manufacturing industry processes are often characterized by many important characteristics. Accordingly, [5] proposed multivariate extensions of three commonly used MSA-approval criteria using the volume of constant-density contours to characterize the variability of the measurement system. These multivariate MSA-metrics require a multivariate analysis of variance (MANOVA) for 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 by 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., [6]). Here we propose the following approach. In the initial set up, after the multivariate measurement instrument is assessed 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 a multivariate gauge study, our proposal can be a useful tool for reducing the costs of multivariate MSA carried out with a certain frequency.

Here is the outline of the paper. The next section introduces the multivariate measurement error model,

describes the multivariate MSA-approval criteria proposed in recent literature and explains the multivariate analysis of variance (MANOVA) method for estimating the covariance matrices of interest. Section 3 develops the test for assessing the multivariate measurement instruments. Section 4 studies the performances of the proposed method. Finally, the last section contains a discussion and the conclusions.

2. MULTIVARIATE MEASUREMENT SYSTEM ANALYSIS

Let

$$\mathbf{X}' = [X_1, X_2, \dots, X_m] \quad (1)$$

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

$$\mathbf{LSL}' = [LSL_1, LSL_2, \dots, LSL_m], \quad (2)$$

$$\mathbf{USL}' = [USL_1, USL_2, \dots, USL_m], \quad (3)$$

and

$$\mathbf{T}' = [T_1, T_2, \dots, T_m] \quad (4)$$

the m -vectors values of the lower specification limits, upper specification limits and target values, respectively.

MSA methodology assumes the model

$$\mathbf{Y} = \mathbf{X} + \mathbf{e} \quad (5)$$

where \mathbf{Y} is the vector of the observable quality characteristics, which is usually obtained from some physical measurements, \mathbf{X} is the true quality characteristics vector and \mathbf{e} is the multivariate measurement error vector. It is assumed that

$$\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e) \quad (6)$$

with $\boldsymbol{\Sigma}_e$ positive definite and that \mathbf{X} and \mathbf{e} are independent. As a result,

$$\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_y) \quad (7)$$

where $\boldsymbol{\Sigma}_y = \boldsymbol{\Sigma}_e + \boldsymbol{\Sigma}$.

Let us denote with λ_i , λ_{ei} and λ_{yi} ($i=1,2,\dots,m$) the eigenvalues of $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_y$, respectively.

In the multivariate framework, [5] developed multivariate versions of three univariate gauge-approval criteria. The author proposed the following statistics.

The **multivariate precision-to-tolerance ratio**, which is defined as the m -th root of the ratio of $(1-\alpha)100\%$ volume of the multivariate error distribution and the volume of the tolerance region. This ratio, according to the specification of the tolerance region, simplifies to

$$\frac{P}{T_{1m}} = \frac{\left(\frac{\left(\prod_{i=1}^m \sqrt{\chi_{\alpha,m}^2 \lambda_{ei}} \right) \pi^{m/2}}{\left(\prod_{i=1}^m TOL_i \right) \Gamma(1+m/2)} \right)^{1/m}}{\quad} \quad (8)$$

when a hypercube-shaped tolerance region is used, and

$$\frac{P}{T_{2m}} = \left(\prod_{i=1}^m \frac{2\sqrt{\chi_{\alpha,m}^2 \lambda_{ei}}}{TOL_i} \right)^{1/m} \quad (9)$$

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 $\chi_{\alpha,m}^2$ is the $100(1-\alpha)$ -th percentile of the χ^2 distribution with m degrees of freedom with $(1-\alpha)$ usually fixed at 0.99. Therefore the P/T_{1m} and P/T_{2m} criteria compare the multivariate instrument variability, computed on the base of the constant-density contour ellipsoid, with the multivariate tolerance region (hypercube or hyperellipsoid).

The **multivariate percent R&R ratio**, which is defined taking the m -th root of the $(1-\alpha)100\%$ volumes of the gauge error distribution and measured-values distribution. The statistic in question simplifies in

$$\%R \& R_m = \left(\prod_{i=1}^m \sqrt{\frac{\lambda_{ei}}{\lambda_{yi}}} \right)^{1/m} 100 \quad (10)$$

and expresses the relative widths of the multivariate distributions of the error \mathbf{e} and the measured values \mathbf{Y} .

The third multivariate approval-metric is the **multivariate-signal-to-noise ratio** which compares the $(1-\alpha)100\%$ volume of the gauge-error distribution. This statistic is

$$SNR_m = \sqrt{2} \left(\prod_{i=1}^m \sqrt{\frac{\lambda_i}{\lambda_{ei}}} \right)^{1/m} \quad (11)$$

The author in [5] also gave the 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 $\leq 30\%$, while based on SNR_m a measurement system is adequate when $SNR_m \geq 5$.

2.1. Multivariate MSA in practice

The covariance matrices $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_e$ and $\boldsymbol{\Sigma}_y$ are usually unknown, for this reason [5] also proposes a multivariate analysis of variance (MANOVA) method of estimating the covariance matrices for one-factor and two-factor gauge studies.

According to the adopted notation, $\mathbf{Y}' = [Y_1, Y_2, \dots, Y_m]$ represents the measured values or data generated by the gauge. Let us consider a random-effects MANOVA model [7] where the factors in question are all m -dimensional vectors.

Let us denote the factors as \mathbf{p}_i ($i=1,2,\dots,p$) for part, \mathbf{o}_j ($j=1,2,\dots,o$) for operator, $\mathbf{p}\mathbf{o}_{ij}$ for part-operator interaction. The error term is denoted by $\boldsymbol{\varepsilon}_{ijk}$, where $k=1,2,\dots,r$ indicates the repeated reading of the same part by the same operator. Therefore \mathbf{Y}_{ijk} is an m -vector containing the k -th reading, by operator j , of the i -th part for the m quality characteristics.

In the two-factor gauge study the MANOVA model is:

$$\mathbf{Y}_{ijk} = \mathbf{p}_i + \mathbf{o}_j + \mathbf{p}\mathbf{o}_{jk} + \boldsymbol{\varepsilon}_{ijk} \quad (12)$$

where the random components \mathbf{p}_i , \mathbf{o}_j , $\mathbf{p}\mathbf{o}_{ij}$ and $\boldsymbol{\varepsilon}_{ijk}$ are mutually independent with distributions $\mathbf{p}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_p)$, $\mathbf{o}_j \sim N(\mathbf{0}, \boldsymbol{\Sigma}_o)$, $\mathbf{p}\mathbf{o}_{ij} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{po})$ and $\boldsymbol{\varepsilon}_{ijk} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$, respectively.

Within this framework,

$$MSP = \frac{or}{p-1} \sum_{i=1}^p (\bar{Y}_{i..} - \bar{Y}_{...}) (\bar{Y}_{i..} - \bar{Y}_{...})' \quad (13)$$

is the mean-square for the part matrix, where $\bar{Y}_{i..} = \frac{1}{or} \sum_{j=1}^o \sum_{k=1}^r Y_{ijk}$ and $\bar{Y}_{...} = \frac{1}{por} \sum_{i=1}^p \sum_{j=1}^o \sum_{k=1}^r Y_{ijk}$. The mean square for the operator matrix is

$$MSO = \frac{pr}{o-1} \sum_{j=1}^o (\bar{Y}_{.j.} - \bar{Y}_{...}) (\bar{Y}_{.j.} - \bar{Y}_{...})' \quad (14)$$

where $\bar{Y}_{.j.} = \frac{1}{pr} \sum_{i=1}^p \sum_{k=1}^r Y_{ijk}$. The mean square for the part-operator interaction matrix is given by

$$MSPO = \frac{r}{(o-1)(p-1)} \times \sum_{i=1}^p \sum_{j=1}^o (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}) (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})' \quad (15)$$

and the MSE matrix is

$$MSE = \frac{1}{po(r-1)} \sum_{i=1}^p \sum_{j=1}^o \sum_{k=1}^r (Y_{ijk} - \bar{Y}_{ijk}) (Y_{ijk} - \bar{Y}_{ijk})'. \quad (16)$$

The covariance matrices are estimated using the expected mean squares. The parts covariance matrix is estimated by

$$\hat{\Sigma}_p = \frac{MSP - MSPO}{or}, \quad (17)$$

the operator factor covariance matrix is estimated by

$$\hat{\Sigma}_o = \frac{MSO - MSPO}{pr}, \quad (18)$$

the part-operator interaction covariance matrix is estimated by

$$\hat{\Sigma}_{po} = \frac{MSPO - MSE}{r} \quad (19)$$

and the covariance matrix of the error terms ε_{ijk} is estimated by

$$\hat{\Sigma}_\varepsilon = MSE. \quad (20)$$

By adopting the gauge R&R notation, Σ^p corresponds to Σ , the covariance matrix of the quality characteristic. Repeatability and reproducibility are given by Σ_ε and $\Sigma_o + \Sigma_{po}$, respectively. The sum of repeatability and reproducibility gives the gauge measurement error covariance matrix

$$\Sigma_e = (\Sigma_o + \Sigma_{po}) + \Sigma_\varepsilon. \quad (21)$$

Therefore, the estimators of the covariance matrices of interest are:

$$\hat{\Sigma} = \hat{\Sigma}_p, \quad (22)$$

$$\hat{\Sigma}_e = (\hat{\Sigma}_o + \hat{\Sigma}_{po}) + \hat{\Sigma}_\varepsilon, \quad (23)$$

and

$$\hat{\Sigma}_y = \hat{\Sigma} + \hat{\Sigma}_e. \quad (24)$$

3. A TEST FOR MULTIVARIATE MEASUREMENT SYSTEMS

The multivariate MSA-approval criteria described in the previous section are based on constant-density contours of the multivariate normal distribution. A change in the

precision of a measurement instrument will cause a change in the corresponding ellipsoid of constant density. Therefore, since we are interested in the detection of a worsening in the measurement instrument precision, we will focus on significant reduction of the ellipsoid coverage.

Let us suppose that at the beginning of the manufacturing activity, which 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 Σ_{e0} the precision of the measurement instrument, with Σ_0 the covariance matrix of the true quality characteristic and with $\Sigma_{y0} = \Sigma_0 + \Sigma_{e0}$ the covariance matrix of the measurements, at time $T=0$.

Assuming the multivariate normality an ellipsoid of constant density is defined by

$$U_0 = \left\{ Y : (Y - \mu)' \Sigma_{y0}^{-1} (Y - \mu) = C \right\}. \quad (25)$$

If we assume $C = \chi^2_{\alpha, m}$, then U_0 is the boundary of the multivariate region in which 100 (1- α)% process fall.

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 Σ_{et} . Usually, the process variability is monitored by a suitable control chart thus if no out of control signals occur we can assume the stability of the process *i.e.* $\Sigma_0 = \Sigma_t$. Under these assumptions the variability of Y is

$$\Sigma_{yt} = \Sigma_0 + \Sigma_{et}. \quad (26)$$

In this framework, differences in the variability of the observed measures are only caused by changes in the precision since $\Sigma_{yt} = \Sigma_{y0}$ if and only if $\Sigma_{et} = \Sigma_{e0}$.

At time $T=t$, the $Y \in U_0$ define the quadratic form

$$Q_t(Y) = (Y - \mu)' \Sigma_{yt}^{-1} (Y - \mu) = (Y - \mu)' (\Sigma_0 + \Sigma_{et})^{-1} (Y - \mu). \quad (27)$$

$Q_t(Y)$ under the normality assumption is distributed as a χ^2 with m degrees of freedom and has several useful properties. $Q_t(Y)$ is not constant, $Q_t(Y) \leq C$ with equality only when $\Sigma_{yt} = \Sigma_{y0}$, *i.e.* $\Sigma_{et} = \Sigma_{e0}$. The minimum and the maximum values of $Q_t(Y)$ can be determined analytically.

Using results from [8] we find

$$\max(Q_t(Y)) = [\text{maximum eigenvalue of } \Gamma] \times C \quad (28)$$

$$\min(Q_t(Y)) = [\text{minimum eigenvalue of } \Gamma] \times C \quad (29)$$

where $\Gamma = \Sigma_{y0} \Sigma_{yt}^{-1}$.

Therefore, if at time $T=t$ the measurement instrument is worse than at time $T=0$, then the $Y \in U_0$ define ellipsoids with coverage ranging from

$$P_{\min} = \Pr(\chi_m^2 \leq \min(Q_t(Y))) \quad (30)$$

to

$$P_{\max} = \Pr(\chi_m^2 \leq \max(Q_t(\mathbf{Y}))). \quad (31)$$

The difference between $(1-\alpha)$, the ellipsoid coverage at time $T=0$, and P_{\min} (or P_{\max}), the ellipsoid coverage at time $T=t$, quantifies the possible worsening in the instrument precision at instant t .

A test for detecting the decreasing of the coverage can be derived as shown below.

Let us consider the null hypothesis that the instrument precision at instant $T=t$ is equal to the precision at instant $T=0$

$$H_0 : \Sigma_{et} = \Sigma_{e0} \quad (32)$$

If H_0 holds, then $\Sigma_{yt} = \Sigma_{y0}$, $\Gamma = \mathbf{I}$, $Q_t(\mathbf{Y}) = C$ and therefore the ellipsoid coverage is $P_{\min} = 1 - \alpha$.

Let the alternative hypothesis be

$$H_1 : \Sigma_{et} - \Sigma_{e0} \text{ is positive definite} \quad (33)$$

If H_1 holds, then $Q_t(\mathbf{Y}) \leq C$ hence the ellipsoid coverage is smaller than $1 - \alpha$: $P_{\min} < 1 - \alpha$.

Let

$$\lambda_{T1} \geq \lambda_{T2} \geq \dots \geq \lambda_{Tm} \quad (34)$$

be the eigenvalues of the matrix Γ . Under hypothesis H_1 we have

$$\lambda_{Tm} \leq \lambda_{T1} \leq 1 \quad (35)$$

and the (upper) limit 1 is reached only under H_0 .

Let $\lambda_{\Omega 1}$ be the largest eigenvalue of the matrix

$$\Omega = \Gamma^{-1} = \Sigma_{yt} \Sigma_{y0}^{-1}. \quad (36)$$

Then, when H_1 holds

$$\lambda_{\Omega 1} = \lambda_{Tm}^{-1} \geq 1 \quad (37)$$

with $\lambda_{\Omega 1} = 1$ only under H_0 .

Let \mathbf{S} be the sample covariance matrix of a random sample of size n from \mathbf{Y} at time t . If H_0 holds, then

$$(n-1)\Sigma_{y0}^{-1/2} \mathbf{S} \Sigma_{y0}^{-1/2} \sim W(\mathbf{I}, n-1) \quad (38)$$

where $W(\mathbf{I}, n-1)$ denotes a Wishart distribution with parameters \mathbf{I} and $n-1$. From [8] we have that matrix $\Sigma_{y0}^{-1/2} \mathbf{S} \Sigma_{y0}^{-1/2}$ and matrix $\Sigma_{y0}^{-1} \mathbf{S}$, have the same eigenvalues.

Furthermore, also matrices $\Sigma_{y0}^{-1} \mathbf{S}$ and $\mathbf{S} \Sigma_{y0}^{-1}$ have the same eigenvalues. Therefore, if we denote with $\hat{\lambda}_{\Omega 1}$ the largest eigenvalue of the matrix $\mathbf{S} \Sigma_{y0}^{-1}$, 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 the measurement instrument is assessed by comparing its performance instant t with those at instant 0, without the necessity of performing a multivariate gauge study (MANOVA): the sample covariance matrix \mathbf{S} can be estimated using the data available by the routine use of the measurement device at no additional costs.

4. CASE STUDIES

In this Section we discuss the ability of the test for detecting worsening in the measurement instrument performances.

Before entering in 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 \mathbf{Y} ; SNR_m compares the width of the multivariate distribution of the true quality characteristics \mathbf{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 behaviour similar to those of $\%R \& R_m$ and SNR_m . For this reason we shall compare the outcomes of the test only with the MSA-metrics $\%R \& R_m$ and SNR_m .

We consider as the situation at time $T=0$ (the benchmark) the case discussed by [5], 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 2 and we design suitable simulation experiments for studying the performances of the proposed test.

Let us therefore consider the case discussed by [5] where the data come from an automotive body panel gauge-study involving $m=4$ quality characteristics, with $p=5$ parts, $o=2$ operators and $r=3$ repeated measurements (see Table 1 in [5] for the original data). Using a two-factor MANOVA method the matrices estimates are

$$\hat{\Sigma} = \begin{bmatrix} 0.01811 & 0.01600 & -0.02180 & -0.00763 \\ 0.01600 & 0.25163 & -0.15732 & 0.35463 \\ -0.02180 & -0.15732 & 0.20856 & -0.39249 \\ -0.00763 & 0.35463 & -0.39249 & 0.98631 \end{bmatrix}$$

$$\hat{\Sigma}_e = \begin{bmatrix} 0.00094 & 0.00168 & -0.00141 & 0.00189 \\ 0.00168 & 0.00632 & -0.00475 & 0.00702 \\ -0.00141 & -0.00475 & 0.00486 & -0.00581 \\ 0.00189 & 0.00702 & -0.00581 & 0.00852 \end{bmatrix}$$

and

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

The eigenvalues of the covariance matrices $\hat{\Sigma}$, $\hat{\Sigma}_e$ and $\hat{\Sigma}_y$ are reported in Table 1. Using equations (10) and (11) we obtain $\%R \& R_m = 12.26061$ and $SNR_m = 11.30385$ respectively. The results show that the measurement instrument is assessed as acceptable by both the multivariate

Table 1. Eigenvalues of the estimated covariance matrices.

$\hat{\Sigma}$			
λ_1	λ_2	λ_3	λ_4
1.29428	0.11185	0.05438	0.00410
$\hat{\Sigma}_e$			
λ_{e1}	λ_{e2}	λ_{e3}	λ_{e4}
0.01908	0.00081	0.00050	0.00025
$\hat{\Sigma}_y$			
λ_{y1}	λ_{y2}	λ_{y3}	λ_{y4}
1.311189	0.11392	0.05557	0.00457

criteria, therefore it can be used in the manufacturing process and for our purposes we can assume $\hat{\Sigma}_e = \Sigma_{e0}$, $\hat{\Sigma} = \Sigma_0$ and $\hat{\Sigma}_y = \Sigma_{y0}$.

Now we examine several scenarios where a realistic worsening of the measurement instrument after a period of use is considered. We base this discussion on the spectral decomposition of Σ_{e0}

$$\Sigma_{e0} = U_{e0} D_{e0} U_{e0}' \quad (39)$$

where $U_{e0} = (\mathbf{u}_{e01}, \mathbf{u}_{e02}, \dots, \mathbf{u}_{e0m})$ is the matrix of eigenvectors with columns \mathbf{u}_{e0i} ($i=1,2,\dots,m$), and $D_{e0} = \text{diag}(\lambda_{e01}, \lambda_{e02}, \dots, \lambda_{e0m})$ is the diagonal matrix of the eigenvalues. The diagonal matrix D_{e0} is the covariance matrix of the latent independent factors that represent the primary independent sources of variability introduced by the instrument at time $T=0$.

The instrument after a period of use is characterised by a measurement error covariance matrix:

$$\Sigma_{et} = U_{et} D_{et} U_{et}' \quad (40)$$

where $U_{et} = (\mathbf{u}_{et1}, \mathbf{u}_{et2}, \dots, \mathbf{u}_{etm})$ with columns \mathbf{u}_{eti} ($i=1,2,\dots,m$) and $D_{et} = \text{diag}(\lambda_{et1}, \lambda_{et2}, \dots, \lambda_{etm})$.

Many alternative cases for Σ_{et} worse than Σ_{e0} can be considered, however Σ_{et} cannot be attained by changing the elements of Σ_{e0} arbitrarily. It is realistic to examine for Σ_{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 Σ_{et} with practical meaning should be those involving the eigenvalues, $\lambda_{eti} \geq \lambda_{ei0}$ for ($i=1,2,\dots,m$), but maintaining unchanged the eigenvectors ($U_{e0} = U_{et}$). A change in the eigenvectors can be interpreted as the presence of serious problems in the instrument such that the independent sources of variability become dependent. It is worth noting that this concept has been used also by other authors. For instance [9] used this definition of plausible changes in a process capability analysis framework.

In what follows, we consider three cases for Σ_{et} where the eigenvectors remain unchanged. Furthermore, for the sake of completeness, we will also consider the case of a change in the eigenvectors.

4.1. Case 1

Now we examine the case where the eigenvalues of Σ_{e0} are increased by the same additive term δ :

$$D_{et} = D_{e0} + \delta I. \quad (41)$$

Note that this case is equivalent to add the diagonal matrix δI directly to Σ_{e0}

$$\Sigma_{et} = \Sigma_{e0} + \delta I, \quad (42)$$

since from the spectral decomposition of Σ_{e0} we get the expression

$$\begin{aligned} U_{e0}'(\Sigma_{e0} + \delta I)U_{e0} &= U_{e0}'\Sigma_{e0}U_{e0} + \delta U_{e0}'U_{e0} = \\ &= D_{e0} + \delta I = D_{et} \end{aligned} \quad (43)$$

Thus, the eigenvectors remain unchanged, $U_{e0} = U_{et}$, and the eigenvalues can be expressed as $\lambda_{eti} = \lambda_{e0i} + \delta$ for $i=1,2,\dots,m$. Within this case we consider scenarios where the worsening term δ ranges from 0 to 0.006 with a step of 0.0001. Therefore, for each value of δ : a) we compute the multivariate MSA approval criteria $\%R$ & R_m and SNR_m and the results are shown in Figure 1; b) using the R -software we generate 10^4 samples ($n=50,75,100,150$) from $Y \sim N(\mu, \Sigma_0 + \Sigma_{et})$ where $\Sigma_{et} = \Sigma_{e0} + \delta I$. For each sample we compute the statistic $(n-1)\hat{\lambda}_{\Delta 1}$, where $\hat{\lambda}_{\Delta 1}$ is the largest eigenvalue of the matrix $S\Sigma_{y0}^{-1}$ 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 function $q\text{WisbartMax}$ of the R -package **RMTstat** [10] for computing the critical values of the test. The simulation results are summarized in Figure 2 where, for each sample size, the rejection rates of hypothesis H_0 as a function of δ are reported.

Examining the results we note that SNR_m assesses the 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 the 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 the instrument as inadequate 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 .

4.2. Case 2

Next, we consider the case where the eigenvalues at time $T=t$ are proportional to those at time $T=0$

$$D_{et} = \delta D_{e0}. \quad (44)$$

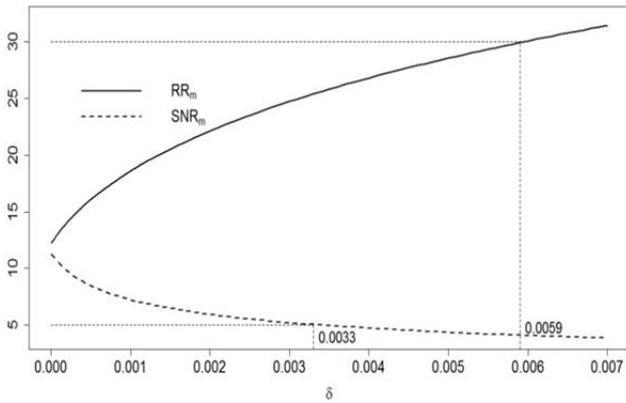


Figure 1. Multivariate MSA-approval criteria as a function of δ for Case 1.

Note that (44) is equivalent to consider the measurement error covariance matrix of instrument at instant t to be proportional to the instrument precision at instant 0

$$\Sigma_{et} = \delta \Sigma_{e0} \quad (45)$$

since we can write

$$U_{e0} (\delta D_{e1}) U_{e0}' = \delta (U_{e0} D_{e0} U_{e0}') = \delta \Sigma_{e0} = \Sigma_{et}. \quad (46)$$

Also in this case, the eigenvectors do not change, $U_{e0} = U_{et}$, and the eigenvalues are expressed as $\lambda_{eti} = \delta \lambda_{e0i}$, for $i=1,2,\dots,m$. We perform our study by considering scenarios where the worsening factor δ ranges from 1 to 10 with a step of 0.1.

For each value of δ we proceed as for case 1: a) we compute the multivariate MSA approval metrics, the results are displayed in Figure 3; b) we generate for each sample size ($n=50, 75, 100, 125$), 10^4 samples from $Y \sim N(\mu, \Sigma_0 + \Sigma_{et})$ where $\Sigma_{et} = \delta \Sigma_{e0}$. For each sample we compute the test statistic $(n-1)\hat{\lambda}_{\Delta 1}$ and, fixed as before an α -level of 5%, we evaluate the test power by computing the proportion of the 10^4 replications where H_0 is rejected in favour of H_1 . The rejection rates for hypothesis H_0 as a function of δ are shown in Figure 4.

In this case $\%R \& R_m$ assesses the measurement instrument as inadequate for $\delta \geq 7.4$. The test concludes that the instrument at instant t is worse than the instrument

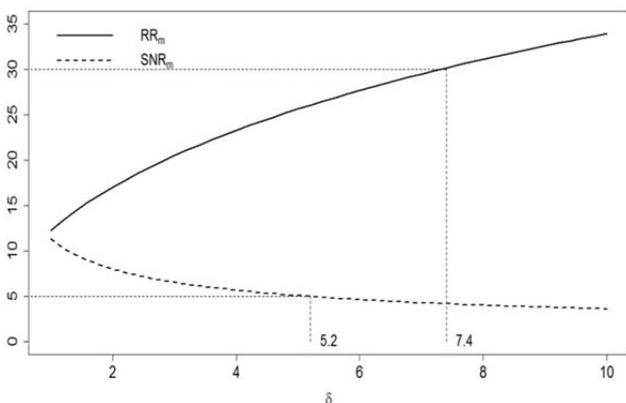


Figure 3. Multivariate MSA-approval criteria as a function of δ for Case 2.

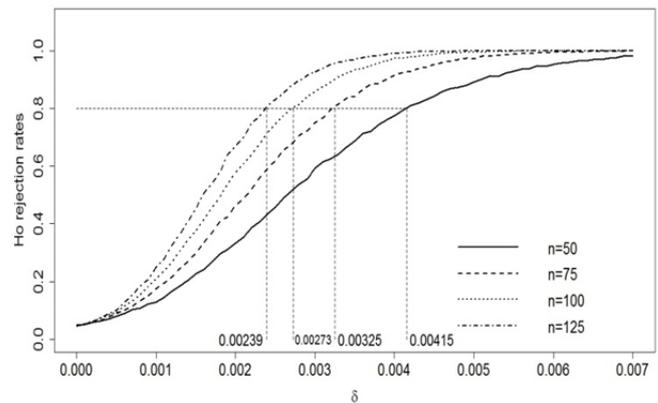


Figure 2. H_0 rejection rates versus δ for Case 1.

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 $\%R \& R_m$.

4.3. Case 3

Since in the previous cases we examined cases where all the eigenvalues simultaneously change, now we discuss the case where only several eigenvalues change their values. Let us consider the scenario where three eigenvalues are increased by a factor δ ranging from 1 to 20 with a step of 0.25:

$$D_{et} = \text{diag}(\lambda_{e01}, \delta \lambda_{e02}, \delta \lambda_{e03}, \delta \lambda_{e04}) \quad (47)$$

and the covariance matrix Σ_{et} is given by $\Sigma_{et} = U_{e0} D_{et} U_{e0}'$. For each value of δ we have followed the same procedure as for Cases 1 and 2: the results are summarized in Figures 5 and 6.

The pattern of the results is similar to that observed for Case 1. SNR_m assesses the instrument as unacceptable for $\delta \geq 9$ and $\%R \& R_m$ for $\delta \geq 16$, while the test concludes, that the instrument at time $T=t$ is worse than the instrument at time 0, with a power greater than 80%, when $\delta \geq 8$ and $\delta \geq 10$ for $n=75$ and $n=50$ respectively.

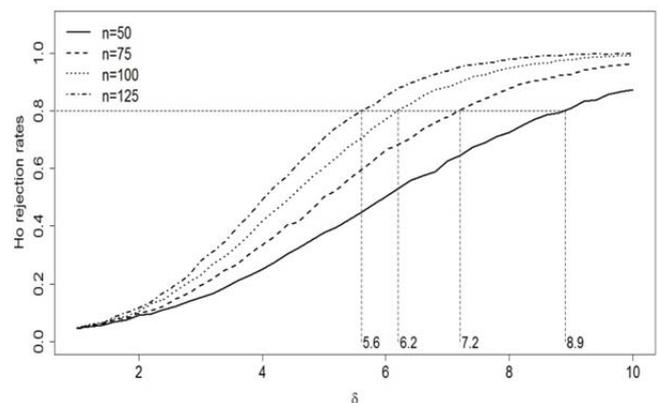


Figure 4. H_0 rejection rates versus δ for Case 2.

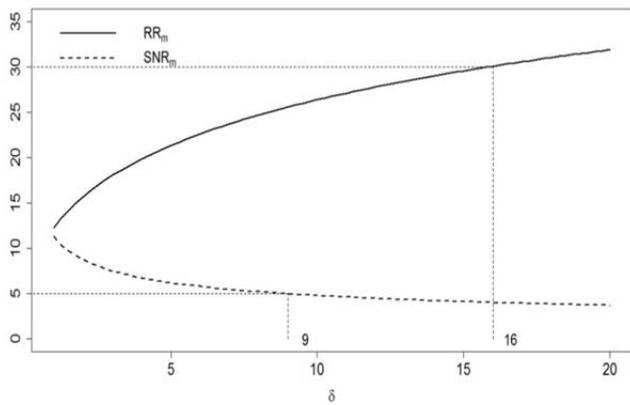


Figure 5. Multivariate MSA-approval criteria as a function of δ for Case 3.

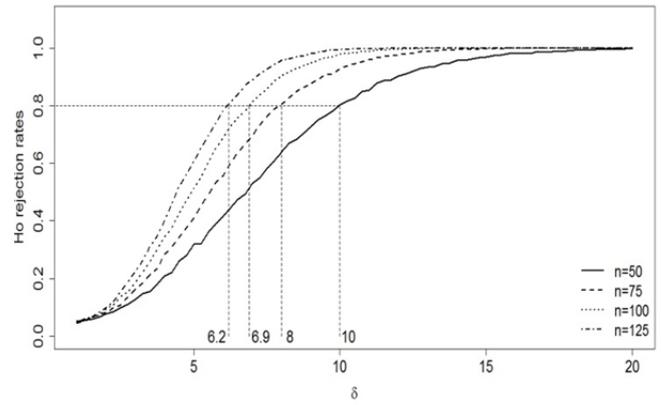


Figure 6. H_0 rejection rates versus δ for Case 3.

4.4. Case 4

Finally, let us examine the situation where the variations involve also the eigenvectors, which 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 Σ_{et} increase, while the other matrix elements are equal to the corresponding elements of Σ_{e0} :

$$\Sigma_{et} = \begin{bmatrix} \delta 0.00094 & 0.00168 & -0.00141 & 0.00189 \\ 0.00168 & \delta 0.00632 & -0.00475 & 0.00702 \\ -0.00141 & -0.00475 & 0.00486 & -0.00581 \\ 0.00189 & 0.00702 & -0.00581 & 0.00852 \end{bmatrix} \quad (48)$$

In the analysis the term δ ranges from 1 to 8 with a step of 0.1. Figure 7 shows the multivariate MSA metrics computed for each values of δ . The Monte Carlo experiment has been performed as in the previous cases and the results are summarized in Figure 8.

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 $\delta \geq 6.2$. Note that in this case, although the multivariate metric $\%R \& R_m$ detects the worsening in the measurement instrument, it assesses the instrument as acceptable for all the values considered of δ .

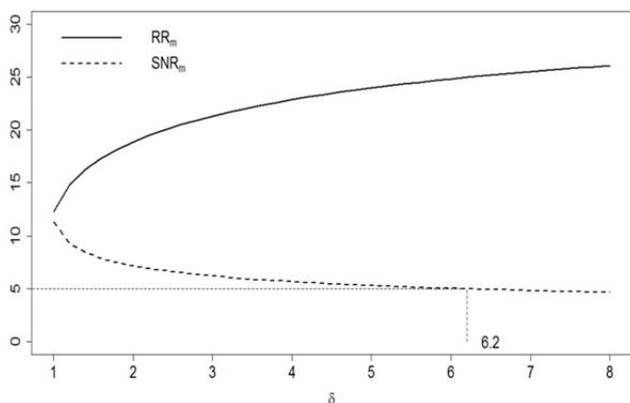


Figure 7. Multivariate MSA-approval criteria as a function of δ for Case 4.

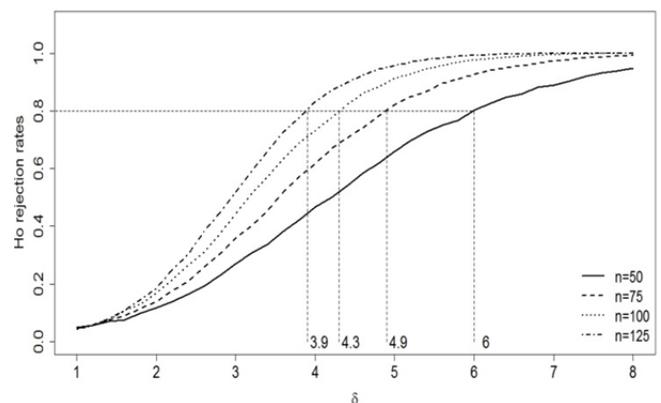


Figure 8. H_0 rejection rates versus δ for Case 4.

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 .

5. DISCUSSION AND CONCLUSIONS

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 statistical properties of 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 (MANOVA).

Since the illustrated strategy can be implemented at almost no additional costs it may be 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.

ACKNOWLEDGEMENT

The author would like to thank the two anonymous reviewers for their helpful comments.

REFERENCES

- [1] R.K. Burdick, C.M. Borror, D.C. Montgomery, A review of methods for measurements systems capability analysis, *J. Qual. Technol.* 35 (2003) pp. 342-354.
- [2] R.K. Burdick, C.M. Borror, D.C. Montgomery, *Design and Analysis of Gauge R&R Studies: Making Decisions with Confidence Intervals in Random Mixed ANOVA Models*, ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, (2005), ISBN 0-89871-588-1.
- [3] Automotive Industry Action Group (AIAG), *Measurement Systems Analysis*, 4th ed. Southfield, MI, (2010), ISBN 9781605342115
- [4] W.H. Woodall, C.M. Borror, Some relationships between gage R&R criteria, *Qual. Reliab. Eng. Int.*, 24 (2008) pp. 99-106.
- [5] K.D. Majeske, Approval criteria for multivariate measurement systems, *J. Qual. Technol.* 40 (2008) pp. 140-153.
- [6] O. Danila, S.H. Steiner, R.J. MacKay, Assessment of a Binary Measurement System in Current Use, *J. Qual. Technol.* 42 (2010) pp.152-164.
- [7] N.H. Timm, *Applied Multivariate Analysis*, Springer-Verlag, New York, 2002, ISBN 0-387-95347-7.
- [8] K.V Mardia, J.T. Kent, J.M. Bibby, *Multivariate Analysis*. Academic press, San Diego, 1979, ISBN 9780124712522.
- [9] I. González, I. Sánchez, Capability indices and nonconforming proportion in univariate and multivariate processes, *Int. J. Adv. Manuf. Tech.* 44 (2009) pp.1036-1050.
- [10] I.M., Johnstone, Z. Ma, P.O. Perry, M. Shahram, "Package RMTstat: Distributions, Statistics and Tests derived from Random Matrix Theory". (2009) <http://cran.rproject.org/web/packages/RMTstat/>