STUDY OF THE PROPAGATION OF ERRORS IN MULTIVARIATE PRODUCTION PROCESSES IN STAGES

Josiane da Silva Jesus
Instituto Militar de Engenharia - IME
Praça General Tibúrcio, 80, 22291-270 - Praia Vermelha, Rio de Janeiro – RJ.
josianesj@id.uff.br

Paulo Henrique Coelho Maranhão
Instituto Militar de Engenharia – IME
Praça General Tibúrcio, 80, 22291-270 - Praia Vermelha, Rio de Janeiro – RJ.
maranhao@ime.eb.br

RESUMO

Atualmente, uma parte considerável dos processos produtivos contém dois ou mais estágios. Assim a qualidade final de um produto está relacionada de como possíveis erros se propagam nos diversos estágios. Além disso, não é raro que as variáveis de interesse desses processos sejam correlacionadas, de forma que um erro em uma das variáveis afeta todas as outras e em consequência, todo o processo é afetado. Este trabalho propõe um modelo multivariado para estudar a propagação dos erros entre estágios. O estudo foi realizado por meio de simulações, que mostrou resultados bastante promissores.

PALAVRAS CHAVE. Processo produtivo, propagação dos erros, multivariado.

Tópicos (4. Resultados, 2. Métodos, 5. Conclusão, 1. Introdução, 3 Metodologia).

ABSTRACT

Currently, a considerable part of the productive processes contains two or more stages. Thus, the final quality of a product is related to how possible errors propagate in the several stages. Moreover, it is not uncommon that the variables of interest of these processes to be correlated, so that an error in one of the variables affects all the others and consequently the whole process is affected. This work proposes a multivariate model to study the propagation of errors between stages where the variables are correlated. The study was carried out through simulations, which showed quite promising results.

KEYWORDS. Production process, propagation of variation, multivariate.

Paper topics (4. Results, 2. Methods, 5. Conclusion, 1. Introduction, 3. Methodology).

1. Introduction

The current world, globalized and competitive, it causes a constant search for quality, efficiency, reduction of costs and greater financial return in any organization. This set of goals became a basic need and crucial to the survival of companies and industries.

A product or service can be adversely affected by variability during the manufacturing process or execution. This variability are deviations that represent the difference between the obtained result and the desired result in any process. It is also known, that a large part of the production processes have two or more stages, and often these stages are correlated. In this way, to study the propagation of errors between stages correlated can be an important contribution to check which process step is responsible for the greatest transmission of the variance, so that it can act preventively in this step.

Generally, the analysis of a process takes into account more than one variable of interest. However, it is not rare to occur cases in which the variables of interest measured in each stage also have a high degree of correlation and techniques multivariate are needed for the development of models that are able to deal with transmission of the error for these cases. The main objective of this work is to develop a study on the propagation of errors in processes in stages, whose quality characteristics of interest are correlated.

The first studies on the transmission of variance were carried out in the mid-nineties. This way, in 1998 there was a study of the variability in a system flexible mounting of doors of vehicles with the aid of the layout provided by the CAD program, and a technique in multivariate statistics called Principal Component Analysis, which describes the variability of the process through a linear transformation of the variables observed originally [Ceglarek 1998]. In 1999, it was proposed a model autorregressivo of order 1 (AR1) to estimate the propagation of the variation in stages of a production process in relation to a single variable of interest, identifying which step in production contributes significantly to the variability of the process [Lawless et al. 1999]. Subsequently, it was developed a method to evaluate the propagation of variation in a machining process by means of a model of the state-space, based on the information of the product design and process, was also used the concept of a virtual operation to isolate faults and determine the causes of the same [Huang et al. 2002].

Already in 2004, was evaluated a machining process through a matrix of homogeneous transformation, whose elements represent the spatial difference between the two systems of coordinates, and had been employed a method of Analyzing Components Designated (DCA). The method DCA is similar to the method of Principal Components, the difference lies in the fact that the DCA is able to identify failures of fixation simultaneous in sheet metal and does not consider any kind of interaction between them and the manufacturing process [Camelio et al. 2004]. In the same year, was also implemented the Six Sigma methodology in a production process of automobiles, with the aim of reducing the deviations generated. For this, it was analyzed the clearance present in the rear doors of 17 vehicles in two stages of manufacture, plating and assembly [Gaio e Sá 2004].

In the following year, [Liao and Wang, 2005] developed a new method from the fractal geometry, which characterizes phenomena, spatial or temporal of the continuous bodies, and the use of the method of finite elements, with the objective of analyzing the variation in the surface micro-geometric of the components of these bodies. The fractal geometry has been expressed by a function called the Weierstrass-Mandelbrot, able to represent the characteristics of the surface micro geometry of the parts used in the assembly process. The finite element method was used to analyze the deformation of the components of the production system. This variation deserves attention in processes of assembly of high precision. Subsequently, [Zhang et al. 2007] presented the methodology of the flow variation based on the description of the project (CAD) and in the process. This methodology uses the modeling of the space-state (linear and that establishes the relationship between the errors and their causes), sensitivity analysis to verify the degree of impact of deviations on the quality of the automotive cylinders, and establish the optimization process in order to minimize them. The study of the decomposition of the error assists the engineers to identify which stage contributes most to the variability of the process.

In 2012, was conducted a study aiming to the minimization of the propagation of variances. The idea was to evaluate stage by stage of the process by minimizing the sum squared of the variations in all the phases [Yang et al. 2012]. In the following year it was proposed a model of quality forecast (CQPM), able to deal with complex variables present in multi-stage manufacturing, through to technical data such as Principal Components (ACP), which extracts information from several variables interrelated and transform them in a new set of variables orthogonal [Arif et al. 2013].

The proposed paper is organized in 5 sections. In section 2, are discussed the main concepts and definitions of tools of Multivariate Statistics and the Model used. Section 3 describes the methodology applied in the simulations. The fourth section presents the analysis of the results. Finally, the fifth section contains the conclusion of the work.

2. Methods

2.1 Autoregressive Model Propagation

In 1999, [Lawless et. al, 1999], have proposed an autoregressive model (AR (1)) to study the transmission of variation in processes of multiple stages. The choice of this model is justified since we wished to study the influence of the variance between the subsequent steps. In this way, the following equation illustrates the model AR (1):

$$Y_i = \alpha_i + \beta_i Y_{i-1} + \varepsilon_i \quad i = 2, \dots, k \tag{1}$$

where Y_i is the variable that represents the measurement observed in step i and Y_{i-1} is the observed measure in the step immediately preceding. Besides Y_i follows a Normal distribution with mean μi and variance σ_i^2. The terms α_i and β_i represent, respectively the linear and angular coefficients of the model. The waste εi´s represent the random errors and also follow a Normal distribution with mean 0 and constant variance σ_ε^2. In addition, it is assumed that the errors and Y_i have covariance equal to zero, that is, they are independent.

The main attribute of the models AR (1) is that the current value is strongly related to the immediately preceding, that is, there is a correlation in the first lag, between Y_i and $Y_{(i-1)}$. This contributes to the fact that the errors can be obtained in each lag considered.

Thus, the variability that is transmitted between two stages is established by the following equation:

$$\sigma_i^2 = \beta_i^2 \sigma_{i-1}^2 + \sigma_{i,A}^2 \tag{2}$$

where σ_i^2 is the total variance of step i; $\sigma_{i,A}^2$ is the variance added in the current step; σ_{i-1}^2 is the variance in the previous step; β_i^2 is the coefficient responsible for the transmission of the variability between the two stages.

If we have k steps, the equation (2) can still be generalized by the following expression:

$$\sigma_k^2 = \beta_k^2 \sigma_{k-1,A}^2 + \cdots + \beta_k^2 \beta_{k-1}^2 \dots \beta_2^2 \sigma_1^2 + \sigma_{k,A}^2 \tag{3}$$

2.2 Principal Components analysis

Principal Components analysis (PCA) is a multivariate technique, in which a number of related variables is transformed into a smaller set of variables not correlated, called Principal Components (CP). In this method are generated as many components as variables, but the large advantage of the technique is that, in general, few components explain most of the variability of the set of original variables. Thus, if p components represent the total variability of a set of p original variables, the greater part of this variability can be explained by k components (k<p). Soon, k components replace the p variables initials, reducing the size of the data set. In addition, the first CP is a linear combination with variance maximum.

So is X a vector random vector of averages μ and the matrix of covariance or correlation, Σ_{pXp}, with λ_i eigenvalues and e_i the eigenvectors (normalized) that obey the conditions: $e_{i'} e_j = 0$ (i≠j); $e_{i'} e_i = 1$ (for all i) and $\Sigma_{pXp}\, e_i = \lambda_i\, e_i$ (for all i). Soon, the j-th principal component is given by the following:

$$Y_j = e_j' X = e_{j1} X_1 + e_{j2} X_2 + \cdots + e_{jp} X_p \quad (4)$$

The Figure 1 below depicts a scheme of how it works the method.

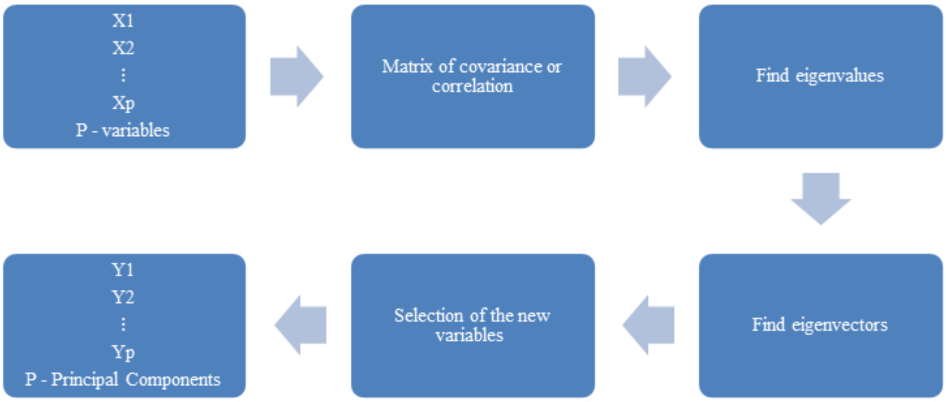

Figure 1: Scheme of operation of the method PCA.

In general, it is observed that a large part of the variability of the original variables is explained by a few principal components. Thus, we can say that if more than 80% of the total variance can be explained by two or three components, then these components can replace the original variables without loss of information [Johnson and Wichern 2007].

2.3 Autoregressive Model in the Principals Components

The autoregressive model of order one, described in the previous section, will serve as the basis for the determination of the autoregressive model in the principals components. In this way, the main component of stage *i* can be modeled as a function of the main component from the previous step *(i-1)*, as can be seen in the following equation:

$$CP_i = \alpha_i + \beta_i CP_{i-1} + \varepsilon_i \quad (5)$$

Applying the variance in the i-th principal component, we have:

$$Var(CP_i) = \sigma_i^2 = \beta_i^2 \sigma_{i-1}^2 + \sigma_{i,A}^2 \quad (6)$$

By analogy, all of the assumptions and definitions applied to the model autorregressivo given by the equation (1) are still valid for the equation (5). Thus, the terms αi and βi are the linear and angular coefficients of the model and the terms εi´s are the residuals. It is also considered that the residuals and the CPs are independent. In practice, the terms of equations (5) and (6) can be estimated by means of equations (7) to (11), which respectively represent the estimators of the covariance of the CPs between the stages, the variance of the CPs, the angular and linear coefficients, and the variance added in the own stage.

$$S_{i-1,i} = \frac{\sum_{j=1}^{n}(CP_{j,i-1}-\overline{CP}_{i-1})(CP_{j,i}-\overline{CP}_{i})}{n} \quad (7)$$

And the variance is given by the equation (8):

$$S_{ii} = \frac{\sum_{j=1}^{n}(CP_{j,i}-\overline{CP}_{i})^2}{n} \quad (8)$$

The parameter of spread is calculated using the equation (9).

$$\beta_i = \frac{S_{i-1,i}}{S_{i-1,i-1}} \quad (9)$$

The linear coefficient and the variance added between steps are obtained by the equations (10) and (11), respectively.

$$\alpha_i = \overline{CP}_i - \beta_i \overline{CP}_{i-1} \quad i=2,\dots,k \quad (10)$$

$$\sigma_{i,A}^2 = S_{ii} - \beta_i(S_{i-1,i}) \quad i=2,\dots,k \quad (11)$$

3. Methodology

In this topic, is presented the methodology used in a study of the simulation. In this way, the idea is to simulate a production process formed by two stages, where each stage is composed of four variables as shown in the Figure 2, in which the same are correlated on different levels. Thus, for each stage were generated 100 samples of these four variables, all following a normal distribution with mean zero and variance 1.

It was also considered three types of correlation between the variables, namely: strong, moderate and weak, and three other types of relationships between the stages, as can be seen in Figure 3. Thus, for the first simulation, Figure 3, is assumed that the variables in each stage are strongly correlated with each other, considering also three types of correlation between the stages. For the second and third simulation the same routine is followed.

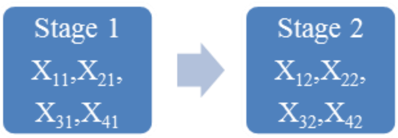

Figure 2: representation of variables in stages.

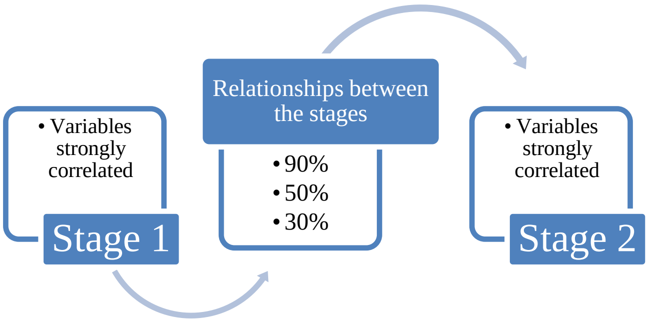

Figure 3: Schematic of the simulated cases considering strongly correlated variables in each stage.

From each group of variables at each stage were generated principal components, selecting the component that best explains the variability of the original variables. Then, it is applied the proposed model between the components of each stage, where the component of the second stage is regressed in the function of the component of the first stage, as shown in Figure 4.

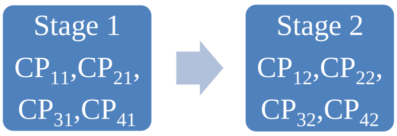

Figure 4: representation of the principals components in stages.

4. Results

In this section, the results are presented related to the study of the simulation described in section 3.

4.1 Results of Simulation 1

After you have obtained the variables strongly correlated in each stage according to the methodology, we have obtained the principals components for the first and the second stages, considering the three relationships between the stages. The Tables 1 and 2 show the variances explained by the principal components in stage 1 and stage 2, respectively.

Table 1: Percentage of variance explained by the CPs simulation 1 Stage 1.

% CP1	% CP2	% CP3	%CP4
99.10%	0.50%	0.30%	0.10%

The values relating to the proportions of the main components in each case in the stage 2 are shown in Table 2.

Table 2: Percentage of variance explained by the CPs simulation 1 Stage 2.

Case	% CP1	% CP2	% CP3	%CP4
1 (90%)	98.90%	0.60%	0.40%	0.10%
2 (50%)	96.70%	1.70%	1.30%	0.30%
3 (30%)	91.80%	4.20%	3.20%	0.80%

Analyzing Tables 1 and 2, it is noted that when the variables originally generated are strongly correlated the first principals components of each stage are responsible for almost the entire variability of the original variables. In this way, we applied the proposed model considering the first components in each stage. Figures 5, 6 and 7 show how the variance propagates between stages.

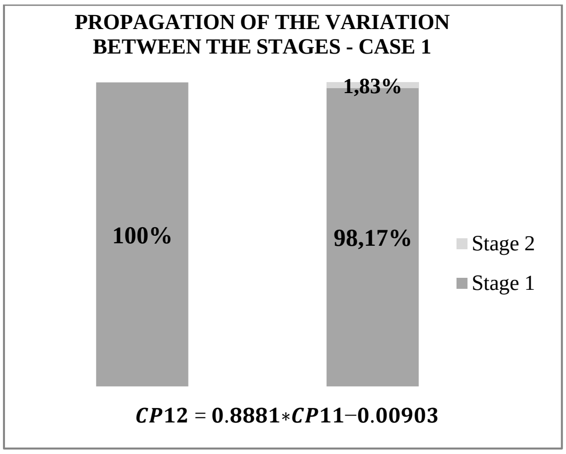

Figure 5: Variability between stages Simulation 1 case 1.

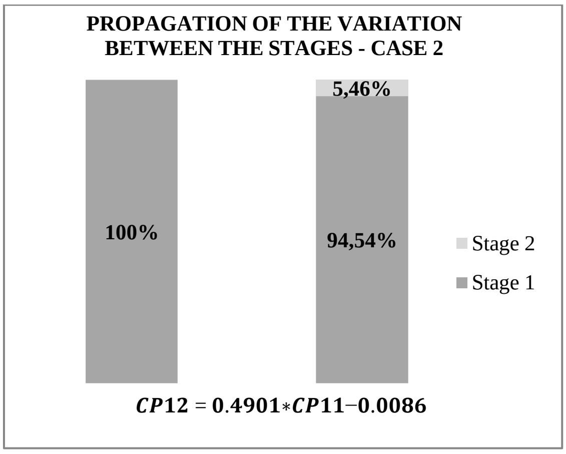

Figure 6: Variability between stages Simulation 1 case 2.

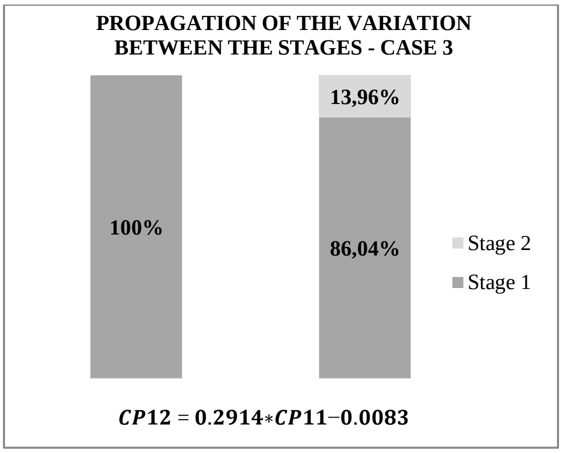

Figure 7: Variability between stages Simulation 1 case 3.

Figure 5 reveals that 98,17% of the variability, present in the stage 2, has been originated in the previous stage. This means that an error in the previous stage, is almost entirely propagated to the next step. Still, it is possible to see by means of the Figures 6 and 7 that as the relationship between the stages decreases, it decreases the transmission of the error.

4.2 Results of Simulation 2

The variances explained by the principals components in stage 1 for simulation 2, are presented in Table 3. Still, the Table 4 presents the values of the variances explained in the stage 2 at each level of the correlation between the stages.

Table 3: Percentage of Variance explained by the CPs simulation 2 Stage 1.

% CP1	% CP2	% CP3	%CP4
96.60%	1.50%	1.40%	0.50%

Table 4: Percentage of Variance explained by the CPs simulation 2 Stage 2.

Case	% CP1	% CP2	% CP3	%CP4
1 (90%)	95.60%	2.20%	1.70%	0.50%
2 (50%)	87.90%	5.90%	4.80%	1.40%
3 (30%)	75.00%	11.70%	9.90%	3.40%

Note that, although the variables in the stages are moderately correlated, the firsts principals components are still capable of a high degree of explanation, even for the worst case, where the relationship between the stages is only 30%. Figures 8, 9 and 10 show the propagation of errors between the two stages in terms of the first principal component, since it represents the largest part of the variability of the variables.

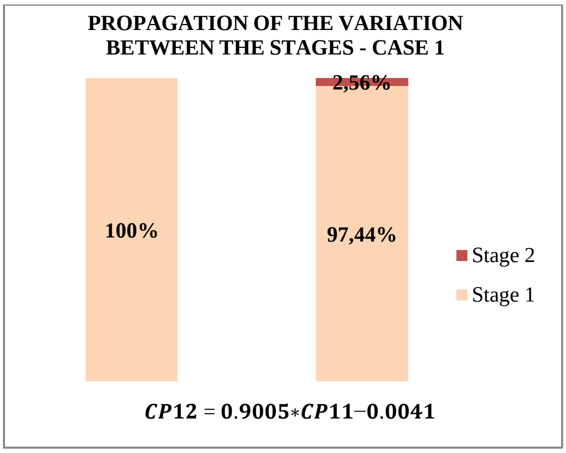

Figure 8: Variability between stages Simulation 2 case 1.

From Figure 8, it is observed that the propagation of the variability between steps was 97,44% for the case 1 of the simulation 2 and Figure 9 shows that in the second stage, 93,63% of the variance was derived from the previous stage and only 6,37% of the errors were generated in the own stage 2. For the correlation between stages of 30%, has that 15,04% of the deviations present in the second stage were generated by the same, as shown in Figure 10.

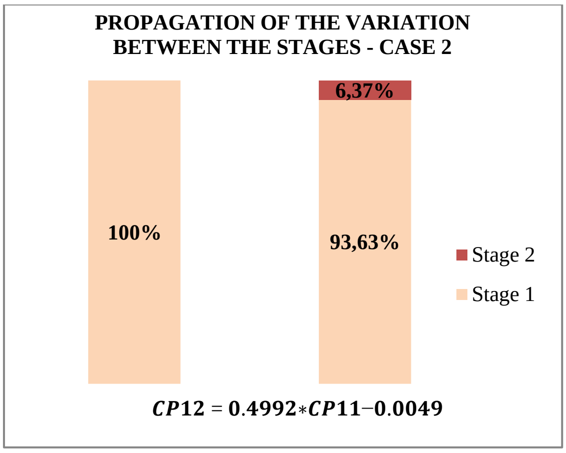

Figure 9: Variability between stages Simulation 2 case 2.

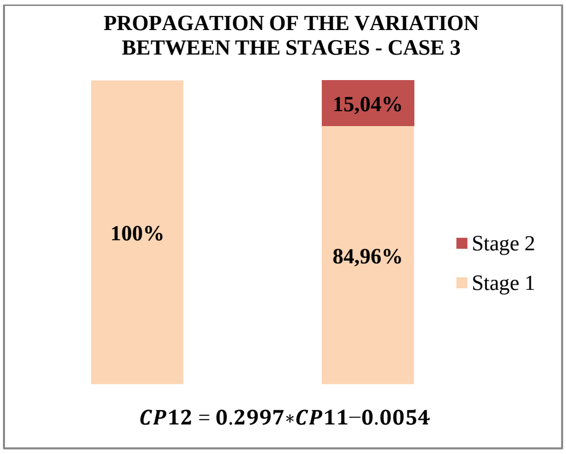

Figure 10: Variability between stages Simulation 2 case 3.

4.3 Results of Simulation 3

With the principal components analysis, we obtained the values of variability, maximum and minimum provided by the principals components in stage 1, according to Table 5.

Table 5: Percentage of Variance explained by the CPs simulation 3 Stage 1.

% CP1	% CP2	% CP3	%CP4
66.70%	19.20%	10.60%	3.50%

The Table 6 gives the values of the variability explained by each principal component in stage 2 in each situation.

Table 6: Percentage of Variance explained by the CPs simulation 3 Stage 2.

Case	% CP1	% CP2	% CP3	%CP4
1 (90%)	64.90%	17.90%	12.50%	4.70%
2 (50%)	49.50%	23.40%	17.60%	9.50%
3 (30%)	40.80%	26.00%	19.60%	13.60%

Analyzing the Figure 11 it is observed that the propagation of errors between the stages was 95,10%. In addition, Figure 12 shows that the variance transmission between the stages was of 88,55% of the first component in the case 2 of simulation 3 and in case 3 (Figure 13) propagation of variances was even smaller, about 75%.

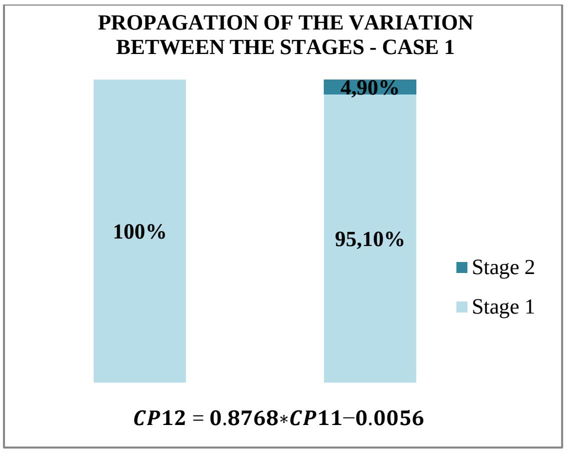

Figure 11: Variability between stages Simulation 3 case 1.

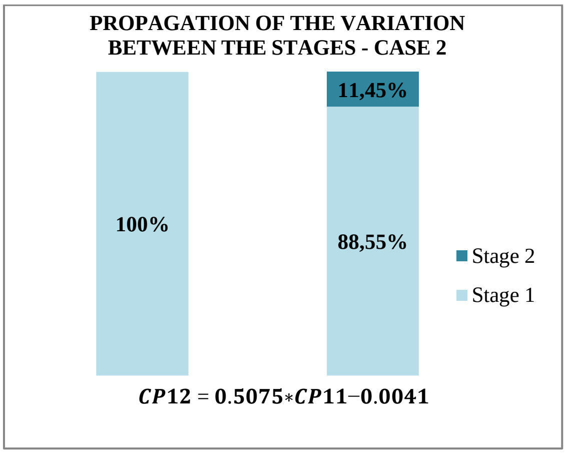

Figure 12: Variability between stages Simulation 3 case 2.

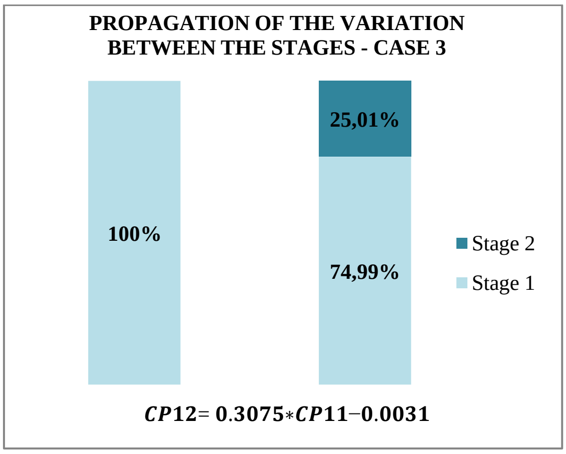

Figure 13: Variability between stages Simulation 3 case 3.

5. Conclusion

After analyzing the cases simulated, it is concluded that the autoregressive model in the Principal Components was able to capture the variation coming from the previous stage, showing promise for application in real cases. In addition, it was observed that the higher the correlation between the data, the greater the ability of the proposed model in reducing the number of variables, due to the high percentage of variance explained by the principals components, in particular for the first component. In consequence, the greater will be the model's capacity to predict the transmission of deviations between stages.

In summary, it can be inferred that the models autoregressivos in the Principals Components are valid to evaluate the behavior of the propagation of errors in cases with large or small volume of data. Through them, it is possible to identify which stage is responsible for the major variability of the process.

References

Ceglarek, D.J. Multivariate analysis and evaluation of adaptive sheet metal assembly systems. The University of Michigan. 1998.

Lawless, J.F., Mackay, R.J., Robinson, J.A. Analysis of variation transmission in manufacturing processes – part 1. Journal of quality technology, v.31, n.2, 1999.

Huang, Q., Zhou, S., Shi, J. Diagnosis of multi-operational machining processes through variation propagation analysis. Robotics and computer integrated manufacturing 18, p.233-239, 2002.

Camelio, J., Hu, S. J., Zhong, W. Diagnosis of multiple fixture faults in machining processes using designated component analysis. Journal of manufacturing systems, v.23, n.4, 2004.

Gaio, A. P.C., Sá, F.C. O uso da metodologia seis sigma em um processo da indústria automotiva por meio de um estudo de caso. Trabalho graduação em engenharia de produção. UERJ. Resende. 2004.

Liao, X.G., Wang, G. Employing fractals and FEM for detailed variation analysis of non-rigid assemblies. International Journal of Machine Tools & Manufacture 45, p.445-454, 2005.

Zhang, M., Djurdjanovic, D., Ni, J. Diagnosibility and sensitivity analysis for multi-station machining processes. International Journal of Machine Tools & Manufacture 47,p.646-657, 2007.

Yang , Z., Popov, A.A., Mcwilliam, S. Variation propagation control in mechanical assembly of cylindrical components. Journal of manufacturing systems 31, p 162-176, 2012.

Arif, F., Suryana, N., Hussin, B. Cascade quality prediction method using multiple pca+id3 for multi-stage manufacturing system. International conference on electronic engineering and computer science. Ieri procedia 4, p.201-207, 2013.

Johnson, R.A., Wichern, D.W. Applied multivariate statistical analysis. 6ª ed. Pearson prentice hall. 2007.

Montgomery, D.C. Introdução ao controle estatístico de qualidade. 4ªed. Gen/ltc. Rio de janeiro. 2013.

Mingoti, S.A. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. 2ª reimpressão. Editora UFMG. Belo horizonte. 2013.