

ROBUST DESIGN IN GENERALISED LINEAR MODELS FOR IMPROVING THE QUALITY OF POLYURETHANE SOLES

A. Mares^{1,2*} & J. Dominguez³

¹Department of Industrial Engineering Postgraduate/ ²Industrial- Electromechanical Area
CIATEC, A.C./ Universidad Tecnológica de León, Guanajuato, México
amares.picyt@ciatec.mx

³Department of Mathematics and Statistics
CIMAT, Aguascalientes, México
jorge@ciamat.mx

ABSTRACT

In a process that manufactures polyurethane soles by casting, a number of problems lead to different types of defects on the sole, causing significant economic losses for the company. In order to improve the product quality and decrease the number of defects, this study conducts an experimental design in the context of robust design. Since the response variable is binary, the statistical analysis was performed using generalised linear models. The operational methodology reduced the percentage of defects, while combining the experimental technique and control systems to achieve superior quality and a consequent reduction in costs.

OPSOMMING

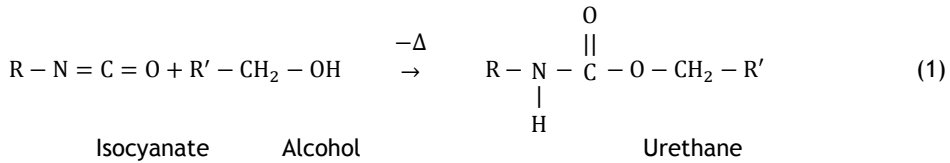
'n Groot aantal probleme word in 'n poliuretaansoolgietproses ervaar en dit lei tot verskillende defekte in die sool. Hierdie defekte veroorsaak noemenswaardige verliese vir die maatskappy. Om die produk gehalte te verbeter en die aantal defekte te verminder word 'n eksperimentele robuuste ontwerp uitgevoer. Aangesien die reaksieveranderlike binêr is, is veralgemeende lineêre modelle vir die statistiese analise gebruik. Die operasionele metodiek het die persentasie defekte verminder en die eksperimentele tegniek en beheerstelsels gekombineer om sodoende 'n hoër gehalte produk en 'n vermindering in kostes te verkry.

¹ The author was enrolled for a PhD (Industrial Engineering) degree in the Department of Industrial Engineering Postgraduate, CIATEC-CONACYT.

* Corresponding author

1 INTRODUCTION

The manufacturing process of polyurethane soles is carried out in rotary machines with 60 moulds, into which a liquid mixture of isocyanate components and liquid Polyol resins is poured. In addition to the above components, a component is added that gives colour to the material. The polyurethane-forming reaction occurs between a polymer isocyanate and an alcohol, as shown in Equation 1:



The case study presented here is the sole model 'Ucrania', which has presented a significant number of defective parts since its introduction. To highlight the impact of the failure, the production data was analysed over a period of six months, during which 20,858 pieces of a total of 47,872 were defective. This meant that 43.57 per cent of the production was defective, which at a cost of 74 Mexican pesos per piece accounts for a loss of approximately 1.5 million Mexican pesos. The vast majority of defects that occur on the sole are qualitative, and relate mainly to failures in its visual presentation. The literature provides a detailed list of defects that occur in the process of polyurethane, and their possible causes [1].

The process includes a noise factor, which is the environmental temperature during the morning and evening shifts. The response variable is binary because the quality inspectors rate the sole as either defective or not, based on pre-defined criteria. For this reason, the use of quality engineering methods under the Robust Parameter Design (RPD) is contemplated, allowing us to design experiments to make the process more efficient, improve quality, and reduce cost. Robinson et al. [2] provide a review of the application of the design RPD.

Response surface methodology (RSM) is a technique used in experimental design to optimise a process. Model-building with RSM requires that the response has a normal probability distribution. One area of great interest is when this assumption is not fulfilled, in particular with the problem of making soles; the random response variable has a binomial or Bernoulli distribution. Given this scenario, it is appropriate to use generalised linear models (GLMs); a variety of examples can be found in Hamada and Nelder [3], Myers and Montgomery [4], and Myers et al. [5].

The purpose of this work is to develop the methodology to construct the response surface model for binomial response in the manufacturing process of polyurethane soles in the context of robust design and generalised linear models. Once the design had been completed, the statistical analysis was performed. Finally, improvements that were made to the process were also indicated.

2 METHODOLOGY

The main part of the sole in which occur defects in the process of the polyurethane sole is shown in Figure 1; this is very important because it represents the final view of the product.

In research papers about improving the quality of polyurethane soles [6-8], it was observed that eight factors have a significant impact on product quality. These factors and their selected levels are shown in Table 1.



Figure 1: Part of the sole in which defects occur

Table 1: Relevant control and noise factors for this case-study

	Low	High	
Air pressure	0.5 Bar	1 Bar	Control
Time drain fluid (Flow)	55 grams/second	70 grams/second	Control
Mould height	15 cm	14 cm	Control
Stroke	110	90	Control
Polyol/Isocyanate ratio	0.9	0.92	Control
Mould temperature	36 °C	45 °C	Control
Air pressure over flow	1.8 Bar	2 Bar	Control
Environmental temperature	22 °C	33 °C	Noise

The above factors are quantitative. Some of these are regulated by the operator from the control panel, such as time of fluid drain (flow), stroke, polyol/isocyanate ratio, and mould temperature. The air pressure and air pressure over the flow are part of the painting process. Mould height can be adjusted manually in each mould. The noise factor is the environmental temperature.

Previous sampling in mould temperatures showed that it is a significant source of variation. By design, the machine model used in the company only has a pyrometer and thermocouple, which control the 60 moulds of the machine; this implies that the moulds handle an average temperature, and not the real temperature that is indicated in the pyrometer. Based on this analysis, it was determined that the variation of this element could be reduced through the application of an improved control system. One goal of the study is to generate information that may justify the investment in a control method based on individual pyrometers with a lower resolution, in order to allow a greater regulation of the temperature in the moulds, and therefore, a reduction of variation in temperature.

The operational methodology, which relates to the statistical procedures to be used for the analysis of the information obtained in the experiment, is presented in Figure 2. The process begins with an engineering analysis to determine the factors that could have any effect on the response. Next, the combined orthogonal array selection arises [9], which includes the control and noise factors. The experiment is then performed with these factors. In this process, the response received is binary, due to the classification of each sole as being either defective or non-defective. The evaluation is performed with the classification criteria of the quality staff. At the end, the number of defective soles determines the percentage of defects counted.

GLMs are appropriate to study the binary responses. The proposed plan was to build the model and obtain response models for the mean and variance. With these models, the goal is to obtain the best level combination from each factor in the process, which in turn produces the lowest number of defects.

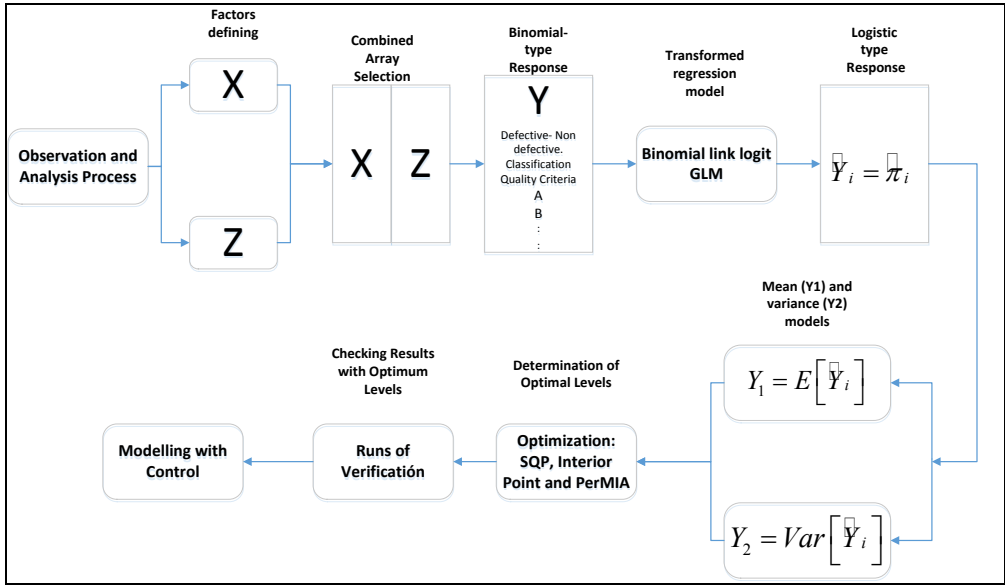


Figure 2: Operational methodology

2.1 Selection of the combined array:

The scheme of the combined array includes the control and noise factors in one orthogonal array [9]. Wu and Hamada [10] provided another reference for selecting the designs that provided more information on the effect of each factor and the interactions among them. The general linear model is shown in Equation 2.

$$Y_{xz} = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\mathbf{\Delta}\mathbf{z} + \varepsilon \quad (2)$$

where $\mathbf{x}' = (x_1, \dots, x_k)$ with k control factors, $\mathbf{z} = (z_1, \dots, z_q)$ q noise factors, β_0 is the constant, and $\beta' = (\beta_1, \dots, \beta_k)$ and $\gamma' = (\alpha_1, \dots, \alpha_q)$ are vectors of parameters. $\mathbf{B} = (\beta_{11}, \dots, \beta_{1k}, \beta_{k1}, \dots, \beta_{kk})$ and $\mathbf{\Delta} = (\alpha_{11}, \dots, \alpha_{1q}, \dots, \alpha_{k1}, \dots, \alpha_{kq})$ are the matrices of second order parameters, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. For the mean and variance models, the conditional expectation and variance are calculated from Equation 2 as follows:

$$E(Y_{xz}) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} \quad (3)$$

$$Var(Y_{xz}) = \sigma_z^2(\delta + \mathbf{x}'\mathbf{\Delta})'(\delta + \mathbf{x}'\mathbf{\Delta}) + \sigma^2 \quad (4)$$

In the process, the \mathbf{z} noise terms are considered as random variables with zero mean and variance σ_z^2 ; the term σ^2 is estimated with the mean square error of the regression model. The term is $(\delta + \mathbf{x}'\mathbf{\Delta})' = \mathbf{I}(\mathbf{x})$, which is defined as the vector of partial derivatives of the response Y_{xz} , with respect to \mathbf{z} , i.e., $\mathbf{I}(\mathbf{x}) = \partial Y_{xz} / \partial \mathbf{z}$.

The array selected for the experiment is a fractional factorial design 2_{IV}^{8-4} . To assign columns of the seven control factors and the noise factor, the design required 16 runs and had the following generators: $E = BCD$, $F = ACD$, $G = ABC$, and $H = ABD$. The design allowed the estimation of seven double-interactions. It is of particular interest to know the effect of noise control interactions that will serve for the time of optimisation. Because of this, care must be taken in how the assignment of the elements is given to the columns, which is according to Table 2.

Because of the factorial design, 2_{IV}^{8-4} is equivalent to Taguchi's $L_{16}(2^{15})$ design; the desired interactions can be obtained from the use of interaction tables [11]. According to the way in which the factors were assigned to the columns, it is possible to know four control-noise interactions that include the environmental temperature:

Table 2: Combined orthogonal array and allocation of factors to the columns:

	Air pressure		Environmental temperature	Time drain fluid (Flow)	Mould height	Stroke	Polyol/ isocyanate ratio	Mould temperature	Air pressure over flow
Run order	A	X1	B Z1	C X2	D X3	E X4	F X5	G X7	H X6
1	-1		-1	-1	-1	-1	-1	-1	-1
2	1		-1	-1	-1	-1	1	1	1
3	-1		1	-1	-1	1	-1	1	1
4	1		1	-1	-1	1	1	-1	-1
5	-1		-1	1	-1	1	1	1	-1
6	1		-1	1	-1	1	-1	-1	1
7	-1		1	1	-1	-1	1	-1	1
8	1		1	1	-1	-1	-1	1	-1
9	-1		-1	-1	1	1	1	-1	1
10	1		-1	-1	1	1	-1	1	-1
11	-1		1	-1	1	-1	1	1	-1
12	1		1	-1	1	-1	-1	-1	1
13	-1		-1	1	1	-1	-1	1	1
14	1		-1	1	1	-1	1	-1	-1
15	-1		1	1	1	1	-1	-1	-1
16	1		1	1	1	1	1	1	1

1. Environmental temperature with mould temperature.
2. Environmental temperature with polyol/isocyanate ratio.
3. Environmental temperature with stroke.
4. Environmental temperature with time drain fluid flow.

The other three interactions are control-control interactions:

5. Time drain fluid flow with mould temperature.
6. Mould temperature with stroke.
7. Air pressure with Air pressure overflow.

The rationale for using a two-level arrangement is based on costs and resources; it should be noted that when experiments with binomial responses are performed, large runs must be made. This implies that the process should be left to run, and if there is an inadequate level selection of the factors, this can cause an uncontrolled growth in the number of defects and the consequent cost of quality that this implies.

2.2 Logistic regression models

In order to find the best conditions to solve the problem of defects in polyurethane soles, we use the logistic regression model. Such models fall in the context of GLMs [12]. From the essential aspects of the manufacturing process, we need to raise the issue of robust design and control systems. We propose the features of the model, for which we propose that the response variable takes values from 0 to 1. In this framework, it is necessary to find equivalent equations to Equations 3 and 4. Supposing the model is:

$$y_i = \mathbf{x}'_i \boldsymbol{\beta} + \varepsilon_i \tag{5}$$

where $\mathbf{x}'_i = [1, x_{i1}, x_{i2}, \dots, x_{ik}]$, $\boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2, \dots, \beta_k]$ and the response variable y_i takes the values from 0 to 1. It is assumed that the response variable y_i is a Bernoulli random variable with probability distribution: $P(y_i = 1) = \pi_i$ if $y_i = 1$, and $P(y_i = 0) = 1 - \pi_i$ if $y_i = 0$. The expected value of y_i is $E(y_i) = 1(\pi_i) + 0(1 - \pi_i) = \pi_i$, since $E(\varepsilon_i) = 0$. Thus from Equation 5, it follows that $E(y_i) = \mathbf{x}'_i \boldsymbol{\beta} = \pi_i$.

From this description, a variance of y_i is obtained, i.e., $\sigma_{y_i}^2 = \pi_i(1 - \pi_i)$. It shows that the mean and variance of the response variable are related, which is a property of the logistic model [11].

In summary, the logistic regression model is in the form of the Bernoulli distribution, and the variable y_i is independent in each trial and expected value:

$$\pi = \frac{\exp(\beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\Delta\mathbf{z})}{1 + \exp(\beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\Delta\mathbf{z})} = \frac{1}{1 + \exp[-(\beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\Delta\mathbf{z})]} \quad (6)$$

The model parameters in Equations 5 and 6 are estimated by the maximum likelihood method. Equation 6 is re-written as follows:

$$g(\pi) = \log\left(\frac{\pi}{(1-\pi)}\right) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\Delta\mathbf{z} \quad (7)$$

Thus the maximum likelihood function is expressed by:

$$l(\boldsymbol{\pi}; \mathbf{y}) = \sum_{i=1}^n y_i \log\left(\frac{\pi_i}{1 - \pi_i}\right) + m_i \log(1 - \pi_i)$$

For estimation of the equivalent response surfaces from Equations 3 and 4, it is assumed that the noise factors vector has an expectation of 0, and this leads to:

$$E(\widehat{Y_{xz}}) = \frac{1}{1 + \exp[-(\beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x})]} \quad (8)$$

$$Var(\widehat{Y_{xz}}) = \frac{1}{1 + \exp[-(\sigma_z^2(\delta + \mathbf{x}'\Delta)'(\delta + \mathbf{x}'\Delta) + \sigma^2)]} \quad (9)$$

2.3 Optimisation schemes

The schemes for optimising the RPD have evolved in various methods, such as mean square error minimisation, fuzzy optimisation, maximum likelihood function, simulation, and nonlinear programming methods such as the simplex method of Nelder-Mead, among others [13-14]. For the proposed problem, the objective function is the minimisation of the variance, while it is intended that the (desired) mean response is zero (the % defective). An optimisation method that meets these characteristics is based on constrained nonlinear programming:

$$\begin{aligned} & \text{Min } Var[\widehat{Y_{xz}}] \\ & \text{subject to } E[\widehat{Y_{xz}}] = 0 \\ & \quad \quad \quad x_i > 0 \end{aligned}$$

Furthermore, the scheme requires that it restrict the area of solutions within the levels at which the experiment is coded. The algorithm of Sequential Quadratic Programming (SQP) [15] adequately meets these objectives. The SQP algorithm starts from the problem of constrained equality:

$$\begin{aligned} & \min f(x) \\ & \text{subject to } c(x) = 0 \end{aligned}$$

The $Var[\widehat{Y_{xz}}]$ function is used instead of $f(x)$, and the function $E[\widehat{Y_{xz}}]$ is used instead of the restriction function $c(x)$. The Interior Point algorithm [16] also suitably meets the optimisation approach for the optimisation problem from the objective function:

$$\begin{aligned} & \min f(x) \\ & \text{s. a } h(x) = 0 \\ & \quad \quad \quad g(x) \leq 0 \end{aligned}$$

$Var[\widehat{Y_{xz}}]$ is used instead of $f(x)$, and $E[\widehat{Y_{xz}}]$ is used instead of constraint function $g(x)$.

2.4 Modelling with control

One of the latest trends in RPD research is the integration with control systems. Miller and Wu [17], Joseph [18], and Dasgupta et al. [19] have developed algorithms to analyse and optimise the systems that include the use of control systems for reducing variation in some quality characteristic if it is justified on a cost-benefit basis.

Joseph [18] proposes the use of a performance measure from Taguchi's dynamic signal-to-noise ratio, dividing initially the Z noise into an online controlled noise Q and random noise R; the Taguchi's signal-to-noise ratio will be a function of the control factors X, controlled noise R, and random noise Q:

$$SR(\mathbf{X}, \mathbf{Q}, \mathbf{R}) = \frac{\beta^2(\mathbf{X}, \mathbf{Q}, \mathbf{R})}{\sigma^2(\mathbf{X}, \mathbf{Q}, \mathbf{R})} \quad (10)$$

The optimal solution for the levels of control variables X can be obtained by minimising the inverse of the expected value from the signal-to-noise ratio (10). $E\{1/SR(\mathbf{X}, \mathbf{Q})\}$ terms in Equation 10 can be rewritten as:

$$\beta^2(\mathbf{X}, \mathbf{Q}, \mathbf{R}) = E^2[\beta(\mathbf{X}, \mathbf{Q}, \mathbf{R})]$$

$$\sigma^2(\mathbf{X}, \mathbf{Q}, \mathbf{R}) = E[V(\mathbf{X}, \mathbf{Q}, \mathbf{R})] + Var[\beta(\mathbf{X}, \mathbf{Q}, \mathbf{R})]$$

Therefore, signal-to-noise ratio in Equation 10 can be expressed as:

$$SR(\mathbf{X}, \mathbf{Q}, \mathbf{R}) = \frac{\beta^2(\mathbf{X}, \mathbf{Q}, \mathbf{R})}{\sigma^2(\mathbf{X}, \mathbf{Q}, \mathbf{R})} = \frac{E^2[\beta(\mathbf{X}, \mathbf{Q}, \mathbf{R})]}{E[V(\mathbf{X}, \mathbf{Q}, \mathbf{R})] + Var[\beta(\mathbf{X}, \mathbf{Q}, \mathbf{R})]} \quad (11)$$

PM is the measure of system performance; the objective is to minimise the inverse of the signal-to-noise ratio in Equation 11, thereby minimising

$$PM = E \left\{ \frac{E[V(\mathbf{X}, \mathbf{Q}, \mathbf{R})] + Var[\beta(\mathbf{X}, \mathbf{Q}, \mathbf{R})]}{E^2[\beta(\mathbf{X}, \mathbf{Q}, \mathbf{R})]} \right\} \quad (12)$$

The optimal values for the levels of the control factors (X) are obtained.

3 RESULTS

Experimental runs were conducted during work shifts in the company; the process was allowed to run normally while leaving the control factors fixed. For the noise (environmental temperature), runs were made during both the day shift and the night shift, where extreme temperatures occur. The results obtained are shown in Table 3.

The last column shows the percentages obtained in the experiment, which is the response of a proportion binomial type; from these values, we proceed to obtain the logistic model.

3.1 Obtaining the logistic regression model

The regression model was obtained using Statgraphics Centurión software. The columns of the desired interactions were added to the orthogonal array 2^{8-4}_{IV} ; the model estimates are obtained through the maximum likelihood method. Originally, 15 elements in the regression model were obtained: eight main elements (seven control and one noise), three control-by-control interactions, and four noise-by-control interactions. There were elements that are not significant in the model, particularly the control-by-control interaction X2*X7. The selection of the significant elements was performed using the step-back algorithm of selection by elimination [20]. The reduced model is shown in Table 4, and the deviation analysis table is shown in Table 5.

Table 3: Orthogonal array and values obtained in the experiment

Order	Run	A	B	C	D	E	F	G	H	Defects	Totals	% Def
1	1	-1	-1	-1	-1	-1	-1	-1	-1	119	212	0.56132
2	2	1	-1	-1	-1	-1	1	1	1	85	295	0.28814
3	3	-1	1	-1	-1	1	-1	1	1	39	192	0.20313
4	4	1	1	-1	-1	1	1	-1	-1	35	130	0.26923
5	5	-1	-1	1	-1	1	1	1	-1	61	286	0.21329
6	6	1	-1	1	-1	1	-1	-1	1	48	171	0.28070
7	7	-1	1	1	-1	-1	1	-1	1	67	174	0.38506
8	8	1	1	1	-1	-1	-1	1	-1	84	247	0.34008
9	9	-1	-1	-1	1	1	1	-1	1	76	266	0.28571
10	10	1	-1	-1	1	1	-1	1	-1	71	224	0.31696
11	11	-1	1	-1	1	-1	1	1	-1	34	295	0.11525
12	12	1	1	-1	1	-1	-1	-1	1	103	284	0.36268
13	13	-1	-1	1	1	-1	-1	1	1	68	306	0.22222
14	14	1	-1	1	1	-1	1	-1	-1	118	272	0.43382
15	15	-1	1	1	1	1	-1	-1	-1	59	203	0.29064
16	16	1	1	1	1	1	1	1	1	28	274	0.10219

Table 4: Reduced estimated regression model

Parameter	Estimate	Standard Error	Odds Ratio Estimated
CONSTANT	-0.954439		
Z1	-0.190845	0.0380451	0.82626
X3	-0.145302	0.0384132	0.864761
X4	-0.221633	0.0375987	0.801209
X5	-0.185119	0.0380134	0.831005
X7	-0.341925	0.0380342	0.710402
X6	-0.103138	0.0374876	0.902003
X4*X7	0.131183	0.0383578	1.14018
X5*Z1	-0.100796	0.0380655	0.904117
X2*Z1	0.142873	0.0380923	1.15358
X1*X6	-0.100534	0.0375118	0.904355

Table 5: Analysis of deviance: Reduced estimated regression model

Source	Deviance	Df	P-Value
Model	238.551	10	0.0000
Residual	4.95749	5	0.4211
Total (corr.)	243.508	15	

The deviation explained by the model is 97.9641 per cent, and the adjusted value is 88.9295 per cent. The selection algorithm eliminated two major controlling factors (X1 and X2) and three interactions. Among the interactions that remained in the model are these two factors. The deviation percentages and adjusted percentages are high, as the P-value in the 'analysis of deviance' table is less than 0.05. There is a statistically significant relationship between the variables, with a confidence level of 95.0 per cent.

Table 6 shows the likelihood ratio tests for the reduced model; it can be observed that all elements are significant with P values lower to 0.05, so it is not desirable to remove any more elements from the model.

Table 6: Likelihood ratio tests for the reduced model

Factor	Chi-Squared	Df	P-Value
Z1	25.5058	1	0.0000
X3	14.2965	1	0.0002
X4	35.2203	1	0.0000
X5	23.9625	1	0.0000
X7	81.3237	1	0.0000
X6	7.5633	1	0.0060
X4*X7	11.7184	1	0.0006
X5*Z1	7.05806	1	0.0079
X2*Z1	14.1347	1	0.0002
X1*X6	7.19041	1	0.0073

After obtaining the most suitable model for the analysis, the fit of the model is verified. In Figure 3, the subplot on the left shows that the residuals do not follow any apparent triangular shape, or funnel pattern; this indicates that there is no direct relationship between the predicted values and the residuals. Therefore, the residuals meet both independence and homoscedasticity. The subplot on the right shows the normality test. The normal probability plot shows that the points have an adequately fit around the straight line, and that there are no points particularly away at the ends; this shows that we have a good model fit.

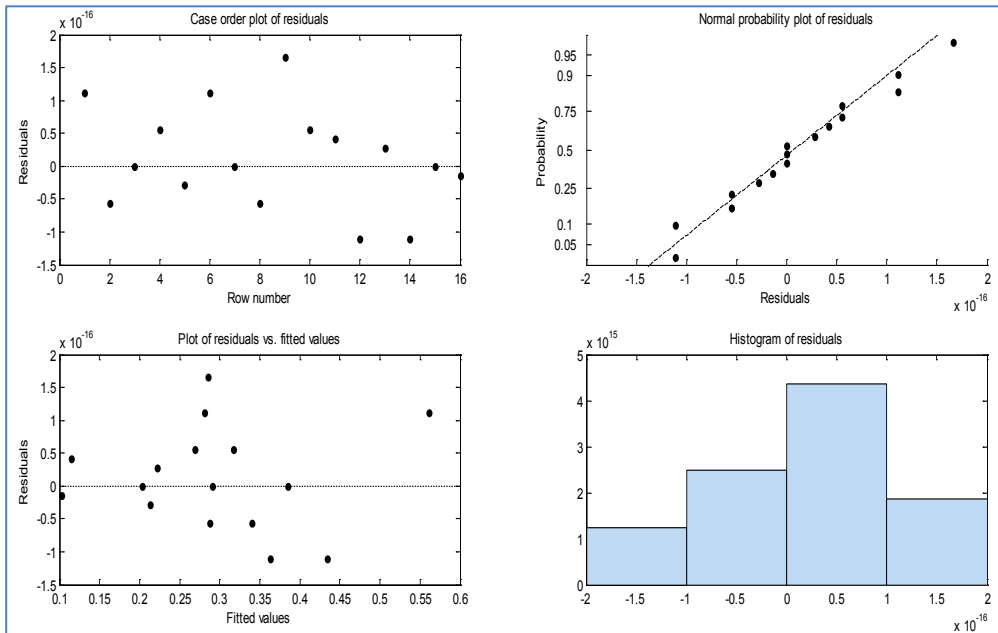


Figure 3: Residual analysis for the model

3.2 Models for the mean and the variance

Once the overall regression model is obtained, it is expressed generally as in Equation 2:

$$Y_{xz} = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\Delta\mathbf{z} + \varepsilon$$

$$\hat{Y}_{xz} = \frac{1}{1 + \exp\left[\frac{-0.953217 - 0.145302x_3 - 0.221633x_4 - 0.185119x_5 - 0.103138x_6 - 0.341925x_7}{-0.190845z_1 + 0.131183x_4x_7 - 0.100534x_1x_6 - 0.100796x_5z_1 + 0.142873x_2z_1}\right]} \quad (13)$$

From the full model, the mean model is calculated from the expectation (Equation 3) and the variance (Equation 4):

Model for the mean $E(Y_{xz}) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'\mathbf{B}\mathbf{x}$

$$E(\hat{Y}_{xz}) = \frac{1}{1 + \exp\left[\frac{-0.953217 - 0.145302x_3 - 0.221633x_4 - 0.185119x_5 - 0.103138x_6}{-0.341925x_7 + 0.131183x_4x_7 - 0.100534x_1x_6}\right]} \quad (14)$$

Model for the variance $Var(Y_{xz}) = \sigma_z^2(\delta + \mathbf{x}'\mathbf{\Delta})'(\delta + \mathbf{x}'\mathbf{\Delta}) + \sigma^2$

$$Var(\hat{Y}_{xz}) = \frac{1}{1 + \exp[(-0.190845 + 0.142873x_2 - 0.100796x_5)^2 + 1.00484]} \quad (15)$$

Note: To estimate the value of σ^2 , the MSE value was used in the ANOVA table by the weighted least squares method.

3.3 Parameter optimisation

The parameters were optimised with the use of the optimisation toolbox in the MATLAB software. Once the functions were declared, limits were proposed for the case; these were defined as -1 and 1. The results obtained are:

Table 7: Parameter estimation results

Variable	Level
x1	1.00000
x2	-1.00000
x3	-1.00000
x4	-1.00000
x5	-1.00000
x6	-1.00000
x7	-1.00000

The algorithm converges to an infeasible point in the tenth iteration, with a function value of 0.25748; the behaviour of the iterations in the algorithm is shown in Figure 4.

The next step is optimised by the interior point algorithm, with the results converging to an infeasible point. The run for the algorithm yields the following results:

Table 8: Interior point algorithm results

Variable	Level
x1	1.00000
x2	-0.02600
x3	-1.00000
x4	-1.00000
x5	-1.00000
x6	-1.00000
x7	-1.00000

The value of the function the interior point algorithm generates a result of 0.26627, which is similar to the result of SQP. The values are consistent on the levels in all the factors except factor X2 (i.e., flow). Figure 5 shows the behaviour of the iterations. Given that the two optimisation schemes presented infeasible solutions, no verification runs were conducted with those levels. Instead, we proceeded to perform the optimisation through the analysis with the PM approach.

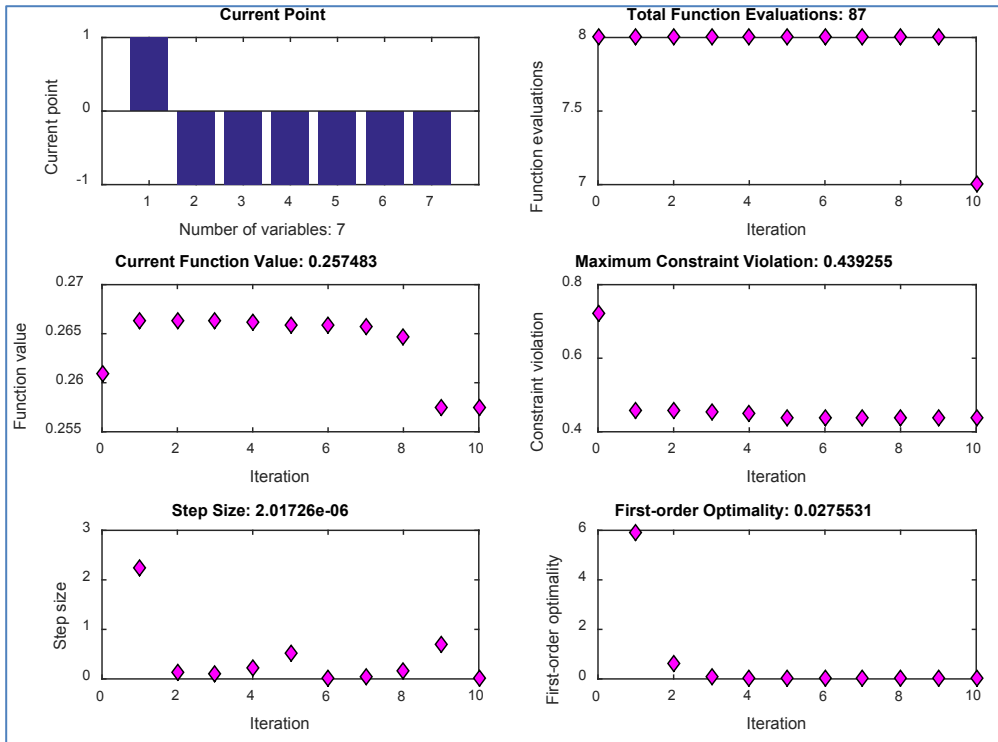


Figure 4: Optimisation plots for the SQP algorithm

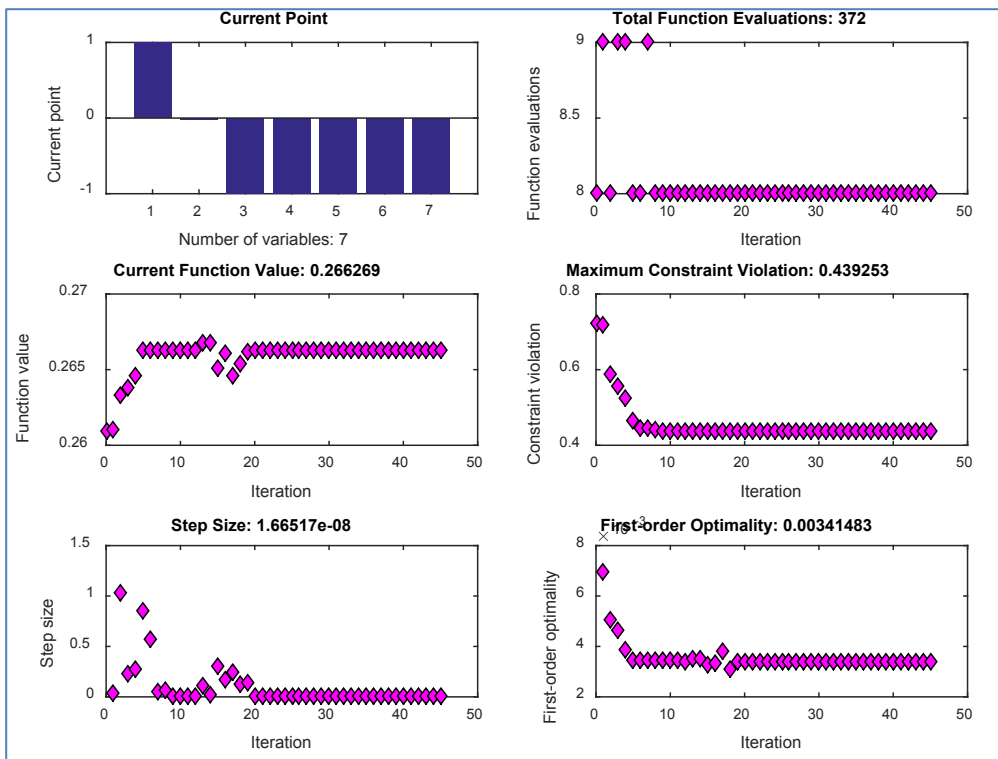


Figure 5: Optimisation plots for the interior point algorithm

3.4 Modelling without and with control by the PM scheme

One way to reduce the existing variation in the moulds' temperature is by implementing various thermocouples and pyrometers. This is done in order to make continuous corrections to the values of temperature, and this way to reduce the variation. The selected factor for the reduction in variation is X7 (i.e., mould temperature); the X5 factor (i.e., Polyol/isocyanate ratio) is also a factor that presents a significant variation in the model, and can be considered to be a random noise factor. According to Equation 10, the performance measure used is $PM = E \left\{ \frac{E[V(X, Q, R)] + Var[\beta(X, Q, R)]}{E^2[\beta(X, Q, R)]} \right\}$, where the elements are defined as:

$$E[V(X, Q, R)] = Var(\hat{Y}_{xz}) = \frac{1}{1 + \exp[(-0.190845 + 0.142873x_2 - 0.100796x_5)^2 + 1.00484]}$$

For the mean model, an artificial variable is added (q_7) to the elements containing the controlled variable X7:

$$E^2[\beta(X, Q, R)] = E^2(\hat{Y}_{xz}) = \left[\frac{1}{1 + \exp \left[\begin{array}{l} -0.953217 - 0.145302x_3 - 0.221633x_4 - 0.185119x_5 - 0.103138x_6 \\ -0.341925(x_7 + q_3) + 0.131183x_4(x_7 + q_7) - 0.100534x_1x_6 \end{array} \right]} \right]^2$$

The expected variance of the regression model takes the values of the constants in the regression model, corresponding to the element to be controlled (X7) and its interactions multiplied by a variance of σ_7^2 , which can be found in practice.

$$Var[\beta(X, Q, R)] = \frac{1}{1 + \exp[(-0.341925)^2\sigma_7^2 + (0.131183x_4)^2\sigma_7^2]}$$

The complete model for measuring performance is:

$$PM = \frac{\left[\frac{1}{1 + \exp[(-0.190845 + 0.142873x_2 - 0.100796x_5)^2 + 1.00484]} \right] + \left[\frac{1}{1 + \exp[(-0.341925)^2\sigma_7^2 + (0.131183x_4)^2\sigma_7^2]} \right]}{\left[\frac{1}{1 + \exp \left[\begin{array}{l} -0.953217 - 0.145302x_3 - 0.221633x_4 - 0.185119x_5 - 0.103138x_6 \\ -0.341925(x_7 + q_7) + 0.131183x_4(x_7 + q_7) - 0.100534x_1x_6 \end{array} \right]} \right]^2} \quad (16)$$

Assuming that (q_7) is normally distributed with zero mean and variance $\sigma_7^2 = 1$, and minimising Equation 14, the following levels are obtained for the factors:

Table 9: Results from Equation (14)

Variable	Level
x1	1.00000
x2	-1.00000
x3	1.00000
x4	-1.00000
x5	1.00000
x6	1.00000
x7	1.00000

With a function value of 0.709406, a minimum was found that satisfies the constraints in iteration 23. This was found from the model:

$$Y = \beta(X, Q, R)C/C_0 \quad (17)$$

The control law can be obtained for the model; by substituting the values in the model, we have

$$Y = \left[\frac{1}{1 + \exp \left[\begin{array}{l} -0.953217 - 0.145302(1) - 0.221633(-1) - 0.185119(1) - 0.103138(1) \\ -0.341925((1) + q_7) + 0.131183x_4(1 + q_7) - 0.100534(1)(1)x_6 \end{array} \right]} \right] \frac{C}{C_0}$$

$$= \left[\frac{1}{1 + \exp[-1.476419 - 0.210742q_7]} \right] \frac{C}{C_0}$$

The control law is given by:

$$C = TC_0 (1 + \exp(-1.476419 - 0.210742q_7)) \quad (18)$$

where T is the target value of the temperature. Without control, the performance measure is:

$$PM = \frac{\left[\frac{1}{1 + \exp[(-0.190845 + 0.142873x_2 - 0.100796x_5)^2 + 1.00484]} \right] + \left[\frac{1}{1 + \exp[(-0.185119)^2\sigma_6^2 + (-0.341925)^2\sigma_7^2 + (0.131183x_4)^2\sigma_7^2]} \right]}{\left[\frac{1}{1 + \exp \left[\begin{array}{l} -0.953217 - 0.145302x_3 - 0.221633x_4 - 0.185119x_5 - 0.103138x_6 \\ -0.341925x_7 + 0.131183x_4x_7 - 0.100534x_1x_6 \end{array} \right]} \right]^2} \quad (19)$$

By optimising Equation 19 with the same levels as the control equation, a reduction in the variation of 20.12 per cent with the use of control can be estimated. These were the levels used for the verification runs.

3.5 Verification runs

For the verification runs, the results of the PM method were used. The main objective of the study was to reduce the percentage of defects generated in the process. The verification runs were made with optimal levels obtained according to the process optimisation. The first run was performed during the morning shift with a production of 214 pieces; of these, 39 were defective, or 18.22 per cent. The second verification run was performed during the night shift with a production of 204 pieces; of these, 43 were defective, or 21.08 per cent. Compared with the initial ratio, which was 43.57 per cent of the pieces being defective, this was a reduction in the defective ratio of about 21.00 per cent after applying the methodology. This reduction resulted in significant cost savings, due to better quality. The methodology showed a marked improvement in the quality of the soles.

4 DISCUSSION AND CONCLUSION

The proposed methodology had a positive impact on the manufacturing process of the sole through a significant reduction of the proportion of defective pieces, and an improvement in their quality. It is important to emphasise that the experiment was conducted online - a novel situation that allowed the production of one shift in each treatment. The results were analysed using GLMs. The analysis and statistical optimisation of the process allowed the reduction of flaws in the process, thereby achieving better production. This strategy integrated a control system.

The flow works best at a low level; so it is desirable that the liquid is emptied at a rate of 55 grams/sec in order for the material to expand appropriately in the mould and so that no material splashes out of the mould. The mould's height works best at 16 cm, which allows the material to be distributed, especially at the front of the sole. The stroke should be 110 mm, and 70 per cent of the material should be poured at the front of the sole for better distribution of the material in the mould. The ideal polyol/isocyanate ratio for the job is 0.92; this favours the physical properties of the material and the expansion reaction of the gas molecules in forming the foam. The mould temperature should preferably be 45°C to promote the adequate curing process of the material. As noted in the experimental phase, the implementation of an improved control system should reduce the variation of the temperature in the moulds, thus improving control of the process and the quality of the polyurethane sole. The factors of air pressure and air pressure overflow which are used in the painting process, will work best at its high levels; a higher air pressure favours a

uniform painting, which improves the appearance of the sole and its presentation to the consumer.

The results of the levels in the optimisation show congruence with those observed during the process. An increase in defects, such as a rigid sole, was observed when the moulds were handled at low temperatures. When the height of the mould is set at high levels, it tends to leave parts with 'bites'. When a short stroke is handled when pouring the mixture, it can mark the sole because the material does not fit adequately. Although the pouring can be done more quickly, it can cause the material to splash, which results in some weight being lost and thus in poor quality soles with marks.

In industry, many processes require the use of control systems. In this project, a control system was applied in order to obtain a higher quality and better productivity in the process based on a reduction in variation. This strategy entailed obtaining a lower production cost and a reduced re-work due to poor quality. Here, the control was applied to the temperature of the moulds; this is a critical factor in the process, due to the large volume of production. When implementing the analysis and optimisation strategy, the objectives were achieved. Finally, we conclude that the impact on the economic state of the company included reduced costs and fewer defects.

Reports on the Ucraina model's quality over a period of six months from its introduction were based on a production of 47,872 pieces with a total of 20,858 defective pieces - that is, 43.57 per cent. The approximate cost of the sole is 74 Mexican pesos, representing an estimated loss of 1,543,492 Mexican pesos. According to the results of the confirmatory tests, if the production of the first six months were projected on to the scale of the defective percentage from the confirmatory runs, the savings would be projected to be 898,038.13 Mexican pesos, while savings from the confirmatory runs during the night shift were projected to be 796,721.03 Mexican pesos. The resulting levels of defects in the process adopted for all models like the Ucraina sole remained low in subsequent production.

Statgraphics Centurion was the software used to obtain the algorithms for the maximum likelihood method of GLMs, and MATLAB software was used for the optimisation algorithms.

REFERENCES

- [1] Herrington, R. and Hock, K. 1997. *Flexible polyurethane foams*. Midland, MI: The Dow Chemical Company.
- [2] Robinson, T.J., Borror, C.M. and Myers, R.M. 2004. Robust parameter design: A review. *Quality and Reliability Engineering International*, 20(1), pp. 81-101.
- [3] Hamada M. and Nelder, J.A. 1997. Generalized linear models for quality-improvement experiments. *Journal Quality Technology*, 29(3), pp. 292-304.
- [4] Myers, R.H. and Montgomery, D.C. 2009. *Response surface methodology*. USA: Wiley.
- [5] Myers, R.H., Montgomery, D.C., Vining G.G. and Robinson T.J. 2010. *Generalized linear models with applications in engineering and the sciences*. USA: Wiley.
- [6] Amaro, W., Diviani, L., Montorfano, D., Oberrauch, E., Depinto, G., Segalini, S., Levi, M. and Turri, S.. Controlling the shrinkage of polymers for customized shoe sole manufacturing. *International Journal of Computer Integrated Manufacturing*, 17(7), pp. 633-644.
- [7] Park, C.C., Choi, W.S. and Lee, J.N. 2007. Effects of hardness and thickness of polyurethane foam midsoles on bending properties of the footwear. *Fibers and Polymers*, 8(2), pp. 192-197.
- [8] Li, K.W. and Chen, C.J. 2004. The effect of shoe soling tread groove width on the coefficient of friction with different sole materials, floors, and contaminants. *Applied Ergonomics*, 37(6), pp. 449-507.
- [9] Welch, W.J., Yu, T.K., Kang, S.M. and Sacks J. 1990. Computer experiments for quality control by parameter design. *Journal Quality Technology*, 22(1), pp. 15-22.
- [10] Wu, C.F.J. and Hamada, M. 2000. *Experiments: planning, analysis and parameter design optimization*. New York: Wiley- Interscience.
- [11] Phadke, M.S. 1989. *Quality engineering using robust design*. Cliffs, New Jersey: Prentice Hall.
- [12] McCullagh, P. and Nelder, J.A. 1989. *Generalized linear models*. 2nd edition, London: Chapman and Hall.

- [13] Mares, A. and Domínguez, J. 2013. Conditional expectation and variability in the industrial problem solution. *IIE Annual Conference Proceedings*, pp. 3338-3347.
- [14] Mares, A. and Domínguez, J. 2015. Robust design with dynamic features: Laboratory level case. *International Journal of Engineering Innovation and Research*, 4(4), pp. 628-635.
- [15] Nocedal, J. and Wright, S.J. 2006. *Numerical optimization*. Springer Science+Business Media, LLC.
- [16] Byrd, R.H., Gilbert, J.C. and Nocedal, J. 2000. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 89(1), pp. 149-185.
- [17] Miller, A.E. and Wu, C.F.J. 1996. Parameter design for signal-response systems: A different look at Taguchi's dynamic parameter design. *Statistical Science*, 11(2), pp. 122-136.
- [18] Joseph, V.R. 2003. Robust parameter design with feed-forward control, *Technometrics*, 45(4), pp. 284-291.
- [19] Dasgupta, T., Miller, A. and Wu, C.J.F. 2010. Robust design of measurement systems. *Technometrics*, 52(1), pp. 80-93.
- [20] Montgomery, D., Peck, E.A. and Vining, G.G. 2012. *Introduction to linear regression analysis*. 5th edition, USA: Wiley.