This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019031, IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

A Gage Study Through the Weighting of Latent Variables Under Orthogonal Rotation

Fabrício Alves de Almeida[1,2], Simone Carneiro Streitenberger[1], Alexandre Fonseca Torres[1], Anderson Paulo de Paiva[1], José Henrique de Freitas Gomes[1]

[1]Institute of Industrial Engineering and Management, Federal University of Itajubá, Itajubá, MG 37500-903, Brazil
[2]Departament of Economics, Faculty of Economic Sciences Southern Minas, Itajubá, Brazil

Corresponding author: Fabrício Alves de Almeida (e-mail: fabricio.alvesdealmeida@gmail.com)

This work was supported in part by FAPEMIG, in part by CNPq, in part by CAPES, and in part by FAPEPE.

ABSTRACT A new approach to identify and diagnose the quality of extensive and multivariate data is presented, using the gage repeatability and reproducibility (GR&R) study through the weighting of rotated factor scores. The proposal uses axis rotation to improve the explanation and interpretations of latent information, providing a statistically appropriate alternative when dealing with two or more correlated data sets. To analyze data with a significant variance-covariance structure, factor analysis (FA) is applied for calculating the eigenvalues and extracting of the rotated scores. Once obtained, these scores are then weighted with their respective eigenvalue for each factor. This procedure results in a single response vector, which is capable of properly interpreting all of the quality responses analyzed. To illustrate an application of the method, a real data set from a resistance spot welding process is selected, and two different types of rotation are compared. The proposed method provided an output that contemplated all of the significant variability of the data in a unique and significant way. In addition, the method enabled a reduction in the data dimensionality, thus minimizing the time for analysis and computational effort.

INDEX TERMS Multivariate measurement system, repeatability and reproducibility, orthogonal rotation, weighted factor analysis, resistance spot welding.

I. INTRODUCTION

Multivariate statistical techniques are widely used to analyze data that has a significant variance-covariance structure [1]. Such methods have been applied in many engineering problems to improve the interpretation of extensive and correlated data. In fact, several studies already use multivariate strategies in a handful of applications, such as flux-cored arc welding process [2], moving average control chart [3], design of experiments on clustering methods [4] and applications in process monitoring [5,6]. Such approaches are also used in the energy [7], healthcare [8] and economy [9] sectors. Among several methods, some of them stand out in view of their characteristics. The principal component analysis (PCA), for instance, is a multivariate strategy that reduces the data dimensionality and promotes uncorrelated vectors, considering its variance-covariance structure [10,11]. PCA has been used in several applications focused on quality improvement, such as the studies of [12–16].

Another widely used approach is the factor analysis (FA), which promotes the grouping of characteristics based on their explanation level [17]. FA has some advantages over the PCA technique. FA provides a better interpretation and explanation of the data with a simpler structure [1]. FA also enables the reduction of repetitive information between variables, using a smaller amount of latent variables [18]. Another advantage is that FA allows the grouping of the variables observed in relation to the factor loads. For example, in a suitable application, one factor would have a high factor load value, while the other factors would have small or moderate loads [1]. Such a characteristic would favor the simplicity of the structure and, consequently, the explanation of the data. However, this structure is not always obtained [17], so it is often recommended to use methods to rotate the axes of the factors to improve the explanation of the variables. The purpose of this rotation approach is to acquire a simple data structure, with easy interpretation of the observed variables [19].

The use of multivariate strategies in engineering problems is a modern practice, as it is in optimization methods [20,21]. However, searching to improve the process using optimization and other strategies may not bring enough results, since the variability is often attributed to the measurement process [22]. If this variability is not identified and properly diagnosed in the measurement process, this portion of variance can contaminate the decision-making process made based on the data, which may lead to results that do not correspond to the reality of the process. Among the techniques developed to analyze the measurement system, Woodal and Borror [23] highlight the gage repeatability and reproducibility study (GR&R) as the best option to analyze its capability. This technique allows the analysis of variability within each system and also between them, in addition to analyzing the consistency of the measurements of the operators with themselves [17].

When verifying the methods used in GR&R studies, Burdick et al. [24] state that the analysis of variance (ANOVA) method is the most used. However, when performing a statistical process control for correlated data, using univariate techniques, a type I error may occur. Industrial processes have multiple responses of interest and the ANOVA method promotes univariate analyzes, that is, one variable at a time. In addition to requiring a longer amount of time (depending on the number of responses), the ANOVA method neglects the variance-covariance structure of the data [17]. As an alternative, many authors have proposed the use of different multivariate methods in the GR&R study, such as the multivariate analysis of variance (MANOVA) [25], principal component analysis [26] and factor analysis [17,27]. Among the variations of these techniques, Almeida et al. [27] presented a combination of the factor analysis strategy and GR&R study, weighting the factor scores by the eigenvalue to improve the precision of a textured fiber bobbins measurement system, called weighted factor scores method (WF). However, the authors considered the factor scores without rotation, that is, without improving the explanation of the observed variables.

In order to contribute to GR&R strategies applied with factor analysis, this study presents a new approach based on rotated factor scores (*quartimax* and *varimax*), weighting by their respective eigenvalues. For this application, a data collection that follows the guidelines of a measurement system analysis is considered. Then, it should be analyzed whether the data are suitable for the application of FA. If data are suitable, some rotation method is then applied to improve the explanation of latent variables, i.e., simplifying its structure based on the principles of parsimony [28]. Finding a simpler structure, one should extract the rotated scores and calculate the eigenvalues for each factor. From this information, each factor is weighted by its respective eigenvalue, creating a unique response vector, capable of adequately representing the critical-to-quality characteristics (CTQ). Such procedure enables the estimation of the variation components and then the calculation of the multivariate indicators to evaluate the measurement system. This approach is called weighting of rotated factor scores (WRF). Based on this, the quality of the data will be evaluated in an appropriate way, given its structure of variance-covariance. In addition, this approach will promote a minimization in time and computational effort, due to the reduced dimensionality of the data and the evaluation. As a rotation strategy, the authors used the orthogonal rotation methods most applied in the literature, such as the *quartimax* and *varimax* method. To demonstrate the behavior of this proposal in real industrial processes, we will apply this approach in a resistance spot welding (RSW) process, analyzing the following geometric characteristics: indentation depth, penetration and nugget width. A study that performs the weighting of rotated factor scores applied to the GR&R study or any other application in industrial processes has not been found in the literature yet.

In general, the contributions of this paper can be summarized as follows:

1) A new proposal to verify the measurement system for extensive and correlated data is presented;
2) The use of orthogonal rotation methods promotes a better interpretation of latent variables, providing a simpler loading structure to assess data quality;
3) The weighting through the eigenvalues of each factor gives the corresponding degree of importance to each response cluster. In addition, the data set is properly represented by a single vector of responses;
4) The proposed method reduces the computational effort (representing several responses in a few factors), in addition to minimizing the analysis time (in which all CTQ's are represented in a single vector, which explains all responses appropriately).

The following section presents the theoretical background for the strategies used. In Section 3, a discussion of the new approach with rotated and weighted factor scores is presented. Section 4 describes the application and results in the RSW process. Section 5 presents the conclusions.

II. THEORETICAL BACKGROUND

A. GAGE REPEATABILITY AND REPRODUCTIBILITY

The GR&R method is an extensively used strategy to analyze the two precision components of a measurement system: repeatability and reproducibility. Repeatability refers to the variation resulting from the measurement device, while reproducibility indicates the variation generated by the measurement system [29]. GR&R can be treated as a particular case of the two-way analysis of variance with random effects [10]. Suppose factor **A** denotes a set of several parts while factor **B** indicates a certain number of operators that carry out the measurements. Since operators and parts can be selected from a large set of options, they can be considered random factors, and the model for this research can be observed in Eq. (1),

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1,2,..., a \\ j = 1,2,..., b \\ k = 1,2,..., n \end{cases} \tag{1}$$

where y expresses the response variable, μ represents the mean value and $\tau_i \sim N(0,\sigma_\tau)$, $\beta_j \sim N(0,\sigma_\beta)$, and $\tau\beta_{ij} \sim N(0,\sigma_{\tau\beta})$ represent, respectively, the random variable for each part, for the operator and for the interaction. Finally, $\varepsilon_{ijk} \sim N(0,\sigma_\varepsilon)$ indicates the estimated error term and a, b and n refer, in the following order, to the number of parts, operators, and replicas. The total variance σ_y^2 can be defined as per Eq. (2). The formula considers the independent normally distributed data with null mean and the variances $\sigma_\tau^2, \sigma_\beta^2, \sigma_{(\tau\beta)}^2$ indicating the variation components.

$$\sigma_y^2 = \sigma_\tau^2 + \sigma_\beta^2 + \sigma_{(\tau\beta)}^2 + \sigma_e^2 \tag{2}$$

Considering the random model for the analysis of variance, it is easy to verify that it is similar to the fixed models. It is important to highlight, however, that they differ regarding the nature of the random effects mean square. Failing to reject the null hypothesis indicates that there is not enough evidence of a significant variability between the population and the sample data, whereas the null hypothesis rejection evinces otherwise [30]. Adopting these concepts for the GR&R analysis allows the generalization of operators and sets of parts. Eq. (3) shows the sum of squares and the associated mean squares expected for random effects terms. This estimation leads to the components of variance. In a similar way, Table 1, Table 2 and Table 3 present the expressions required to perform the GR&R study.

$$y_{\bullet\bullet\bullet} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} = abn\mu + bn\sum_{i=1}^{a}\tau_i + an\sum_{j=1}^{b}\beta_j + n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\tau\beta_{ij}\right)_{ij} + \varepsilon_{\bullet\bullet\bullet} \tag{3}$$

TABLE 1. **Two-way ANOVA for random effects (Part I).**

Source	DF	SS	MS
(A)	$(a-1)$	$SS_A = \frac{1}{(bn)}\sum_{i=1}^{a} y_{i\bullet\bullet}^2 - \frac{y_{\bullet\bullet\bullet}^2}{(abn)}$	$MS_A = \frac{SS_A}{(a-1)}$
(B)	$(b-1)$	$SS_B = \frac{1}{(an)}\sum_{j=1}^{b} y_{\bullet j\bullet}^2 - \frac{y_{\bullet\bullet\bullet}^2}{(abn)}$	$MS_B = \frac{SS_B}{(b-1)}$
(AB)	$(a-1)(b-1)$	$SS_{AB} = SS_P - SS_A - SS_B$	$MS_{AB} = \frac{SS_{AB}}{df_{AB}}$
Error	$ab(n-1)$	$SS_E = SS_T - SS_A - SS_B - SS_{AB}$	$MS_E = \frac{SS_E}{ab(n-1)}$
Total	$(abn-1)$	$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{i=1}^{n} y_{ijk}^2 - \frac{(y_{\bullet\bullet\bullet})^2}{abn}$	

TABLE 2. **Two-way ANOVA for random effects (Part II).**

Source	Var Comp	F	P-Value
(A)	$\sigma_\tau^2 = \frac{MS_A - MS_{AB}}{bn}$	$F_{(A)} = \frac{MS_A}{MS_{AB}}$	$1 - F_{(A)}^{-1}\left[F_A;(a-1);ab(n-1)\right]$
(B)	$\sigma_\beta^2 = \frac{MS_B - MS_{AB}}{an}$	$F_{(B)} = \frac{MS_B}{MS_{AB}}$	$1 - F_{(B)}^{-1}\left[F_B;(b-1);ab(n-1)\right]$
(AB)	$\sigma_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n}$	$F_{(AB)} = \frac{MS_{AB}}{MS_E}$	$1 - F_{(AB)}^{-1}\left[F_{AB};df_{AB};ab(n-1)\right]$
Error	σ_e^2		
Total	σ_y^2		

Eq. (4) shows how to obtain the percentage of variation which allows to adequately evaluate and classify the measurement system. The number of distinct categories (ndc) identified by the measurement system is calculated using Eq. (5). Table 4 indicates the acceptance criteria of the established measurement system [31].

$$\%R\&R = \left(\frac{\sigma_{MS}}{\sigma_T}\right)100 \tag{4}$$

$$ndc = \sqrt{\frac{2\sigma_P^2}{\sigma_{MS}^2}} = 1.41\frac{\sigma_P}{\sigma_{MS}} \tag{5}$$

TABLE 3. Percentage of contribution and study variation for GR&R.

Source	Var Comp	% Contribution	%Study Var
Total Gage R&R	$\sigma_\beta^2+\sigma_{\tau\beta}^2+\sigma_\varepsilon^2$	$\frac{\left(\sigma_\beta^2+\sigma_{\tau\beta}^2+\sigma_\varepsilon^2\right)}{\sigma_\mathbf{T}^2}$	$\left(\frac{1}{\sigma_\mathbf{T}}\right)\sqrt{\left(\sigma_\beta^2+\sigma_{\tau\beta}^2+\sigma_\varepsilon^2\right)}$
Repeatability	σ_e^2	$\frac{\sigma_\varepsilon^2}{\sigma_\mathbf{T}^2}$	$\frac{\sigma_\varepsilon}{\sigma_\mathbf{T}}$
Reproductibility	$\sigma_\beta^2+\sigma_{\tau\beta}^2$	$\frac{\left(\sigma_\beta^2+\sigma_{\tau\beta}^2\right)}{\sigma_\mathbf{T}^2}$	$\left(\frac{1}{\sigma_\mathbf{T}}\right)\sqrt{\left(\sigma_\beta^2+\sigma_{\tau\beta}^2\right)}$
Operators	σ_β^2	$\frac{\sigma_\beta^2}{\sigma_\mathbf{T}^2}$	$\frac{\sigma_\beta}{\sigma_\mathbf{T}}$
Part-to-Part	σ_τ^2	$\frac{\sigma_\tau^2}{\sigma_\mathbf{T}^2}$	$\frac{\sigma_\tau}{\sigma_\mathbf{T}}$
Total Variation	$\sigma_\mathbf{T}^2=\sigma_\tau^2+\sigma_\beta^2+\sigma_{\tau\beta}^2+\sigma_\varepsilon^2$	100%	100%

Table 4. Classification criteria for the measurement system

Measurement System Assessment	*%GR&R*
Acceptable	< 10%
Marginal	10% to 30%
Unacceptable	> 30%
ndc	> 5

B. FACTOR ANALYSIS

Factor analysis (FA) is a multivariate statistical technique that describes the covariance relationships among the response variables (y_i, $i = 1, 2, \ldots, p$), gathering variables that are highly correlated in a same factor (f_j, $j = 1, 2, \ldots, m$) [1]. When $m < p$, these factors are unobservable variables, also known as latent variables or common factors. The FA model can be represented through a linear relationship, as shown in Eq. (6),

$$\mathbf{Y}-\boldsymbol{\mu}=\mathbf{LF}+\boldsymbol{\varepsilon} \tag{6}$$

where $\mathbf{Y}$ is an observable random vector with p response variables, $\boldsymbol{\mu}_{(p\times1)}$ is the vector of population means, $\mathbf{L}$ represents the matrix of factor loadings with dimension $p \times m$ (Eq. (7)), $\mathbf{F}_{(m\times1)}$ indicates the random vector containing the unobservable latent variables, and $\boldsymbol{\varepsilon}_{(p\times1)}$ is a random vector of additional sources of variation (errors), also known as vector-specific factors. The matrix $\mathbf{L}$ is composed of the factor loadings l_{ij} that comprises the correlation or the covariance between the response Y_i and the common factor f_j [18].

$$\mathbf{L}=\begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pm} \end{bmatrix} \tag{7}$$

Applying the FA requires the original response variables to be adequate [1]. Such adequacy can be evaluated using the Bartlett sphericity test. This test considers the test statistic $\chi^2_{\alpha;v}$ to verify whether the correlation matrix is an identity matrix, with α as the level of significance and a number of $v = p\,(p-1)/2$ degrees of freedom. This test also assumes that the dataset $\mathbf{Y} = [Y_1, Y_2, ..., Y_p]^\mathrm{T}$ follows a normal multivariate distribution. In this sense, the null hypothesis that the correlation matrix is equal to the identity matrix is not rejected. In other words, the data is considered not correlated when $\chi^2 > \chi^2_{\alpha;[p(p-1)/2]}$, and the value of χ^2 is obtained from Eq. (8) [18].

$$\chi^2 = -\left[n - 1 - \frac{(2p+5)}{6}\right]\ln|\mathbf{R}| \tag{8}$$

Another way to verify the adequacy of the data is through the KMO index as shown in Eq. (9), where r_{ij} are the sample correlation matrices **R** and q_{ij} are the anti-image correlation matrices **Q** [18]. Although it ranges from 0 to 1, values greater than 0.5 already indicates the suitability of the original responses set.

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2} \tag{9}$$

FA's mathematical emphasis is in expressing the population covariance $\mathbf{\Sigma}_{(p \times p)}$ through a matrix in terms of a specific variance matrix, where the terms of the main diagonal contains the errors, while the values outside it are null. The population parameters are not known, so the $\mathbf{\Sigma}$ matrix can be estimated by the sample covariance matrix $\mathbf{S}_{(p \times p)}$. Nevertheless, it is highly recommended to model the sample correlation matrix $\mathbf{R}_{(p \times p)}$ instead of **S**. Most problems involving multivariate analyses contain response variables with diverging scales, which makes the correlation more adequate since it is not sensitive to these discrepancies. The matrix **R** can be calculated by Eq. (10),

$$\mathbf{R} = \mathbf{L}\mathbf{L}^T + \mathbf{\Psi} \tag{10}$$

where **L** represents the matrix of factor loadings with dimension $p \times m$ and ψ_i is a part of the total variance of Y_i explained by the specific factor ε_i [18].

The Principal Component (PC) is one the most used methods to estimate the above-mentioned matrix **L**. It determines the factor loads and the specific variances through the spectral decomposition of **S** or **R** matrices [1]. Thereby, the matrix **L** can be calculated as presented in Eq. (11):

$$\mathbf{L} = \mathbf{P}_m \mathbf{\Lambda}_m^{1/2} = \left[\sqrt{\lambda_1}\mathbf{e}_1, \sqrt{\lambda_2}\mathbf{e}_2, \ldots, \sqrt{\lambda_m}\mathbf{e}_m\right] \tag{11}$$

where $\mathbf{P}_m$ represents the matrix $p \times m$ of the first m normalized eigenvectors (e_i) of **R**, and $\mathbf{\Lambda}_m$ indicates the diagonal matrix $m \times m$ of the eigenvalues (λ_i) of the same matrix **R**. Considering Eqs. (10) and (11) it is possible to obtain the specific variances, as shown in Eq. (12).

$$\mathbf{\Psi} = \operatorname{diag}\left(\mathbf{R} - \mathbf{L}\mathbf{L}^T\right) \tag{12}$$

The FA theory states that the number of factors m is necessarily less than p, which also allows the reduction of the problem dimension. The issue of determining how many factors should be used to represent the data set can be solved using several criteria. Nevertheless, the main requirement is that m must present a cumulative variation rate ≥ 80% and, in view of the sample correlation matrix, m must have eigenvalues greater than the mean eigenvalues, i.e., $\lambda_i \geq 1$ [18].

C. ORTHOGONAL ROTATION METHODS

The rotation of the factor loads is a widespread practice to deal with the difficulty on factor loads interpretation, since it facilitates the association of the common factors to the response variables utilizing a simpler load structure [19]. The rotated factor load matrix $\mathbf{L}^{\circ}$ can reproduce either **S** or **R**, while maintains the estimation of the communalities and specific variances, since $\mathbf{L}^{\circ} = \mathbf{LT}$, and **T** is an orthogonal matrix for rotating **L** [34]. Among the rotation methods, the most used are: *quartimax* and *varimax* method.

An approach commonly used for rotating the axes is the *quartimax* method. The *quartimax* approach is characterized as a type of orthogonal rotation that aims to simplify the columns of a factor matrix [32], minimizing the cross-product term, according to Eq. (13) [33].

$$Quartimax = \sum_{i=1}^{p}\sum_{j=1}^{q} \tilde{l}_{ij}^{\circ 4} + \sum_{i=1}^{p} \sum\sum_{j \neq k}^{q} \tilde{l}_{ij}^{\circ 2}\tilde{l}_{ik}^{\circ 2} \tag{13}$$

However, some rotation methods may perform better than others, depending on the data structure. The *varimax* method selects an orthogonal matrix **T** to create rotational factor loads that promote maximization of the objective function indicated in Eq. (14), where $\tilde{l}_{ij}^{\circ} = l_{ij}^{\circ} / \sqrt{h_i^2}$, in other words, it represents the relation between the rotated factor load and the i^{th} commonality.

$$Varimax = \frac{1}{p}\sum_{j=1}^{m}\left[\sum_{i=1}^{p}\tilde{l}_{ij}^{\circ 4} - \left(\sum_{i=1}^{p}\tilde{l}_{ij}^{\circ 2}\right)^2 \Big/ p\right] \tag{14}$$

Since FA produces latent variables, it is usual to obtain estimated values for them (factor scores) to conduct further analysis. According to Johnson and Wichern [1], minimizing the sum of squared residuals of the factor model leads to the estimation of the common factors. Eq. (15) shows how the rotated factor scores are obtained.

$$\mathbf{F} = \mathbf{Z}\left[\mathbf{L}^{\circ}\left(\mathbf{L}^{\circ T}\mathbf{L}^{\circ}\right)^{-1}\right] \tag{15}$$

where $\mathbf{F}_{(n \times m)}$ is the matrix containing the estimation of the rotated latent variables **R**, $\mathbf{Z}_{(n \times p)}$ is the matrix of the standardized values of the response variables, and n is the number of observations in each response variable.

III. GR&R THROUGH THE WEIGHTING OF LATENT VARIABLES UNDER ORTHOGONAL ROTATION

In this study, the authors propose the combination of the axis rotation strategy with the GR&R measurement system analysis. After properly selecting the operators, collecting parts and measuring the CTQ (correctly defining a GR&R study), one should initially check whether the data is appropriate for the use of the FA strategy. Thus, the Bartlett sphericity test and the Kaiser-Meyer-Olkin (KMO) index are performed. If the data set is not adequate, another multivariate strategy, such as PCA or MANOVA, should be used. However, if the data are suitable, the proposed method can be continued. The number of factors to must be equal of higher than 2 to perform the rotation of the axes.

The ordinary least squares method (OLS) is usually adopted when factor scores are obtained using the principal components [1] and, in this approach, the goal is to minimize the sum of squares of the residuals of the factor model. More specifically, if $\mathbf{y}$-$\boldsymbol{\mu}$ = $\mathbf{LF}$ + $\boldsymbol{\varepsilon}$ is the definition of the factor model, then the residue vector is $\boldsymbol{\varepsilon}$ = $\mathbf{y}$-$\boldsymbol{\mu}$-$\mathbf{LF}$. Then, the minimization function is described in Eq. (16).

$$\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \left(\mathbf{y}-\boldsymbol{\mu}-\mathbf{LF}\right)^T\left(\mathbf{y}-\boldsymbol{\mu}-\mathbf{LF}\right) \tag{16}$$

Thus, before extracting the factor scores (Eq. (15)), axis rotation should be performed to improve data interpretation.

The rotated factor load ($\mathbf{L}^{\circ}$) maintains estimates of specific commonalities and variances, such as $\mathbf{L}^{\circ}$ = $\mathbf{LT}$, where $\mathbf{T}$ is an orthogonal matrix for rotating $\mathbf{L}$. In this sense, the orthogonal rotation metrics can be tried out, envisioning the method that best simplifies the data structure. Therefore, one can contemplate the use of methods consolidated in the literature, such as *quartimax* (Eq. (13)) and *varimax* (Eq. (14)).

Considering one of the rotation methods described previously, the scores of rotated factors should be extracted. Assuming that the first-order partial derivate of Eq. (16) related to the matrix of factor scores $\mathbf{F}$ are null, then $\mathbf{F}$ is estimated by Eq. (15). Then, it is possible to create a new WRF vector, which can be described as Eq. (17).

$$\mathbf{WRF} = \begin{bmatrix} \left(\dfrac{CTQ_{11}-\overline{CTQ_1}}{\sqrt{s_{11}}}\right) & \left(\dfrac{CTQ_{12}-\overline{CTQ_2}}{\sqrt{s_{22}}}\right) & \cdots & \left(\dfrac{CTQ_{1p}-\overline{CTQ_p}}{\sqrt{s_{pp}}}\right) \\ \left(\dfrac{CTQ_{21}-\overline{CQT_1}}{\sqrt{s_{11}}}\right) & \left(\dfrac{CTQ_{22}-\overline{CTQ_2}}{\sqrt{s_{22}}}\right) & \cdots & \left(\dfrac{CTQ_{2p}-\overline{CTQ_p}}{\sqrt{s_{pp}}}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\dfrac{CTQ_{n1}-\overline{CTQ_1}}{\sqrt{s_{11}}}\right) & \left(\dfrac{CTQ_{n2}-\overline{CTQ_2}}{\sqrt{s_{22}}}\right) & \cdots & \left(\dfrac{CTQ_{np}-\overline{CTQ_p}}{\sqrt{s_{pp}}}\right) \end{bmatrix} \times \tag{17}$$
$$\left\{ \begin{bmatrix} l_{11}^{\circ} & l_{12}^{\circ} & \cdots & l_{1m}^{\circ} \\ l_{21}^{\circ} & l_{22}^{\circ} & \cdots & l_{2m}^{\circ} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1}^{\circ} & l_{p2}^{\circ} & \cdots & l_{pm}^{\circ} \end{bmatrix} \times \left(\begin{bmatrix} l_{11}^{\circ} & l_{21}^{\circ} & \cdots & l_{p1}^{\circ} \\ l_{12}^{\circ} & l_{22}^{\circ} & \cdots & l_{p2}^{\circ} \\ \vdots & \vdots & \ddots & \vdots \\ l_{1m}^{\circ} & l_{2m}^{\circ} & \cdots & l_{pm}^{\circ} \end{bmatrix} \times \begin{bmatrix} l_{11}^{\circ} & l_{12}^{\circ} & \cdots & l_{1m}^{\circ} \\ l_{21}^{\circ} & l_{22}^{\circ} & \cdots & l_{2m}^{\circ} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1}^{\circ} & l_{p2}^{\circ} & \cdots & l_{pm}^{\circ} \end{bmatrix} \right)^{-1} \right\} \times \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}$$

It is possible to use the eigenvalues to establish the weights of their respective factors. The result is a univariate variable, as shown in Eq. (18).

$$\mathbf{WRF} = \sum_{i=1}^{m}\left[\lambda_i \mathbf{F}_i\right] = \lambda_1\mathbf{F_1} + \lambda_2\mathbf{F_2} + \ldots + \lambda_m\mathbf{F}_m \tag{18}$$

At this stage, the analysis of variance (ANOVA) for random effects can be applied to the weighted factor under rotation (WRF). Then, the variation components of the GR&R study are obtained once the factor scores are extracted. Eq. (19) presents the WRF model for the measurement system.

$$\mathbf{WRF}_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., b \\ k = 1, 2, ..., n \end{cases} \tag{19}$$

In Eq. (19), μ is the expected value for WRF and τ_i, β_j, $\tau\beta_{ij}$, and ε_{ijk} refer to random effects with null expected values and variances σ_{τ}^2, σ_{β}^2, $\sigma_{\tau\beta}^2$, and σ_{ε}^2 respectively. For scenarios in which the interaction is not significant, then WRF is estimated using Eq. (20).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019031, IEEE Access

$$\mathbf{WRF}_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk} \tag{20}$$

The variance components for the GR&R study are presented in Table 5, where MS_A, MS_B and MS_{AB} represent the mean squares for the part factor, operator factor, interaction term, respectively, and MS_E the mean square for the error term. Based on the previous analyses, the measurement system must be classified by the contributions to the percentage of variability (*%R&R*$_m$) and also the number of distinct categories (ndc_m). These indicators are described respectively in Eqs. (21) and (22). The evaluation criteria are the same as those described in Table 4, based on the AIAG [31].

TABLE 5. Variation components for the GR&Rm study.

			With interaction	Without interaction
Process	$\hat{\sigma}_P^2$	=	$\frac{MS_A - MS_{AB}}{bn}$	$\frac{MS_A - MS_E}{bn}$
Repeatability	$\hat{\sigma}_{repeat}^2$	=	MS_E	MS_E
Reproducibility	$\hat{\sigma}_{reprod}^2$	=	$\frac{MS_B - MS_{AB}}{an} + \frac{MS_{AB} - MS_E}{n}$	$\frac{MS_B - MS_E}{an}$
Measurement System	$\hat{\sigma}_{MS}^2$	=	$\hat{\sigma}_{repeat}^2 + \hat{\sigma}_{reprod}^2$	$\hat{\sigma}_{repeat}^2 + \hat{\sigma}_{reprod}^2$
Total Variation	$\hat{\sigma}_T^2$	=	$\hat{\sigma}_P^2 + \hat{\sigma}_{MS}^2$	$\hat{\sigma}_P^2 + \hat{\sigma}_{MS}^2$

$$\%R\&R_m = \sqrt{\frac{\sigma_{MS}^2}{\left(\sigma_P^2 + \sigma_{MS}^2\right)}} \times 100 \tag{21}$$

$$ndc_m = \sqrt{2\left[\frac{\sigma_\tau^2}{(\sigma_\beta^2 + \sigma_{\tau\beta}^2) + \sigma_\varepsilon^2}\right]} \tag{22}$$

To visually represent the method and, consequently, facilitate its understanding, Fig. 1 illustrates the flowchart of the proposed approach, contemplating the steps for applying the method. Analogously, Table 6 describes the pseudocode for implementing the gage study proposal through weighted of factors scores under orthogonal rotations.

IV. NUMERICAL EXAMPLE: A RSW PROCESS

In order to demonstrate the application of this improvement, the approach was applied in a resistance spot welding process, evaluating the following critical-to-quality characteristics: indentation depth (*ID*), penetration (*P*) and nugget width (*NW*). The planning was carried out from the design of experiments (DOE) strategy, specifically by a fractional factorial design indicated in [34]. Data collection followed appropriate planning for a GR&R study using eight parts (*a*), four different operators (*b*) and three replicates (*n*) for three distinct quality characteristics, totaling 288 measurement data. All measurements were performed randomly, adapted from [17] and available in Table 7. The ID measurement was performed on the upper face of the specimen, for it presents a higher indentation depth value, according to the specification suggested by Almeida et al. [22].

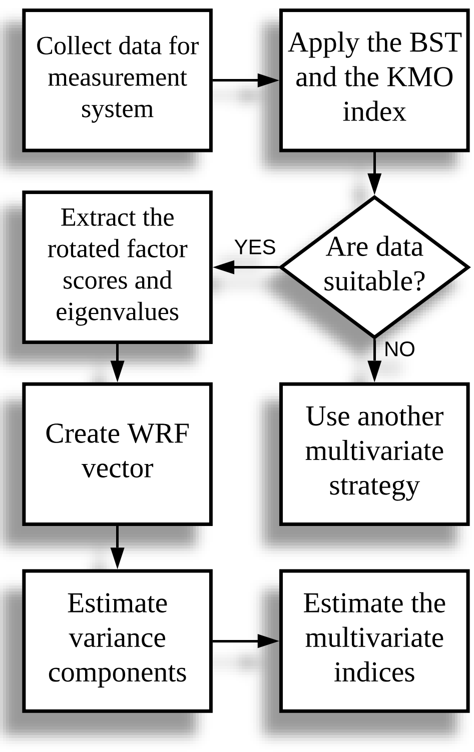


FIGURE 1. Detailed flowchart for conducting WRF approach

TABLE 6. **Pseudocode for implementation of WRF approach**

Pseudocode:

Input: Data for measurement system
Output: Multivariate indices estimations

$y_1, y_2, \ldots, y_p \leftarrow$ original response variables $(i = 1,2,\ldots,p)$
$f_1, f_2, \ldots, f_m \leftarrow$ latent variables $(j = 1,2,\ldots,m)$
$n \leftarrow$ sample size of the dataset
$R \leftarrow$ correlation matrix
BST $\leftarrow -\left[n - 1\frac{(2p+5)}{6}\right]\ln|R|$
$r_{ij} \leftarrow$ sample correlation matrices R
$q_{ij} \leftarrow$ anti-image correlation matrices Q
$u \leftarrow 0$
$w \leftarrow 0$
for $i = 1$ **to** p
    for $j = 1$ **to** m
        if $i \neq j$
            $u = u + r_{ij}^2$
            $w = w + q_{ij}^2$
        end if
    end for
end for
KMO $\leftarrow \frac{u}{u+w}$
if (BST, KMO) indicate suitable data
    $s_1, s_2, \ldots, s_m \leftarrow$ rotated factor scores
    $e \leftarrow$ eigenvalues vector
    $WRF \leftarrow \sum_{i=1}^{m}[\lambda_m \boldsymbol{F}_m]$
    $v \leftarrow$ variance components estimation
    $x \leftarrow$ multivariate indices estimation
    $RR_m \leftarrow \sqrt{\frac{\sigma_{MS}^2}{(\sigma_P^2+\sigma_{MS}^2)}} \times 100$
    if $RR_m < 10$
        $acceptRR \leftarrow 0$ //acceptable
        else if $RR_m > 30$
            $acceptRR \leftarrow 1$ //unacceptable
        else
            $acceptRR \leftarrow 2$ //marginal
        end if
    end if
    $NDC_m \leftarrow \sqrt{2 \times \left[\frac{\sigma_\tau^2}{\left(\sigma_\beta^2+\sigma_{\tau\beta}^2\right)+\sigma_\varepsilon^2}\right]}$
    if $ndc > 5$
        $acceptNDC \leftarrow 0$ //acceptable
    else
        $acceptNDC \leftarrow 1$ //unacceptable
    end if
end if

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019031, IEEE Access

TABLE 7. Measurements of CTQ's for the RSW process

n	*a*	*b* = *A*			*b* = *B*			*b* = *C*			*b* = *D*		
		ID	**P**	**NW**	**ID**	**P**	**NW**	**ID**	**P**	**NW**	**ID**	**P**	**NW**
1	1	0.191	0.988	4.295	0.202	0.978	4.324	0.201	0.949	4.344	0.192	0.969	4.384
2	1	0.202	0.978	4.274	0.191	0.978	4.325	0.192	0.958	4.354	0.191	0.988	4.354
3	1	0.192	0.958	4.353	0.202	0.968	4.294	0.191	0.947	4.354	0.192	0.986	4.384
1	2	0.211	1.116	4.654	0.202	1.139	4.744	0.212	1.117	4.734	0.212	1.119	4.724
2	2	0.212	1.108	4.724	0.212	1.138	4.725	0.222	1.117	4.725	0.202	1.137	4.844
3	2	0.212	1.127	4.784	0.221	1.127	4.754	0.222	1.116	4.734	0.212	1.126	4.854
1	3	0.131	1.026	3.534	0.133	1.057	3.605	0.122	1.018	3.645	0.122	1.049	3.674
2	3	0.121	1.018	3.664	0.122	1.057	3.615	0.121	1.018	3.635	0.131	1.058	3.674
3	3	0.123	1.018	3.644	0.122	1.048	3.565	0.132	1.019	3.644	0.132	1.046	3.655
1	4	0.069	0.977	3.544	0.071	0.988	3.575	0.069	0.969	3.534	0.068	0.968	3.525
2	4	0.072	0.987	3.544	0.069	0.987	3.546	0.069	0.967	3.564	0.072	0.978	3.574
3	4	0.070	0.967	3.564	0.071	0.956	3.534	0.071	0.958	3.514	0.072	0.967	3.574
1	5	0.252	1.189	4.585	0.262	1.176	4.554	0.262	1.188	4.614	0.272	1.198	4.534
2	5	0.252	1.200	4.514	0.261	1.198	4.544	0.251	1.197	4.644	0.271	1.199	4.524
3	5	0.261	1.209	4.524	0.271	1.178	4.564	0.261	1.207	4.585	0.252	1.198	4.624
1	6	0.202	1.098	3.805	0.191	1.097	3.795	0.202	1.089	3.844	0.202	1.108	3.844
2	6	0.192	1.098	3.784	0.192	1.107	3.804	0.201	1.099	3.833	0.202	1.117	3.875
3	6	0.202	1.087	3.795	0.201	1.088	3.825	0.202	1.079	3.834	0.193	1.108	3.874
1	7	0.169	1.159	4.794	0.182	1.147	4.824	0.168	1.147	4.835	0.171	1.168	4.814
2	7	0.172	1.167	4.864	0.172	1.179	4.805	0.173	1.137	4.814	0.164	1.167	4.925
3	7	0.172	1.159	4.794	0.171	1.137	4.814	0.171	1.139	4.825	0.171	1.158	4.844
1	8	0.142	1.139	4.074	0.141	1.158	4.145	0.141	1.127	4.184	0.138	1.128	4.124
2	8	0.143	1.138	4.165	0.141	1.149	4.124	0.142	1.137	4.164	0.142	1.158	4.195
3	8	0.142	1.128	4.085	0.142	1.138	4.134	0.139	1.138	4.094	0.143	1.149	4.184

All values were measured on the millimeter scale

A. METHOD WITH VARIMAX ROTATION

Given the measurement data for the GR&R study, the method described in section 3 was applied to the selected data set. All analyzes were performed using the *Minitab18*®, *R Studio*® and *Visual Basic for Applications* (*VBA*®) software. The first step was to verify that the data are suitable for the application of the FA. Since the data set was not considered as a multivariate normal distribution, the KMO indicator was used. The individual test values showed KMO equal to 0.8; 0.67 and 0.79 for ID, P and NW, respectively. The overall KMO is equal to 0.75, so it is possible to infer that all data are suitable for application of the multivariate FA strategy.

The next step was to apply the multivariate strategy. Given the Kaiser criterion [1], it was verified that the CTQs can be represented by two factors, $RF_{1\text{-}v}$ and $RF_{2\text{-}v}$. Then, the scores of the rotated factors and the eigenvalues were extracted by the *varimax* method considering two factors. To demonstrate the influence of score rotation, Table 8 presents the factor loadings and communalities for the original (unrotated) and rotated method. Factor loadings with values close to 1 or -1 indicate that this factor significantly influences the variable. As a result, it was possible to observe that the rotation of the scores provided a better explanation of the data, where $RF_{1\text{-}v}$ adequately explained the *ID* and *NW* characteristics, while $RF_{2\text{-}v}$ explained the characteristic *P*. With regards to the original approach (unrotated), note that F_1 holds the explanation of all CTQs, keeping F_2 with low explanation in its factor loadings. This behavior can be verified by the variability explained by each factor (variance), i.e., F_1 unrotated has higher value than RF_1 (rotated). However, when evaluating F_2, the original method has a low value (0.4240), because this factor does not adequately explain any of the CTQs. When analyzing $RF_{2\text{-}v}$ with the *varimax* rotation, the variance of this factor is greater than 1, indicating that $RF_{2\text{-}v}$ is significant according to the Kaiser criterion [1]. These results indicate that the rotation of the axes promoted a simpler interpretation of the data, favoring their explanation. It is important to highlight that the variance values are equal to the eigenvalues for the unrotated approach.

TABLE 8. **Factor loadings and communalities for unrotated and varimax rotation**

Unrotated factor scores				*Varimax* Rotation			
Variable	**F_1**	**F_2**	**Communality**	**Variable**	**$RF_{1\text{-}V}$**	**$RF_{2\text{-}V}$**	**Communality**
ID	**0.877**	-0.307	0.862	*ID*	**0.884**	0.285	0.862
NW	**0.885**	-0.206	0.827	*NW*	**0.830**	0.370	0.827
P	**0.843**	0.536	0.997	*P*	0.348	**0.936**	0.997
Variance	2.262	0.424	2.686	*Variance*	1.591	1.095	2.686
% Var	0.754	0.141	0.895	*% Var*	0.530	0.365	0.895

Table 9 presents the rotated factor scores for $RF1_{-v}$ and $RF2_{-v}$, representing the CTQs. Based on these values, the WRF_v vector was obtained using Eq. (23), which represents all the quality characteristics analyzed by a single vector. Given the values of the WRF vector (Table 9), one can estimate the variance components for the WRF_v vector. Based on the analysis of variance, and considering a confidence level of 95%, the interaction term is not significant for the study (*p-value* equal to 0.387), as illustrated in Fig. 2. Thus, by removing the interaction term, new results for the analysis of variance were obtained as presented in Table 10.

TABLE 9. **Rotated factor scores and WRF vector scores for varimax approach**

n	*a*	*b = A*			*b = B*			*b = C*			*b = D*		
		$F_{1\text{-}v}$	**$F_{2\text{-}v}$**	**WRF_v**	**$F_{1\text{-}v}$**	**$F_{2\text{-}v}$**	**WRF_v**	**$F_{1\text{-}v}$**	**$F_{2\text{-}v}$**	**WRF_v**	**$F_{1\text{-}v}$**	**$F_{2\text{-}v}$**	**WRF_v**
1	1	0.9433	-1.6491	1.4347	1.1913	-1.8881	1.8944	1.3725	-2.3369	2.1140	1.1798	-1.9830	1.8282
2	1	1.1305	-1.8748	1.7625	1.0439	-1.8252	1.5875	1.2058	-2.1411	1.8198	1.0225	-1.6808	1.6005
3	1	1.2057	-2.1364	1.8216	1.2030	-2.0270	1.8619	1.2631	-2.3042	1.8804	1.0869	-1.7200	1.7294
1	2	0.9272	0.0760	2.1299	0.7974	0.4422	1.9915	1.0469	0.0453	2.3875	1.0173	0.0840	2.3370
2	2	1.0761	-0.0747	2.4027	0.9118	0.3783	2.2230	1.1685	-0.0055	2.6411	0.9284	0.3901	2.2658
3	2	1.0476	0.1889	2.4500	1.1360	0.1389	2.6289	1.1840	-0.0255	2.6676	1.1419	0.1581	2.6503
1	3	-1.0560	-0.4585	-2.5833	-1.1180	-0.0222	-2.5386	-0.9951	-0.5629	-2.4899	-1.1377	-0.0995	-2.6159
2	3	-0.9754	-0.5656	-2.4465	-1.2602	0.0446	-2.8320	-1.0124	-0.5564	-2.5261	-1.0655	-0.0150	-2.4167
3	3	-0.9815	-0.5719	-2.4630	-1.2744	-0.0776	-2.9159	-0.8706	-0.6025	-2.2250	-1.0127	-0.1900	-2.3717
1	4	-1.6021	-0.8241	-3.9738	-1.5869	-0.6926	-3.8836	-1.5575	-0.9525	-3.9273	-1.5864	-0.9543	-3.9934
2	4	-1.6091	-0.7007	-3.9373	-1.6545	-0.6706	-4.0273	-1.5135	-0.9912	-3.8443	-1.5277	-0.8376	-3.8113
3	4	-1.4938	-0.9989	-3.8030	-1.4634	-1.1565	-3.8010	-1.5036	-1.1195	-3.8763	-1.4590	-1.0071	-3.7278
1	5	0.9719	0.9512	2.6021	1.1362	0.7105	2.8716	1.1482	0.8679	2.9655	1.1209	0.9843	2.9532
2	5	0.8252	1.1389	2.3497	0.9922	1.0474	2.6888	0.9889	1.0669	2.6896	1.0905	1.0171	2.8983
3	5	0.8990	1.2317	2.5560	1.2709	0.6728	3.1603	0.9920	1.1737	2.7418	0.9690	1.0822	2.6510
1	6	-0.1775	0.1164	-0.3521	-0.3230	0.1679	-0.6594	-0.0791	-0.0219	-0.1884	-0.1814	0.2656	-0.2978
2	6	-0.3378	0.1871	-0.6848	-0.3709	0.3267	-0.7006	-0.1520	0.1265	-0.2903	-0.1951	0.3840	-0.2784
3	6	-0.1274	-0.0421	-0.3060	-0.1005	-0.0338	-0.2416	-0.0323	-0.1816	-0.1501	-0.2640	0.3016	-0.4693
1	7	0.2988	0.9393	1.0742	0.5792	0.6691	1.5940	0.4091	0.7527	1.2446	0.3015	1.0626	1.1327
2	7	0.3739	1.0326	1.2836	0.2319	1.2263	1.0446	0.4983	0.5809	1.3736	0.3506	1.0546	1.2402
3	7	0.3417	0.9185	1.1625	0.4815	0.5860	1.3377	0.4787	0.6181	1.3450	0.3930	0.9007	1.2710
1	8	-0.8769	1.0282	-1.5478	-0.9009	1.3026	-1.4856	-0.6776	0.8143	-1.1875	-0.8062	0.8763	-1.4523
2	8	-0.7385	0.9827	-1.2539	-0.8795	1.1721	-1.4926	-0.7480	0.9793	-1.2769	-0.8375	1.2920	-1.3468
3	8	-0.7997	0.8601	-1.4443	-0.7962	0.9997	-1.3774	-0.8789	1.0243	-1.5540	-0.7763	1.1396	-1.2730

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019031, IEEE Access

$$\mathbf{WRF}_{varimax} = \sum_{i=1}^{m}\left[\lambda_i \mathbf{RF}_i\right] = 2.262\,\mathbf{RF}_{\mathbf{1}-V} + 0.424\,\mathbf{RF}_{\mathbf{2}-V} \tag{23}$$

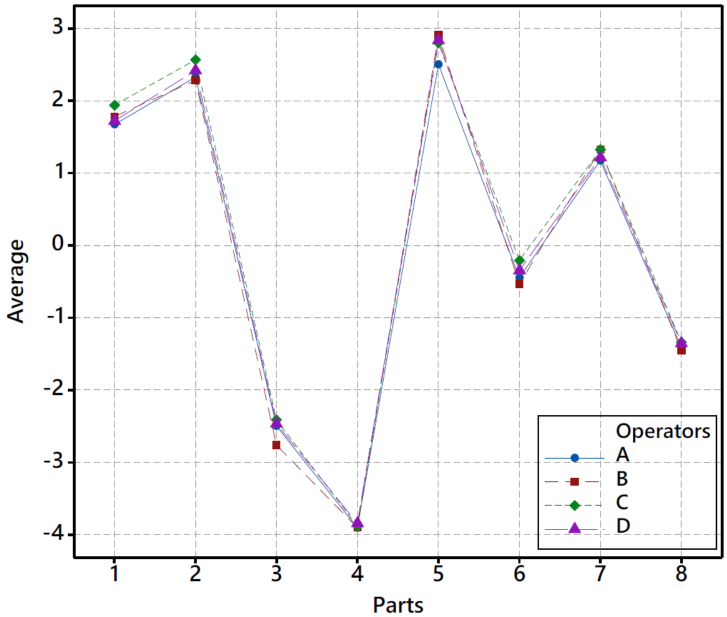


FIGURE 2. **Interaction plot for parts and operators for WRF approach.**

TABLE 10. **Analysis of variance for WRFv scores.**

Source	DF	SS	MS	F	P
Parts	7	500.546	71.507	2631.44	0.000
Operators	3	0.415	0.139	5.10	0.003
Repeatability	85	2.310	0.027		
Total	95	503.271			

Based on Table 10, the parts and the operators reject the null hypothesis that the average of the groups are equal (*p-value* < 0.05). With this information, it was possible to estimate multivariate indicators for the WRF_v vector based on *varimax* rotation. From the available metrics, the indicators show that the repeatability and reproducibility study can be classified as *acceptable*, where the value of *%R&R* equals 7.29% and number of distinct categories identified by the system greater than 5, as suggested by AIAG [31]. Table 11 describes these results.

By evaluating the data individually, the *ID* and *NW* characteristics present less variability compared to the *P* characteristic. This explains the grouping created by the rotation of the scores, favoring the interpretation of the data with greater similarity. In addition, it was possible to verify that the eigenvalue for RF_{1-v} (which explains *ID* and *NW*) presented a higher value, prioritizing the weighting of this factor in relation to RF_{2-v}, which represents only the characteristic with greater variability.

To demonstrate and verify the consistency of the measurement amplitude for the WRF_v vector with *varimax* rotation, Fig. 3 illustrates the *R-control chart*. This chart presents the operators' ranges, which shows that all operators presented measurements within the upper and lower control limits. More specifically, operators C and D presented greater stability in their measurements than operators A and B.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019031, IEEE Access

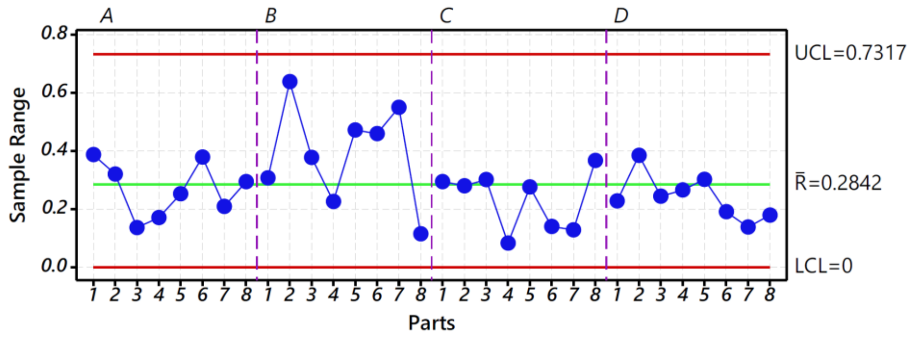


FIGURE. 3. R-control chart for the GR&R-WRF

TABLE 11. Variance components and gage evaluation for WRF

Source	*VarComp*	*%Contribution*	*StdDev (SD)*	*%Study Var*
Total Gage R&R	0.03181	0.53	0.17836	7.29
Repeatability	0.02717	0.45	0.16485	6.74
Reproducibility	0.00464	0.08	0.0681	2.78
Operator	0.00464	0.08	0.0681	2.78
Part-To-Part	5.95662	99.47	2.44062	99.73
Total Variation	5.98843	100	2.44713	100
ndc		19		

To properly represent data variability, Fig. 4 illustrates the confidence regions of the data through the confidence ellipses originally proposed in the WF method. Given the confidence ellipses for ndc_m, proposed by Almeida et al. [27], notice the data for *ID*×*NW* have narrower ellipses, indicating the most precise intervals, as well as non-overlapping ellipses. However, when checking the relation *NW*×*P* and *P*×*ID*, we verified the presence of overlapping ellipses, as well as larger confidence regions, indicating a high variability due to the presence of the quality characteristic *P*.

B. COMPARISON WITH QUARTIMAX ROTATION

The factor scores were also extracted using the *quartimax* method, which is the most used rotation method. Table 12 presents the loadings and communalities factor for *quartimax* rotation. Notice the subtle difference in the structure of the factor loads, where the loads for $RF_{1\text{-}Q}$ present higher values for the characteristics ID and NW, when compared to the *varimax* method (see Table 8). However, there is a disparity when comparing the loadings of the second factor between *quartimax* and *varimax*. The behavior for $RF_{2\text{-}v}$ of *varimax* showed a higher loading, better balancing the factors in relation to the *quartimax* method (such difference is visible due to the balance of the total variance explained). Hence, although the *quartimax* method favors the explanation for ID and NW, compared to the $RF_{1\text{-}Q}$, such rotation presents an unsatisfactory result when loading the variable P, inferring a more confusing structure to interpret this variable. To better illustrate this loading behavior, Fig. 5 and Fig. 6 presents the loads and groupings for both rotation methods.

TABLE 12. Factor loadings and communalities for quartimax rotation

Quartimax Rotation			
Variable	**$RF_{1\text{-}Q}$**	**$RF_{2\text{-}Q}$**	**Communality**
ID	**0.929**	0.008	0.862
NW	**0.903**	0.106	0.827
P	0.611	**0.790**	0.997
Variance	2.051	0.635	2.686
% Var	0.684	0.212	0.895

Given the eigenvalues and the scores rotated by the *quartimax* method (additional data available in the supplementary material), the WRF_Q vector was calculated using Eq. (24). It is important to note that the eigenvalues for *quartimax* rotation remain the same, since the extraction of both is based on principal components. Based on the results, the variance components were estimated, where the interaction term was not significant (*p-value* equal to 0.472), confirming what had already been verified when using *varimax* rotation. In view of the results, it is inferred again that the parts and the operators reject the null hypothesis that the average of the groups are equal (*p-value* equal to 0.000). However, presenting *p-value* for operators subtly lower than that verified in *varimax* rotation.

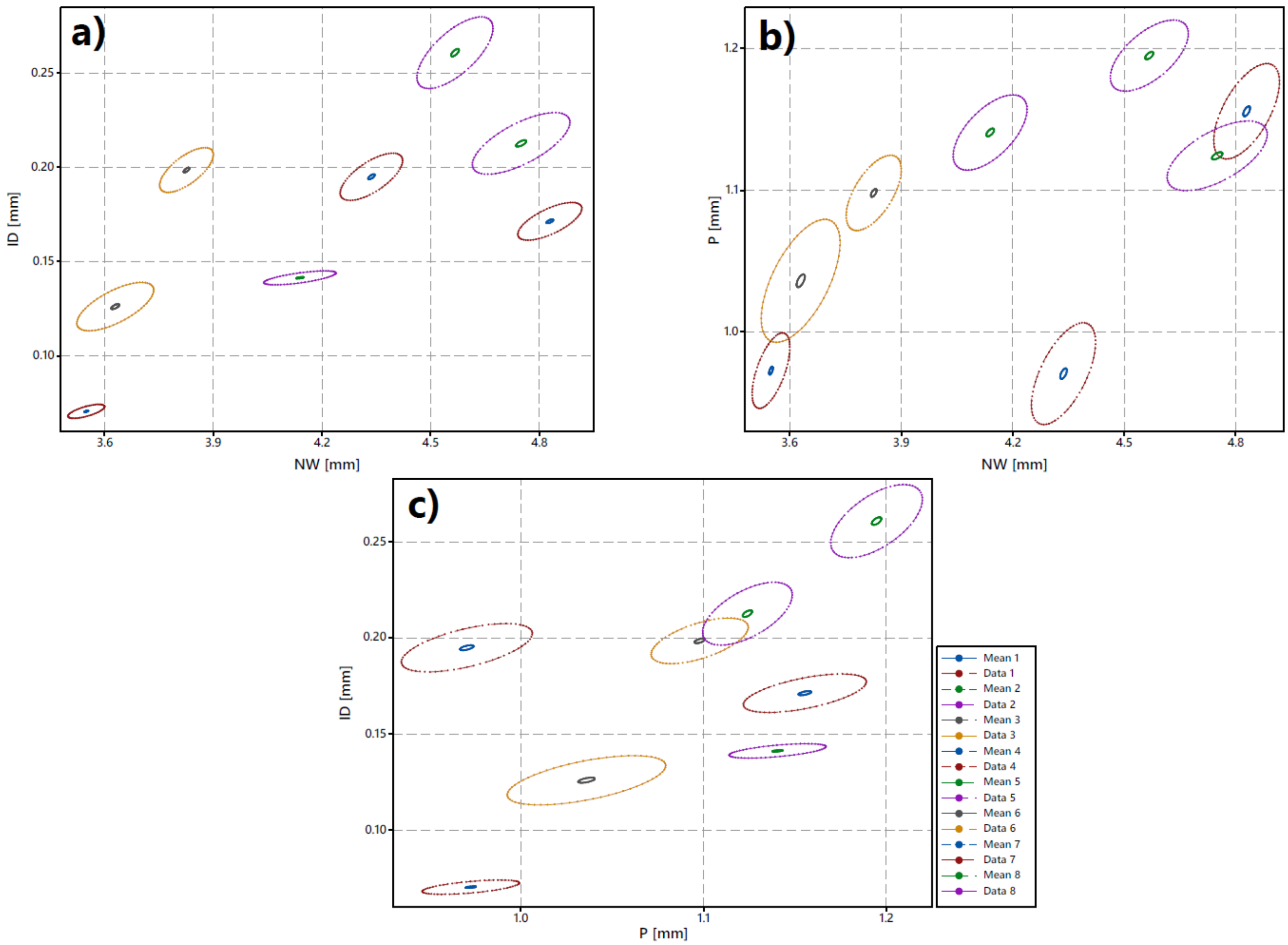


FIGURE. 4. 95% confidence ellipses for (a) ID×NW, (b) P×NW and (c) ID×P.

$$\mathbf{WRF}_{quartimax} = \sum_{i=1}^{m} \left[\lambda_i \mathbf{RF}_i \right] = 2.262\,\mathbf{RF}_{\mathbf{1}-Q} + 0.424\,\mathbf{RF}_{\mathbf{2}-Q} \tag{24}$$

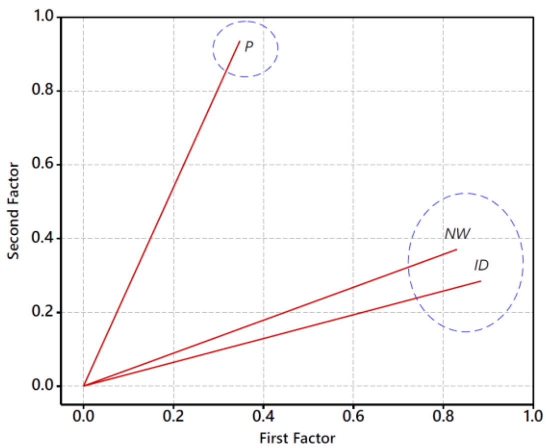


FIGURE 5. Loadings for *varimax* rotation

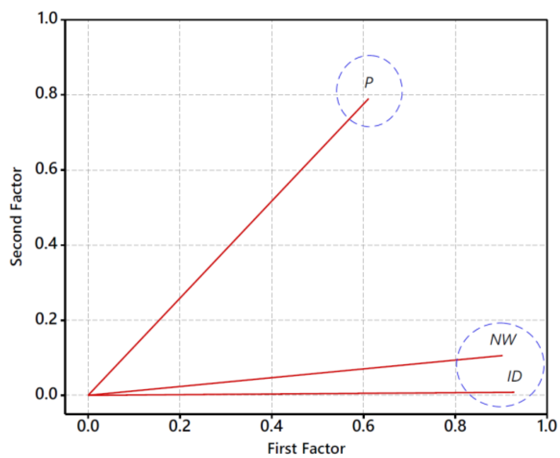


FIGURE 6. Loadings for *quartimax* rotation

Based on the multivariate indicators, the WRF_Q method also presented an *acceptable* classification, with $\%R\&R_m$ equal to 5.58% and ndc_m equal to 25. Although the WRF_Q, values, in this particular case, are more attractive from the manager's point of view (due to the low variability presented in the study), such behavior is not ideal. This can be easily explained by comparing the loading values for *quartimax* rotation, where the *quartimax* method prioritized the loading for the ID and NW responses, which showed less variability in the study. Howeber, the variable P, grouped in the second factor, presented less load with the *quartimax* rotation. The total explained variance values (*%Var* from Tables 8 and 12) confirm this statement. Thus, in this measurement study with *quartimax* rotation, there is a confusing structure and poor simplification, with an imbalance in the prioritization of observable variables. Additionally, the variable P, which has greater variability (see Fig. 4), was less prioritized.

However, it is important to note that such results do not detract from the *quartimax* method for use. The choice of the rotation method varies according to the data analyzed [32], whose degree of explanation and simplicity may vary. Thereunto, one must analyze the data structure before choosing the option that will stand out and present the best results, favoring the decision maker and the evaluation of the quality of extensive and correlated data.

V. CONCLUSIONS

Given the need to use appropriate strategies to consider the structure of variance-covariance matrix of the data, multivariate techniques can enhance studies related to the measurement system and data quality. This study sought to present an approach to contribute to the metrics of multivariate measurement system, using the FA technique with orthogonal rotations. In addition, the weighting of the rotated scores by the respective eigenvalues was added to form a single vector of responses, which adequately represented all the responses of interest. The behavior of the method was demonstrated for the RSW process and, from that, the following conclusions can be reached:

- The method proved to be a suitable alternative to analyze the measurement system for data with a significant variance-covariance structure, improving performance and precision in multivariate measurement system assessment;
- The proposed approach presented the possibility of a single metric, filling the gaps of other methods of GR&R with FA. Thus, WRF contemplates the use of orthogonal rotations (to improve interpretation) and the weighting of factors (attributing the degree of importance associated with the eigenvalue). This procedure creates a unique analysis for the variability of the measurement system;
- The application carried out in the RSW process showed that the method improved the interpretation and explanation of latent variables, simplifying the data structure using the *varimax* method. This rotation method stood out for this data set, as it presented a greater balance in the division of factor loadings for the geometric characteristics of the process;
- When comparing the case study using the *quartimax* method, it was found that it may not be the best choice for this particular data set, since it presented moderate loads for the variable P, which has the greatest variability. As a consequence, the *quartimax* rotation promoted a more confusing loading structure, prioritizing the variables explained in the first factor. However, it is important to note that the *quartimax* method can be a useful option for future applications in different processes and data sets;
- The use of the FA method with extraction by principal components provided a minimization in the data dimension and, consequently, in the computational effort for processing. In addition, the analysis time has been significantly reduced (approximately a 75% reduction), since the method promotes a single response vector that represents all the variables observed in a significant way.

Finally, as future suggestions, the method can be extended to different applications, industrial or from other segments. In addition, the authors suggest comparisons with other unusual rotation methods available in the literature and other applications focused on process quality.

VI. REFERENCES

[1] Johnson, R.A., Wichern D. Applied Multivariate Statistical Analysis. 6th ed. New Jersey: Prentice-Hall; 2007.

[2] Torres AF, Rocha FB, Almeida FA, Gomes JHF, Paiva AP, Balestrassi PP. Multivariate Stochastic Optimization Approach Applied in a Flux-Cored Arc Welding Process. IEEE Access 2020;8:61267–76. https://doi.org/10.1109/ACCESS.2020.2983566.

[3] Adegoke NA, Abbasi SA, Smith ANH, Anderson MJ, Pawley MDM. A Multivariate Homogeneously Weighted Moving Average Control Chart. IEEE Access 2019;7:9586–97.

[4] Puggina Bianchesi NM, Romao EL, Lopes MFBP, Balestrassi PP, De Paiva AP. A design of experiments comparative study on clustering methods. IEEE Access 2019;7:167726–38. https://doi.org/10.1109/ACCESS.2019.2953528.

[5] Song B, Zhou X, Tan S, Shi H, Zhao B, Wang M. Process Monitoring via Key Principal Components and Local Information Based Weights. IEEE Access 2019;7:15357–66.

[6] Song B, Yan H, Shi H, Tan S. Multisubspace Elastic Network for Multimode Quality-Related Process Monitoring. IEEE Trans Ind Informatics 2020;16:5874–83. https://doi.org/10.1109/TII.2019.2959784.

[7] Almeida FA de, Miranda Filho J, Amorim LF, Gomes JH de F, Paiva AP de. Enhancement of discriminatory power by ellipsoidal functions for substation clustering in voltage sag studies. Electr

Power Syst Res 2020;185:106368. https://doi.org/https://doi.org/10.1016/j.epsr.2020.106368.

[8] Mehmood RM, Du R, Lee HJ. Optimal Feature Selection and Deep Learning Ensembles Method for Emotion Recognition From Human Brain EEG Sensors. IEEE Access 2017;5:14797–806.

[9] Almeida LM de L, Balanco PA de F. Application of multivariate analysis as complementary instrument in studies about structural changes: An example of the multipliers in the US economy. Struct Chang Econ Dyn 2020;53:189–207. https://doi.org/https://doi.org/10.1016/j.strueco.2020.02.006.

[10] Almeida FA, De Paula TI, Leite RR, Gomes GF, Gomes JHF, Paiva AP, et al. A multivariate GR&R approach to variability evaluation of measuring instruments in resistance spot welding process. J Manuf Process 2018;36:465–79. https://doi.org/10.1016/j.jmapro.2018.10.030.

[11] Reddy GT, Reddy MPK, Lakshmanna K, Kaluri R, Rajput DS, Srivastava G, et al. Analysis of Dimensionality Reduction Techniques on Big Data. IEEE Access 2020;8:54776–88.

[12] de Almeida FA, Santos ACO, de Paiva AP, Gomes GF, Gomes JH de F. Multivariate Taguchi loss function optimization based on principal components analysis and normal boundary intersection. Eng Comput 2020. https://doi.org/10.1007/s00366-020-01122-8.

[13] Song B, Zhou X, Shi H, Tao Y. Performance-Indicator-Oriented Concurrent Subspace Process Monitoring Method. IEEE Trans Ind Electron 2019;66:5535–45. https://doi.org/10.1109/TIE.2018.2868316.

[14] Belinato G, de Almeida FA, de Paiva AP, de Freitas Gomes JH, Balestrassi PP, Rosa PARC. A multivariate normal boundary intersection PCA-based approach to reduce dimensionality in optimization problems for LBM process. Eng Comput 2019;35:1533–44. https://doi.org/10.1007/s00366-018-0678-3.

[15] Liu C, Ding Q. A Modified Algorithm for the Logistic Sequence Based on PCA. IEEE Access 2020;8:45254–62.

[16] Sun X, Wang X, Cai D, Li Z, Gao Y, Wang X. Multivariate Seawater Quality Prediction Based on PCA-RVM Supported by Edge Computing Towards Smart Ocean. IEEE Access 2020;8:54506–13.

[17] Almeida FA, Leite RR, Gomes GF, Gomes JH de F, de Paiva AP. Multivariate data quality assessment based on rotated factor scores and confidence ellipsoids. Decis Support Syst 2020;129:113173. https://doi.org/10.1016/j.dss.2019.113173.

[18] Alvin C. Rencher. Methods of Multivariate Analysis. 2nd ed. New York: John Wiley & Sons, Inc.; 2002. https://doi.org/10.1002/0471271357.

[19] Costello AB, Osborne JW. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Pract Assess Res Eval 2005:1–9.

[20] Daroz Gaudêncio JH, de Almeida FA, Turrioni JB, da Costa Quinino R, Balestrassi PP, de Paiva AP. A multiobjective optimization model for machining quality in the AISI 12L14 steel turning process using fuzzy multivariate mean square error. Precis Eng 2019;56:303–20. https://doi.org/10.1016/j.precisioneng.2019.01.001.

[21] León-García F, Rodríguez-Lozano FJ, Olivares J, Palomares JM. Data Communication Optimization for the Evaluation of Multivariate Conditions in Distributed Scenarios. IEEE Access 2019;7:123473–89.

[22] de Almeida FA, Gomes JH de F, Gomes GF, Romão EL, Balestrassi PP. Variation causes analysis attributed to different metrological instruments to verify the geometric characteristics of a spot welding process. Soldag e Insp 2018;23:485–504. https://doi.org/10.1590/0104-9224/SI2304.05.

[23] Woodall WH, Borror CM. Some Relationships between Gage R&RCriteria. Qual Reliab Eng Int 2008;24. https://doi.org/10.1002/qre.870.

[24] Burdick RK, Borror CM, Montgomery DC. A Review of Methods for Measurement Systems Capability Analysis. J Qual Technol 2003;35:342–54. https://doi.org/10.1080/00224065.2003.11980232.

[25] Majeske KD. Approval Criteria for Multivariate Measurement Systems. J Qual Technol 2008;40:140–53. https://doi.org/10.1080/00224065.2008.11917721.

[26] Wang F-K, Yang C-W. APPLYING PRINCIPAL COMPONENT ANALYSIS TO A GR&R STUDY. J Chinese Inst Ind Eng 2007;24:182–9. https://doi.org/10.1080/10170660709509032.

[27] Almeida FA de, Gomes GF, Gaudêncio JHD, Gomes JH de F, Paiva AP de. A new multivariate approach based on weighted factor scores and confidence ellipses to precision evaluation of textured fiber bobbins measurement system. Precis Eng 2019;60:520–34. https://doi.org/10.1016/j.precisioneng.2019.09.010.

[28] Thurstone LL. Multiple-factor analysis. Chicago Univ Chicago Press 1947:535.

[29] Al-Refaie A, Bata N. Evaluating measurement and process capabilities by GR&R with four quality measures. Measurement 2010;43:842–51. https://doi.org/10.1016/J.MEASUREMENT.2010.02.016.

[30] Montgomery DC. Design and Analysis of Experiments. 2017.

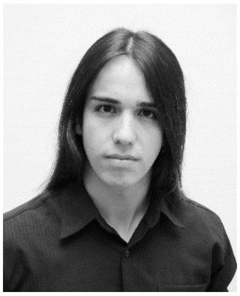
[31] AIAG. Measurement systems analysis: reference manual. 4th ed. Detroid, MI, USA: Automotive Industry Action Group; 2010.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3019031, IEEE Access

[32] Hair JF, Tatham RL, Anderson RE, Black W. Multivariate Data Analysis. 5th ed. Pearson College Div; 1998.

[33] Darton RA. Rotation in Factor Analysis. J R Stat Soc Ser D (The Stat 1980;29:167–94. https://doi.org/10.2307/2988040.

[34] Almeida FA, Gomes GF, Sabioni RC, Gomes JHF, Paula VR, Paiva AP, et al. A Gage Study Applied in Shear Test to Identify Variation Causes from a Resistance Spot Welding Measurement System. Stroj Vestn J Mech Eng 2018;64:621–31. https://doi.org/10.5545/sv-jme.2018.5235.



FABRÍCIO A. ALMEIDA received his B.S. degree in economics from the Faculty of Economic Sciences Southern Minas Gerais (2014). He received his M.S. degree in industrial engineering from the Federal University of Itajubá (2017) and he is currently pursuing a Ph.D. in industrial engineering at UNIFEI. Currently, he is Black Belt in Lean Six Sigma and a professor at the Faculty of Economic Sciences Southern Minas Gerais and a member of the research group of the Nucleus of Manufacturing Optimization and Innovation Technology. His research areas include multivariate statistical analysis, multiobjective optimization, quality engineering, industrial economics, and design of experiments. Metrics since 2018: 150 citations and *h7* factor.

SIMONE C. STREITENBERGER received the B.S. degree in Computer Science from Federal University of Itajubá, Minas Gerais, Brazil, in 2007, the graduation degree in Project Management for Information Technology from São Paulo University, São Paulo, Brazil, in 2013, and the M.Sc. degree in Industrial Engineering from Federal University of Itajubá, Minas Gerais, Brazil, in 2020. From 2008 to 2017, she worked in Information Technology area with enterprises such as Itaú Unibanco bank and SAP Brazil. Her research interests include Design of Experiments, Multiobjective Optimization, Genetic Algorithms and Multivariate Data Analysis

ALEXANDRE F. TORRES received his B.S. degree in industrial engineering from the Federal University of Itajubá (UNIFEI). He received his M.S. degree in industrial engineering from the same university and is currently pursuing his Ph.D. in industrial engineering at UNIFEI. Currently. he is a professor at Centro Universitário de Itajubá, teaching Statistics for decision making for graduate students. He is also a professor at Centro de Ensino Superior em Gestão, Tecnologia e Educação (FAI) in Santa Rita, Brazil, where he teaches quality engineering for undergraduate students of industrial engineering. His research areas include multiobjective optimization, hardened steel turning optimization, and stochastic programming.

ANDERSON P. PAIVA is a professor of the Industrial Engineering and Management Institute at Federal University of Itajubá since 2005. He has 67 journal papers published on Web of Science, 67 journal papers published on Scopus and he has a H9 factor. He got his bachelor's degree in mechanical engineering at Centro Universitário do Sul de Minas in 1996, his master's degree in industrial engineering at Federal University of Itajubá (UNIFEI) in 2004 and his doctor's degree in mechanical engineering at Federal University of Itajubá (UNIFEI) in 2006. His main research areas are: multiobjective optimization, design of experiments and multivariate statistics.

JOSÉ H. F. GOMES is a Mechanical Industrial Engineer from the Federal University of Itajubá (2003-2007), with master's degree (2008-2010) and PhD (2010-2013) in Industrial Engineering from the same institution. Currently Associate Professor I of the Institute of Industrial and Management Engineering (IEPG) of the Federal University of Itajubá, in the undergraduate courses in Production Engineering and Administration, and postgraduate in Production Engineering. It acts in the research lines of modeling, analysis and optimization of production systems and manufacturing operations. His main research areas are: application and improvement of multi-objective methods, statistical methods, mathematical programming methods and materials management.