Proceedings of the 13th Annual
International Conference on Industrial Engineering
Theory, Applications and Practice
Las Vegas, Nevada
September 7-10, 2008

# UNDERSTANDING AND REDUCING PROCESS VARIATION ON MANUFACTURING APPLICATIONS

**Luis G. Resendiz-Zamudio[1], Manuel R. Piña-Monarrez[2], and Pedro Perez-Villanueva[3]**

[1] Corporación Mexicana de Investigación en Materiales.
Ciencia y Tecnología 790, Fracc. Saltillo 400
C.P. 25290, Saltillo, Coah., Mexico.
Corresponding author's e-mail: Luis.Resendiz@mabe.com.mx

**Abstract:** Manufacturing processes are always changing along industries because of costs reduction, new supplier's evaluation, quality improvement or introduction of new products. Between these activities, acquiring new machinery and tooling is a usual activity. A common problem that most of industries must afford is how to make sure this tooling or machinery will met the quality and manufacturing specifications trough the program execution, referring to process output, one of the most important issues is process variation, it is required to understand, analyze and reduce it otherwise the product to be manufactured will have a high probability of being out of specifications. The consequences of an improper definition, execution and acceptance process will be reflected on production problems because of time or costs impact and poor quality results. There are some methods and statistical tools that help engineers analyzing data from any kind of manufacturing process in order to analyze it. Process capability is an indicator that determines if the process will be capable or not of complying specifications, measuring and comparing mean values as well as standard deviation, standardized Z value is used to estimate the probability of defective parts, usually expressed as parts per million. In this case we applied quality control tools and statistical analysis as Múltiple Linear Regressions, Anova analysis, Variation contributions to define, execute and implement a new process on an appliance manufacturing plant that met the target objective for its process capability of 3,4 parts per million defective planned at the beginning of the program.

## 1. INTRODUCTION

When manufacturing processes are being designed, some of the most important targets are: to have them in time, between the budget, fitting the facilities, space & labor and complying quality and manufacturing requirements. On this paper, the process under analysis is machinery for the manufacture of a stainless steel drum that contains clothes on household dryers. The objective is to analyze process variation along its variables and get final parts from the process between the dimensional specifications and demonstrate a process capability at least 3,4 ppm on the critical dimension and keep them under control once the machinery is installed and delivered in to the final manufacturing plant.

The drum manufacturing process consists on taking a stainless steel metal sheet from a stacker, cut and stamp some features, folding the sheet to get a cylinder, then a flanging and grooving process is radially done on the cylinder and finally by seaming both ends, a rear and front parts are assembled. In this process, there is a critical dimension which is the total depth, this is the most important for the analysis since this is the critical to quality dimension on the assembly. Input components variables, as well as process variables are considered for the variation analysis.

Analysis on single variables interaction and relation with the whole product system, tolerance loops of the system and its components and how the process and raw material specification will affect the final product to be manufactured is required. Special interest and analysis must be kept on understanding the process, its variation and which components or steps on the process contribute the most to product or process variation.

Since the target of the program is getting a machinery complying quality requirements, in which process capability is the one that will be calculated, in order to accept or reject the machinery, the statistical analysis performed trough this paper is focused on analyzing process variation and get its process capability. Multiple Analysis were performed to understand the process performance: Measuring method validation (R&R Gages), Process variation, Process capability, Percentage of variations given by individual components and Multiple linear regression.

Multivariate methods are used for data analysis, in this case the main objective of this work is to define the significant factors to the process response, consider the relation of every single variable with the full set of data and get the process variation given by each variable. Usually, the interest on multivariate analysis is to find the relationship between 1)Response variables, 2)Experimental units, 3)Both response variables and experimental units. (Johnson, 2000). The main analysis is done to define which components trough the assembly are contributing to process variation, this by taking single measurements for all components and process features used on each assembly, this in order to analyze them by using multiple linear regression and covariance matrix and its eigenvalues. The components that conform the assembly are: Front drum, wrapper, rear drum, all of them contributing to total height of the assembly. Also the process features will be analyzed in order to determine the combination of components and process contribution to total variation of the response. For manufacturing processes controlled there are some well stablished measures of performance based on statistical model of the process, the most common is the process capability (Hardt, D.,Sin-Siu, T. 2001) The result obtained during site machinery qualification is a probability lower than 3.4 defective parts per million, with a Z level of 6.

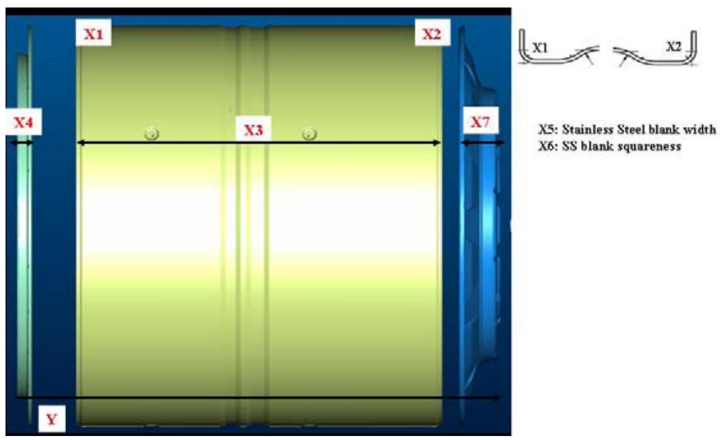


Figure 1. Assembly under analysis and its variables

## 2. PRODUCT AND PROCESS ANALYSIS

The process under analysis involves variables from the input components for the assembly, as well as process variables trough the manufacture of the product in this case the dryer household drum, in which the levels of the factors or variables involved cannot be easily adjusted. Regression methods are frequently used to analyze data were the interest focuses on modeling and exploring the relationship between the response variable and the independent variables (Montgomery, 2005). A multiple linear regression analysis is performed to analyze if there is a relation between the considered variables from the components and process, with the response dimension. In this case, 7 dimensional and process variables, and the response total depth of the drum are analyzed. X1: Front Flange forming, X2: Rear Flange forming, X3: Wrapper height, X4: Front Drum, X5: Stainless Steel blank width, X6: Stainless Steel blank squareness, X7: Rear Drum.

The relation between the process response and its predictor variables can be modeled with Multiple Linear Regression Equation (Montgomery, Peck & Vining, 2001) that in general form is given by:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad (1)$$

In which y is the process response, $\beta_i$ the regression coefficients and $x_i$ the variables considered for the analysis and $\varepsilon$ is the vector error. The Regression coefficients can be estimated by Ordinary Least Square (OLS) given by:

$$\hat{\beta} = (X'X)^{-1}X'y \qquad (2)$$

Data for the analysis is presented in Table 1. The corresponding Anova analysis for the regression by applying Minitab is presented in figure 2.

As we can see from figure 1, the fitted model does not presents collinearity problems all VIF less than 10 (Montgomery et. al 2002), the problem of multicollinearity arises in a regression problem whenever there is a linear dependency among independent variables (Gou & Fyfe, 2002 ) . The model significance (p-value less than 005), with a proper fit (R-Sq(adj)=80.8%) and low standard deviation (S=0.1345).

Table 1: Data for the regression analysis

| Y | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 672.32 | 3.93 | 6.81 | 598.21 | 22.09 | 608.33 | 0.508 | 33.86 |
| 672.10 | 3.90 | 6.70 | 598.18 | 22.05 | 608.41 | 0.203 | 33.95 |
| 672.37 | 4.01 | 6.76 | 598.35 | 22.11 | 608.33 | 0.305 | 34.03 |
| 672.43 | 4.14 | 7.88 | 598.58 | 22.31 | 608.33 | 0.406 | 34.30 |
| 673.13 | 4.14 | 6.66 | 599.12 | 22.41 | 608.33 | 0.000 | 34.32 |
| 672.67 | 4.18 | 6.90 | 598.79 | 22.24 | 608.33 | 0.051 | 34.21 |
| 672.74 | 4.11 | 6.86 | 598.84 | 22.26 | 608.33 | 0.356 | 34.05 |
| 672.81 | 4.21 | 6.88 | 598.54 | 22.22 | 608.33 | 0.254 | 34.05 |
| 672.07 | 4.21 | 7.07 | 598.16 | 22.00 | 608.33 | 0.000 | 33.90 |
| 672.75 | 4.27 | 7.01 | 598.79 | 22.22 | 608.33 | 0.102 | 34.02 |
| 672.56 | 4.27 | 7.13 | 598.61 | 22.27 | 608.33 | 0.406 | 34.18 |
| 672.70 | 4.37 | 7.07 | 598.51 | 22.25 | 608.33 | 0.127 | 34.00 |
| 672.67 | 4.25 | 7.12 | 598.25 | 22.22 | 608.46 | 0.406 | 33.99 |
| 673.30 | 4.62 | 7.02 | 599.01 | 22.43 | 608.33 | 0.051 | 34.41 |
| 672.72 | 4.12 | 6.92 | 598.52 | 22.10 | 608.33 | 0.203 | 34.12 |
| 672.81 | 3.93 | 6.59 | 598.70 | 21.83 | 608.41 | 0.152 | 34.11 |
| 672.70 | 3.90 | 6.81 | 598.61 | 22.08 | 608.41 | 0.076 | 34.04 |
| 672.47 | 3.90 | 6.70 | 598.61 | 22.30 | 608.33 | 0.000 | 34.21 |
| 672.64 | 4.24 | 6.88 | 598.74 | 22.01 | 608.33 | 0.152 | 34.01 |
| 672.84 | 3.93 | 6.72 | 598.82 | 22.29 | 608.33 | 0.203 | 34.20 |
| 673.10 | 4.00 | 6.73 | 598.98 | 22.37 | 608.33 | 0.102 | 34.36 |

```
The regression equation is
Y = -1027 +0.601A -0.349B +0.788C +0.147D +1.99E +0.344F +0.347G

Predictor     Coef  SE Coef      T      P  VIF
Constant   -1027.2    596.5  -1.72  0.109
A           0.6014   0.2130   2.82  0.014  1.7
B          -0.3489   0.1602  -2.18  0.048  2.1
C           0.7882   0.2100   3.75  0.002  3.7
D           0.1466   0.3143   0.47  0.649  2.5
E           1.9937   0.9171   2.17  0.049  1.3
F           0.3440   0.2446   1.41  0.183  1.5
G           0.3475   0.4045   0.86  0.406  4.2
S = 0.134495   R-Sq = 87.5%   R-Sq(adj) = 80.8%

Analysis of Variance
Source          DF       SS       MS      F      P
Regression       7  1.64857  0.23551  13.02  0.000
Residual Error  13  0.23515  0.01809
Total           20  1.88372

Source  DF   Seq SS
A        1  0.26763
B        1  0.30252
C        1  0.92248
D        1  0.01038
E        1  0.10468
F        1  0.02753
G        1  0.01335
```

Figure 2. Analysis for Multiple Linear Regression

Since the model fitted is significant to represent the process data, we could estimate the variation contribution that each variable has over the complete system, by determine the variance of each one, and dividing it by the total variance of the modeled system, this if there is no linear relation between the variables the process variables. In order to detect if there is a relation between the process variables, variance and covariance matrix is computed for the original set of data (Peña, 2002) by applying (3):

$$COV = E[(Xi-\mu_i)(Xj-\mu_j)] = \frac{\sum_{i=1}^{n}(Xi-\mu_i)(Xj-\bar{\mu}_j)}{n-1} \qquad (3)$$

If the matrix defined in (3), is not a diagonal matrix, then the major conclusion is that the set of process variable are not independent each other, and as a consequence the simple index variance of the $X_j$ variable divided by the total variance is not the best way to determine the contribution that each variable has over the total variance. Instead that, we should use the eigenvalues of the matrix defined in (3), for assign this contribution, due that those eigenvalues represent the behavior of the variance that all the process variables has over the modeled system, in the sense that precisely these eigenvalues are the axes of this system. For eigenvalues analysis, the mean have to be substracted from each of the data dimensions (Smith, L. 2002). The center data to construct the matrix defined in (3) are given in table 2 and its corresponding covariance matrix is given in table 3.

Since the covariance matrix given in table 3, it is not diagonal, we estimate the contribution of each variable by use the eigenvalues. The eigenvalues estimated using Mathcad corresponding to the 7 process variables under analysis and its corresponding contribution is presented in table 4.

Table 2. Centered data of the process variables

| X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|---|---|---|---|---|---|---|
| -0.195 | -0.105 | -0.405 | -0.103 | -0.017 | 0.315 | -0.250 |
| -0.225 | -0.215 | -0.435 | -0.143 | 0.059 | 0.010 | -0.160 |
| -0.115 | -0.155 | -0.265 | -0.083 | -0.017 | 0.112 | -0.080 |
| 0.015 | 0.965 | -0.035 | 0.117 | -0.017 | 0.213 | 0.190 |
| 0.015 | -0.255 | 0.505 | 0.217 | -0.017 | -0.193 | 0.210 |
| 0.055 | -0.015 | 0.175 | 0.047 | -0.017 | -0.142 | 0.100 |
| -0.015 | -0.055 | 0.225 | 0.067 | -0.017 | 0.163 | -0.060 |
| 0.085 | -0.035 | -0.075 | 0.027 | -0.017 | 0.061 | -0.060 |
| 0.085 | 0.155 | -0.455 | -0.193 | -0.017 | -0.193 | -0.210 |
| 0.145 | 0.095 | 0.175 | 0.027 | -0.017 | -0.091 | -0.090 |
| 0.145 | 0.215 | -0.005 | 0.077 | -0.017 | 0.213 | 0.070 |
| 0.245 | 0.155 | -0.105 | 0.057 | -0.017 | -0.066 | -0.110 |
| 0.125 | 0.205 | -0.365 | 0.027 | 0.110 | 0.213 | -0.120 |
| 0.495 | 0.105 | 0.395 | 0.237 | -0.017 | -0.142 | 0.300 |
| -0.005 | 0.005 | -0.095 | -0.093 | -0.017 | 0.010 | 0.010 |
| -0.195 | -0.325 | 0.085 | -0.363 | 0.059 | -0.041 | 0.000 |
| -0.225 | -0.105 | -0.005 | -0.113 | 0.059 | -0.117 | -0.070 |
| -0.225 | -0.215 | -0.005 | 0.107 | -0.017 | -0.193 | 0.100 |
| 0.115 | -0.035 | 0.125 | -0.183 | -0.017 | -0.041 | -0.100 |
| -0.195 | -0.195 | 0.205 | 0.097 | -0.017 | 0.010 | 0.090 |
| -0.125 | -0.185 | 0.365 | 0.177 | -0.017 | -0.091 | 0.250 |

Table 3. Variance & Covariances matrix

| | | | | | | |
|---|---|---|---|---|---|---|
| 0.689 | 0.461 | 0.239 | 0.215 | -0.033 | -0.057 | 0.120 |
| 0.461 | 1.475 | -0.241 | 0.196 | -0.023 | 0.302 | 0.096 |
| 0.239 | -0.241 | 1.512 | 0.462 | -0.073 | -0.355 | 0.649 |
| 0.215 | 0.196 | 0.462 | 0.449 | -0.044 | -0.036 | 0.316 |
| -0.033 | -0.023 | -0.073 | -0.044 | 0.027 | 0.016 | -0.033 |
| -0.057 | 0.302 | -0.355 | -0.036 | 0.016 | 0.469 | -0.143 |
| 0.120 | 0.096 | 0.649 | 0.316 | -0.033 | -0.143 | 0.461 |

Table 4.Eigenvalues and its contribution.

| Variable | Eigenvalue | % Cont to Variation |
|---|---|---|
| X1 | 0.287 | 5.65 |
| X2 | 0.217 | 4.27 |
| X3 | 0.021 | 0.41 |
| X4 | 0.083 | 1.63 |
| X5 | 0.48 | 9.44 |
| X6 | 1.809 | 35.59 |
| X7 | 2.186 | 43.01 |
| λj | 5.083 | λj/ Σλj |

The estimation of the contribution of each variables, reflected by the eigenvalues, we use the fact that the total variance is the sum of the eigenvalues, and the relation $\lambda_j / \Sigma\lambda_j$ to measure the proportion from the total variation explained by each $X_j$ variable. The contribution to process variation has been estimated on the variables listed on table 3, it shows how variation considering the variables and its correlation behaves. Single components variation has also been analyzed to focus on components variation reduction for the main contributors, this is variables x6 and x7. Eigenvalues as indicators of variability can be also used to reduce process variables for further analysis as Principal Component Analysis as method to reduce dimension reduction properties (Boil, R. 2002). If the correlated variables need to be monitored and controlled, a control chart for the eigenvalues can be used as proposed on (Piña and Zertuche, 2008).

Now the variability from the seven process variables has been analyzed, X6 and X7 result as main contributors, in order to monitor the process performance and keep it under control a multivariate process control chart can be used. A multivariate chart take the correlation into account in monitoring the mean vector or variance covariance matrix (Oktay, U., Cilan, C. 2001), once the variables are clearly identified and analyzed a multivariate control chart can be used to determine if the process is under control parameters and to detect when an output signal is present and its relation to the single variables analyzed.

Single components variation is also estimated by considering its individual Sum of Squares over the total from the same set of data shown on table 2 by using.

$$SS = \sum_{i}^{n} \frac{(x_i - \bar{x})^2}{n} = \sum_{i}^{n} x_i \tag{4}$$

Table 5. Contribution to total variation considering independent variables

| Variable | % Cont to Variation |
|---|---|
| X1 | 13.56 |
| X2 | 29.02 |
| X3 | 29.75 |
| X4 | 8.84 |
| X5 | 0.54 |
| X6 | 9.22 |
| X7 | 9.07 |

This values show how individual components or part of the process are affecting variation, actions were taken on components processes in order to reduce variation on the dimensional features specially on those with a considerable contribution as X1, X2 and X3. Single variation reduction for the input components to the process was performed on the dimensional features to make sure that the components used to manufacture the drum are between spec and complying with variation specified with the tolerances.

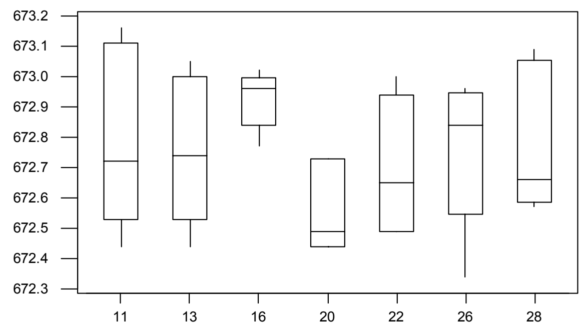


Figure 3. Boxplot for response measurement during machinery certification

Process capability analysis was performed for the response to determine if it complies the target, boxplot of response measurements trough the machinery certification runs are shown on figure 3. Process capability analysis was performed on those measurements and the Z level obtained is 6 over a long run production including labor shifts changes, different batches of raw material letting the process to get normal variation during the manufacture of the drum so the probability of parts out of specification on parts per million are 3.4.

## 3. CONCLUSIONS

Process variation can be analyzed considering both independent or correlated data, the option of considering there is no relation between the variables is helpful to clearly understand individual variations so they can be reduced focusing on one single feature. For multivariate process, variation trough the process should be first analyzed by its correlations, otherwise biased conclusion from the analysis will be achieved. By analyzing variation with eigenvalues estimated from variance and covariance matrix, the correlation among the variables is also included, another method to analyze a process when there are several variables involved is trough eigenvalues and eigenvectors in order to estimate main contributors to variation, then dimension of the data variables can be reduced and easily analyzed with those that result with higher eigenvalues. Techniques as Principal Component Analysis or Factor Analysis can be used.

## 4. REFERENCES

Boik, R. (2002). Lecture Notes: Classical Multivariate Analysis.
Gou, Z., Fyfe, C. (2002) A Canonical Correlation Neural Network for Multicollinearity and Functional Data
Hardt, D., Sin-Siu, T. (2001). Cycle to Cycle Manufacturing Process Control
Johnson, D. (2000). Métodos Multivariados Aplicados al Análisis de Datos, Thomson. ISBN: 968-7529-90-3.
Mongomery, D. (2005). Diseño y Análisis de Experimentos, Limusa Wiley. ISBN: 968-18-6156-6.
Mongomery, D., Peck, E., Vining, G. (2002) Introducción al Análisis de Regresión Lineal. Cecsa. ISBN: 0-471-31565-6.
Oktay, U., Cilan, C., (2001). Multivariate Statistical Process Control Methods and New Approaches. 6th TQM World Congress.
Peña, D. (2002). Análisis de Datos Multivariantes, Mc Graw Hill. ISBN: 84-481-3610-1
Smith, L. (2002). A Tutorial on Principal Components Analysis
Piña, M.R. and Zertuche, F. (2008). A Note on Confidence Intervals Estimation for the Eigenvalues *Proceeding of the 13th International journal of Industrial Engineering Applications and Practice*