SOBRAPO

TUTORIAL FOR MIXTURE-PROCESS EXPERIMENTS
WITH AN INDUSTRIAL APPLICATION

Luiz Henrique Abreu Dal Bello[1] and Antonio Fernando de Castro Vieira[2*]

ABSTRACT. This article presents a tutorial on mixture-process experiments and a case study of a chemical compound used in the delay mechanism for starting a rocket engine. The compound consists in a three-component mixture. Besides the mixture components, two process variables are considered. For the model selection, the use of an information criterion showed to be efficient in the case under study. A linear regression model was fitted. Through the developed model, the optimal proportions of the mixture components and the levels of the process variables were determined.

Keywords: mixture experiments, process variable, linear regression, information criterion, optimization.

1 INTRODUCTION

Formulations of Mixture Experiments (ME) are commonly found in chemical, pharmaceutical and food industries, as well as in other industrial segments. In those experiments, the factors are proportions of the components in a mixture and the response is a variable that characterizes the quality of the product, assumed as a function of component proportion. In those experiments, the sum of component proportions always equals to one. In certain situations, there may be other factors, in addition to the mixture components, that affect the characteristics of the mixture. Such factors are named process variables, which are frequently included in the experiment as factorial designs. Therefore, we intend to determine not only the optimal proportions of the mixture components, but also the optimal levels of the process variables. The aim of this article is to provide a tutorial on Mixture-Process Experiments (MPE).

A chemical mix experiment will be presented, as part of a subsystem of a mechanism to delay the activation of a rocket engine, described in Vieira & Dal Bello (2006) and Dal Bello & Vieira (2008). The aim of the experiment is to find the ideal formulation, so that the expected burning time value is according to the design's specification, that is, 8 seconds. Such time is optimal for

---

*Corresponding author

[1]Centro Tecnológico do Exército (CTEx), Rio de Janeiro, RJ, Brasil. E-mail: dalbello@ctex.eb.br

[2]Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ, Brasil. E-mail: afcv@puc-rio.br

the rocket's range. As it will be seen, several formulations lead to a time of 8 seconds, being thus convenient to find a formulation that provides the lowest variance for future responses. The mixture components are Zarfesil, Ground Glass and Nitrocellulose. In addition to these mixture variables, other two variables that may also affect the characteristics of the mixture were also taken into consideration. The first is the granulometry of the mixture; the second variable consists of the insertion – or not – of an orifice for gas escape during the burn reaction, making this reaction always to occur at a pressure which is approximately the atmospheric pressure.

2   LITERATURE REVIEW

Cornell (2002) is the main reference on ME, being the Chapter 7 dedicated to MPE cases. In it, a comprehensive and detailed exposition can be found. Myers & Montgomery (2002) dedicate Chapters 12 and 13 to ME and MPE, thus comprising a good introduction to the topic. Piepel (2004) summarizes a survey related to mixture experiments for a period of 50 years, ranging from 1955 to 2004. Prescott *et al.* (2002) propose a quadratic model as an alternative to the models traditionally used in ME (Scheffé models). Cornell (2000), Cornell (2002, Chapter 6), Cornell & Gorman (2003) and Khuri (2005) carried out comparative studies between models that they named as slack-variable models and Scheffé models. Piepel (2007) compares the CSLM (Component Slope Linear Model) with the SLM (Scheffé Linear Model) and the CLM (Cox Linear Model). They conclude that the models SLM, CLM and CSLM are mathematically equivalent and provide the same statistics for a given ME. The differences lie in the interpretations of their coefficients.

Goos & Donev (2006) describe an algorithm to plan experiments in blocks involving mixtures. They show that, for restricted and unrestricted experimental regions, the resulting design of experiments is statistically more efficient than the options of experiments in blocks presented in the literature. Goos & Donev (2007) describe an algorithm to plan split-plot experiments in cases involving mixture and process variables. They use an optimization criterion for the choice of experimental points and show that it is preferable to spread the replications all over the experiment region, instead of concentrating them in central points.

Kowalski *et al.* (2002), Prescott (2004) and Sahni *et al.* (2009) analyzed the MPE modeling. Goldfarb *et al.* (2004a) propose the use of a plot method (variance dispersion plot) for MPE planning. The variance dispersion plot presents a visual way of assessing the variance properties of an MPE within the joint mixture and process area. That information may be used to select experiments with an acceptable variance profile.

Goldfarb *et al.* (2003), Goldfarb *et al.* (2004b) and Chung *et al.* (2007) consider the case where, in addition to the mixture components and process variables (controlled factors), there are uncontrolled factors in the productive process (noise variables), although they may be controlled in laboratory experiments. The authors address models that allow them to choose the controllable variable values (mixture and process) that make the process more robust in relation to the noise variables.

3   MIXTURE EXPERIMENTS

Consider x_i, the variables that represent the proportions of the q mixture components. Then:

$$\sum_{i=1}^{q} x_i = 1; \quad x_i \geq 0; \quad i = 1, \ldots, q \tag{1}$$

The restrictions presented in Equation (1) are shown in Figure 1, in the event of a mixture of two and three components.
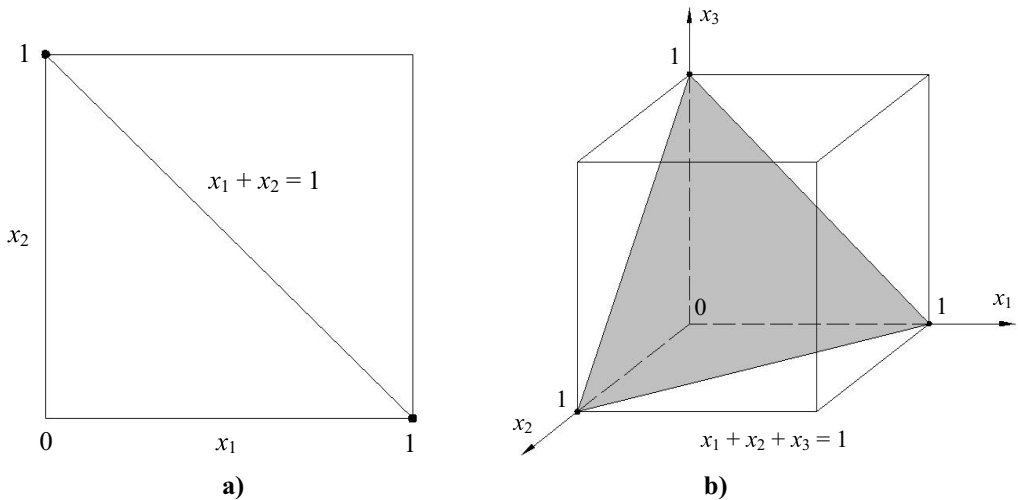
Figure 1 – Restricted factor space for mixtures with a) 2 components and b) 3 components.

The feasible region of the two-component mixture is represented by a line segment and, in case of a three-component mix, it is represented by a triangle. Generally, in case of three-component mixtures, the restricted experimental region is represented by using a 3-coordinate system, as shown in Figure 2. Each edge of the triangle corresponds to a binary mixture, and the vertices of the triangle correspond to the formulations with pure components. The possible ternary mixtures are located inside the triangle. In this case, only two dimensions are needed to represent the experiment. As each component is represented by a vertex, a geometric figure with three apexes and two dimensions – that is, an equilateral triangle – represents the restricted factor space of a ternary mixture.

In many ME there are limitations in the component proportion, making the experimental space a sub-region of the original space. Therefore, upper and/or lower limits are established in the proportions, represented as follows:

$$0 \leq L_i \leq x_i \leq U_i \leq 1 ; \quad i = 1, \ldots, q \tag{2}$$

where L_i is the lower limit and U_i is the upper limit of the component proportion i.

Figure 3 shows an experimental case with a three-component mixture with lower constraints in the proportions of the three components, and a case with four components with upper constraint in the proportion of only one component.
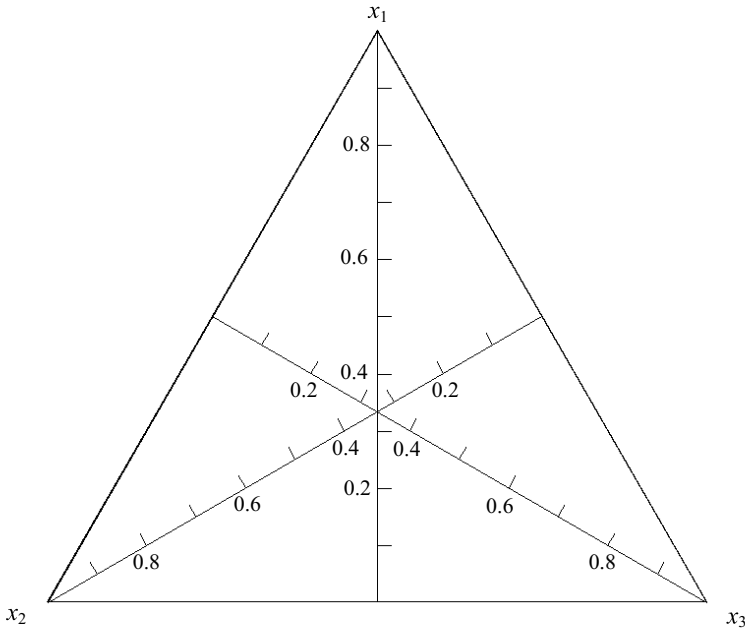

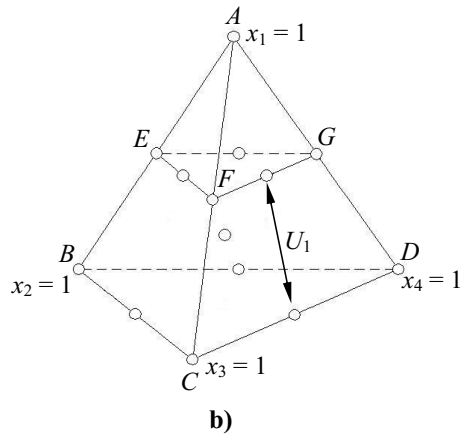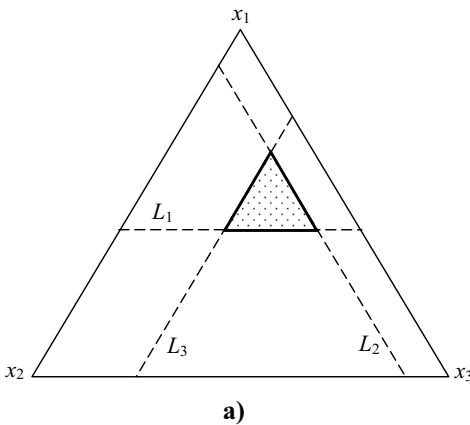
Figure 2 – 3-Coordinate System.

Figure 3 – a) lower and b) upper constraints in the proportion of components.

When the upper and lower limits for the proportions of one mixture are established, the experimental region is reduced to a sub-region of the original region. In these cases, the coordinates of the sub-regions may be redefined in terms of "pseudo"-components. The pseudocomponents are defined as a function of the original components and of one of their limits (lower or upper).

We then have two types of pseudocomponents: the L-pseudocomponents in relation to the lower limit, and the U-pseudocomponents in relation to the upper limit. According to Cornell (2002, p. 134), the main reason to use pseudocomponents is that it is usually easier to plan the experiment and fit the model. Myers & Montgomery (2002, p. 655) recommend the use of pseudocomponents to fit mixture models where there are component constraints, resulting in mild to high levels of multicollinearity. Generally speaking, a mixture model that uses pseudocomponents will have lower levels of multicollinearity than the same model with the original components.

The L-pseudocomponents are defined as:

$$v_i = \frac{x_i - L_i}{1 - L}; \quad i = 1, 2, \ldots, q \tag{3}$$

where $L = \sum_{i=1}^{q} L_i$.

In order to obtain original components (x_i), just use the Equation (3):

$$x_i = L_i + (1 - L) v_i \tag{4}$$

The U-pseudocomponents are defined as:

$$u_i = \frac{U_i - x_i}{U - 1}, \quad i = 1, 2, \ldots, q \tag{5}$$

where $U = \sum_{i=1}^{q} U_i$.

In order to calculate respective original components (x_i), just use the Equation (5):

$$x_i = U_i - (U - 1) u_i \tag{6}$$

Where there are upper and lower constraints simultaneously, the choice of using L-pseudocomponents, v_i, or U-pseudocomponents, u_i, depends on the shape of the experimental region. Where $(1 - L) < (U - 1)$, the choice if for L-pseudocomponents, and where $(1 - L) \geq (U - 1)$, the choice is for U-pseudocomponents (Cornell, 2002, p. 160).

4   MODEL FOR MIXTURE-PROCESS EXPERIMENTS

4.1   Models for Mixture Experiments

The models which are traditionally used in mixture experiments are Scheffé's canonical polynomials (Cornell, 2002). Scheffé's Quadratic Model is as follows:

$$Q(\boldsymbol{\beta}, \mathbf{x}) = \sum_{i=1}^{q} \beta_i x_i + \sum \sum_{i<j} \beta_{ij} x_i x_j \tag{7}$$

where the β's are the models' parameter coefficients. Note that this model does not have the intercept, as it is eliminated by a simplification originated from the basic limitation presented in Equation (1).

Scheffé's Cubic Model is as follows:

$$C(\boldsymbol{\beta}, = \mathbf{x}) = \sum_{i=1}^{q} \beta_i x_i + \sum \sum_{i<j}^{q} \beta_{ij} x_i x_j + \sum \sum \sum_{i<j<k}^{q} \beta_{ijk} x_i x_j x_k \tag{8}$$

$$+ \sum \sum_{i<j}^{q} \beta_{i-j} x_i x_j (x_i - x_j)$$

Scheffé's Special Cubic Model is the same as the Cubic Model, excluding the terms

$$\sum \sum_{i<j}^{q} \beta_{i-j} x_i x_j (x_i - x_j).$$

For the cases where there are high levels of multicollinearity, the use of slack-variable models is recommended (Cornell, 2002, Chapter 6). In a mixture of q components, being aware of the proportion of the first $(q-1)$ components, it is possible to determine also the proportion of component q, since $x_q = 1 - x_1 - x_2 - \cdots - x_{q-1}$. Therefore, x_q is defined as the slack variable, which will not be part of the composition of the model terms. In order to balance out the lack of the slack variable, an intercept is included to the model. Therefore, for a mixture of q components there are q possibilities of slack variables, that is, q different models with slack variable.

The Quadratic Model with slack variable:

$$Q(\boldsymbol{\beta}, \mathbf{x}) = \beta_0 + \sum_{i=1}^{q-1} \beta_i x_i + \sum \sum_{i \leq j}^{q-1} \beta_{ij} x_i x_j \tag{9}$$

The Cubic Model with slack variable:

$$C(\boldsymbol{\beta}, \mathbf{x}) = \beta_0 + \sum_{i=1}^{q-1} \beta_i x_i + \sum \sum_{i \leq j}^{q-1} \beta_{ij} x_i x_j + \sum \sum_{i \leq j \leq k}^{q-1} \sum \beta_{ijk} x_i x_j x_k \tag{10}$$

4.2   Model for Process Variables

An adequate model for r process variables $z_1, z_2, \ldots, z_r$, involving second-order terms is (Cornell, 2002):

$$Q(\boldsymbol{\alpha}, \mathbf{z}) = \alpha_0 + \sum_{l=1}^{r} \alpha_l z_l + \sum_{l=1}^{r} \alpha_{ll} z_l^2 + \sum \sum_{l<m}^{r} \alpha_{lm} z_l z_m \tag{11}$$

where the α's are the model's parameter coefficients for process variables. The experiment for the process variables may be a factorial design at two or more levels. In order to include terms with the variable z_j^2 in the model, it is necessary an experiment with at least three levels of each process variable and a total number of points sufficient to fit and test the model. In order to fit a model without the variable z_j^2, considering only the main effects of the process variables and the interactions among them, only two levels of each process variable are necessary (Myers & Montgomery, 2002).

4.3   Mixture Model Including Process Variables

The experimental design is established through a combination of the mixture variables with the process variables. In addition, any model for the mixture may be combined with the model for the process variables, in order to represent mixture-process-type problems. Consider $f(x)$ the for the mix and $g(z)$ the model for the process variables. Then, the additive model combined for the mixture-process case is (Prescott, 2004)

$$C(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) + g(\mathbf{z}) \tag{12}$$

For example, the additive combined model, which includes Scheffé's cubic model for the mixture presented in Equation (8) and the reduced quadratic model, considering only the main effects of the process variables and the

$$
\begin{aligned}
C(\boldsymbol{\gamma}, \mathbf{x}, \mathbf{z}) = {} & \sum_{i=1}^{q} \gamma_i^0 x_i + \sum \sum_{i<j} \gamma_{ij}^0 x_i x_j + \sum \sum \sum_{i<j<k}^{q} \gamma_{ijk}^0 x_i x_j x_k \\
& + \sum \sum_{i<j}^{q} \gamma_{i-j}^0 x_i x_j \left(x_i - x_j \right) + \sum_{l=1}^{r} \gamma_0^l z_l + \sum \sum_{l<m}^{r} \gamma_0^{lm} z_l z_m
\end{aligned}
\tag{13}
$$

where the γ's are the parameters for the mixture's combined model including process variables. The lower indexes of γ refer to mixture variables, whereas the upper ones refer to process variables. Observe that the combined model presented in Equation (13) does not have the intercept originated from the model for the process variables. The intercept is eliminated from the combined model, since it has a linear dependence with the terms $\gamma_i^0 x$, due to the constraint presented in Equation (1). Those additive models do not take into account the effects of process variables on the linear and non-linear properties of the mixture components. Alternative models were suggested, with the introduction "crossed" terms in $f(\mathbf{x})$ and $g(\mathbf{z})$. The complete cross-over model is

$$C(\mathbf{x}, \mathbf{z}) = f(\mathbf{x}) \times g(\mathbf{z}) \tag{14}$$

The combined multiplicative model that includes Scheffé's cubic model for mixtures and the reduced quadratic model for the process variables. This combined model is represented as follows:

$$
\begin{aligned}
C(\boldsymbol{\gamma}, \mathbf{x}, \mathbf{z}) = {} & \sum_{i=1}^{q} \gamma_i^0 x_i + \sum \sum_{i<j}^{q} \gamma_{ij}^0 x_i x_j + \sum \sum \sum_{i<j<k}^{q} \gamma_{ijk}^0 x_i x_j x_k \\
& + \sum_{l=1}^{r} \left[\sum_{i=1}^{q} \gamma_i^l x_i + \sum \sum_{i<j}^{q} \gamma_{ij}^l x_i x_j + \sum \sum \sum_{i<j<k}^{q} \gamma_{ijk}^l x_i x_j x_k \right] z_l \\
& + \sum_{l<m}^{r} \left[\sum_{i=1}^{q} \gamma_i^{lm} x_i + \sum \sum_{i<j}^{q} \gamma_{ij}^{lm} x_i x_j + \sum \sum \sum_{i<j<k}^{q} \gamma_{ijk}^{lm} x_i x_j x_k \right] z_l z_m
\end{aligned}
\tag{15}
$$

However, the combined models that include the additive and multiplicative model terms simultaneously may also be considered.

5 D-OPTIMAL DESIGN OF EXPERIMENTS

When there are constraints in the proportions of the mixture components, the use of experiments generated according to some optimization criterion is recommended (Cornell, 2002, p. 400). Cornell (2002) describes four alphabetic optimization criteria (A-optimal, D-optimal, G-optimal and V-optimal) for the choice of the experimental points. Such criteria are based on the optimization of some information matrix function $(\mathbf{W'W})$, where $\mathbf{W}$ is a matrix $(n \times p)$, whose elements are the proportions of the mixture components (x_i), the levels of the process variables (z_i) and the x_i and z_i functions (such as interactions), where n is the number of observations in the experiment and p is the number of model parameters. The general combined mixture model with inclusion of process variables is represented in matrix form:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{16}$$

For n observations, $\mathbf{y}$ is a vector $(n \times 1)$ of the observations, $\boldsymbol{\beta}$ is a vector $(p \times 1)$ of the coefficients and $\boldsymbol{\varepsilon}$ is a vector $(n \times 1)$ of the random errors. In the classical linear regression model, $\boldsymbol{\varepsilon}$ is considered with multivariate normal distribution, that is, $\boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \mathbf{I}\sigma^2\right)$.

The estimation vector for the coefficients is $\hat{\boldsymbol{\beta}} = \left(\mathbf{W'W}\right)^{-1}\mathbf{W'y}$, the variance-covariance matrix is $\mathrm{var}(\hat{\boldsymbol{\beta}}) = \sigma^2\left(\mathbf{W'W}\right)^{-1}$. The prediction value for the response at point $\mathbf{w}$ ($\mathbf{w'}$ is a matrix line $\mathbf{W}$) is $\hat{\mathbf{y}}(\mathbf{w})$ and its variance is

$$\mathrm{var}\left[\hat{y}\left(\mathbf{w}\right)\right] = \sigma^2\mathbf{w'}\left(\mathbf{W'W}\right)^{-1}\mathbf{w} \tag{17}$$

The statistical software Design-Expert$^{\circledR}$, developed and distributed by the company Stat-Ease, uses the D-optimal criterion for the choice of the experimental points. Myers & Montgomery (2002) define the D-optimal criterion using the moment matrix $\mathbf{M} = \mathbf{W'W}/n$. According to these authors, a D-optimal experiment is the one that makes the determinant of the moment matrix, $|\mathbf{M}|$, to be maximized. They show that the determinant of the moment matrix has the following property:

$$|\mathbf{M}| = \left|\mathbf{W'W}\right|\big/n^p \tag{18}$$

With all that, and assuming that the error are normally distributed, independent and with constant variance the determinant $\left|\mathbf{W'W}\right|$ is inversely proportional to the squared volume of the confidence region on the regression coefficients. When $\left|\mathbf{W'W}\right|$ is small, it means that the inverse of $\left|\mathbf{W'W}\right|$ is large and, therefore, the volume of the confidence region is large; thus, the estimated $\boldsymbol{\beta}$ is not regarded as good (Myers & Montgomery, 2002, Appendix 7).

Consequently, the D-optimal experiment design is the one that minimizes the volume of the confidence ellipsoid on $\boldsymbol{\beta}$, which is achieved by maximizing the determinant $\left|\mathbf{W'W}\right|$. By observing Equation (18), we may infer that maximizing the determinant $\left|\mathbf{W'W}\right|$ is equivalent to maximizing the determinant of the moment matrix, $|\mathbf{M}|$.

6   DELAY MIX EXPERIMENT

The chemical mix under study is part of a subsystem of a mechanism to delay the ignition of a rocket engine, which evolved from the R&D stage in laboratory scale to be currently produced at industrial scale. The aim of the experiment is to find the ideal formulation, so that the expected burning time value is according to the project's specification, that is, 8 seconds. Such time is optimal for the maximization of the rocket's range. As it will be seen, several formulations lead to a time of 8 seconds, being thus convenient to find a formulation that provides the lowest variance for future responses.

The mixture components are Zarfesil (x_1), Ground Glass (x_2) and Nitrocellulose (x_3), with the following limitations in the proportions: $0.77 \leq x_1 \leq 0.81$, $0.14 \leq x_2 \leq 0.18$ and $0.05 \leq x_3 \leq 0.07$.

As presented in Section 3, where there are simultaneous upper and lower constraints in the mixture components proportions, the L-pseudocomponents are chosen if $(1 - L) < (U - 1)$. In the case concerned, the choice was for the L-pseudocomponents, as $(1 - L) = 0.04$ and $(U - 1) = 0.06$.

Table 1 shows the experimental design of the delay mix, as well as the experiments' responses.

Table 1 – Reduced experiment of D-Optimal delay mix with L-pseudocomponents.

Run	v_1	v_2	v_3	Time (s)
1	0	1	0	12.9
2	0.5	0.5	0	7.3
3	0	1	0	13.7
8	1	0	0	7.2
12	0.3125	0.3125	0.375	10.7
13	0.5	0.5	0	8.7
14	0.5	0	0.5	5.9
17	1	0	0	5.6
21	0.5	0	0.5	6.1
22	0.75	0	0.25	6.0
26	0	0.5	0.5	7.3
36	0.75	0	0.25	6.8
37	0	0.5	0.5	5.6

Considering the figures presented in Table 1 and using the backward method, the following Scheffé model is obtained:

$$\hat{y} = 6,47v_1 + 13,301v_2 + 5,67v_3 - 7,53v_1v_2 - 12,13v_2v_3 + 124,39v_1v_2v_3 \qquad (19)$$

Now, considering the terms from slack variable models v_1, v_2 and v_3 and the backward method, the following models are obtained, respectively:

$$\hat{y} = 6,47 - 0,70v_2 - 0,80v_3 + 99,06v_2v_3 + 7,53v_2^2 - 207,32v_2^2v_3 \tag{20}$$

$$\hat{y} = 13,30 - 14,30v_1 + 10,93v_3 + 58,11v_1v_3 + 7,40v_1^2 - 49,26v_3^2 - 77,42v_1^2v_3 \tag{21}$$

$$\hat{y} = 5,67 + 0,80v_1 - 4,50v_2 + 128,99v_1v_2 + 12,13v_2^2 - 248,79v_1^2v_2 \tag{22}$$

6.1 Delay Mix Experiment with Process Variables

In addition to the mixture variables, other two variables that may also affect the characteristics of the mixture were also taken into consideration. The first one is the granulometry (z_1), which may be regarded as a categorical process variable, and will be varied at two levels. The granulometry currently used in the production of the delay mix is the 20-30 (level [−1]); however, the granulometry 25-30 (level [1]) is believed to provide a reduction in the variability of the burning times. The second variable (z_2) may be regarded as a design variable, as it is a modification of the original delay mechanism design. The insertion of an orifice for gas exhaustion during the burn reaction was suggested, making this reaction always to occur at a pressure which is approximately the atmospheric pressure. Currently, the burning reaction occurs in a confined environment, at a pressure under transient conditions. The insertion of an orifice for gas escape is expected to contribute to the reduction of response variability. That design variable may be seen as a categorical process variable, which will also be experimented in two levels (without orifice [−1] and with orifice [1]). Therefore, the problem of the delay mix may be addressed as a mixture experiment with three components, including two category process variables, which may be represented as $z_l = -1, 1; l = 1, 2$.

With the constraints in the mixture components, the resulting experimental region becomes a sub-region of the original experimental space. Under such conditions, Cornell (2002) recommends the use of some computer algorithm for the choice of experimental points according to some optimization criterion, arising out of some previously selected candidate points. The software Design-Expert® 7 offers the option of planning restricted mixture experiments in the proportions of the components and including process variables. For the choice of experimental points arising out of some previously selected candidate points, the software uses the D-optimal criterion, which was presented in Section 5. For the delay mix problem, Design-Expert® generated a D-Optimal experiment according to Table 2 and Figure 4, which also presented the results obtained in the random execution sequence in which the experiments were performed.

In Figure 4, each quadrant represents one possible combination of the process variables levels. These quadrants represent the plot of restricted experimental region of the mix variables in the 3-coordinate system (see Fig. 2), containing the experimental points.

Table 2 – Experiment of D-Optimal delay mix with L-pseudocomponents.

Run	v_1	v_2	v_3	z_1	z_2	Time (s)
1	0	1	0	1	1	12.9
2	0.5	0.5	0	1	1	7.3
3	0	1	0	1	1	13.7
4	1	0	0	1	−1	4.4
5	0	0.5	0.5	1	−1	6.1
6	0	1	0	−1	1	13.0
7	0.3125	0.3125	0.375	−1	−1	9.5
8	1	0	0	1	1	7.2
9	1	0	0	−1	−1	5.0
10	0	0.75	0.25	−1	−1	12.3
11	0	0.5	0.5	−1	−1	5.9
12	0.3125	0.3125	0.375	1	1	10.7
13	0.5	0.5	0	1	1	8.7
14	0.5	0	0.5	1	1	5.9
15	0.75	0	0.25	1	−1	5.6
16	0.5	0	0.5	1	−1	5.1
17	1	0	0	1	1	5.6
18	0.5	0	0.5	−1	−1	4.9
19	0	0.75	0.25	−1	−1	14.9
20	0.5	0.5	0	−1	−1	6.5
21	0.5	0	0.5	1	1	6.1
22	0.75	0	0.25	1	1	6.0
23	0.5	0	0.5	−1	1	5.9
24	0	0.75	0.25	−1	1	14.6
25	1	0	0	−1	1	7.9
26	0	0.5	0.5	1	1	7.3
27	0	1	0	−1	−1	13.5
28	0.6875	0.1875	0.125	−1	−1	5.3
29	0	1	0	−1	−1	12.7
30	0.3125	0.3125	0.375	1	−1	10.9
31	0	0.75	0.25	−1	−1	14.5
32	1	0	0	−1	−1	6.0
33	0.5	0.5	0	1	−1	7.3
34	0	0.5	0.5	−1	1	8.2
35	0.5	0.5	0	−1	1	8.7
36	0.75	0	0.25	1	1	6.8
37	0	0.5	0.5	1	1	5.6
38	0	1	0	1	−1	10.9
39	0.3125	0.3125	0.375	−1	1	12.2

$z_1 = -1; z_2 = 1$ $z_1 = 1; z_2 = 1$

$z_1 = -1; z_2 = -1$ $z_1 = 1; z_2 = -1$
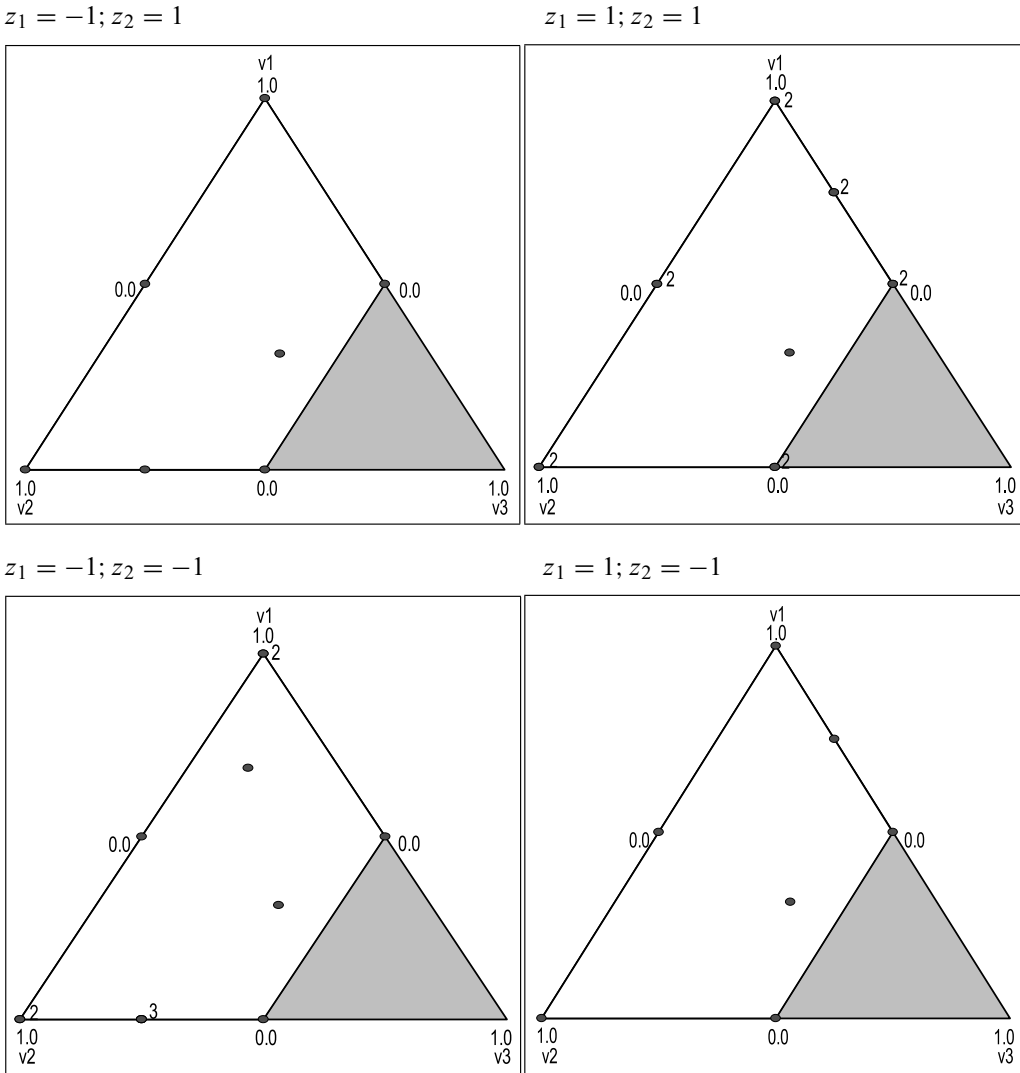

Figure 4 – Experiment of D-Optimal delay mix with L-pseudocomponents.

At first, a model using the stepwise, forward or backward model was fitted. The backward method provided a better model than those obtained through the forward and stepwise methods. The following model was fitted with the backward method.

$$
\begin{aligned}
\hat{y} = {} & 5{,}99v_1 + 12{,}51v_2 - 38{,}48v_3 - 6{,}47v_1v_2 + 87{,}34v_1v_3 \\
& + 78{,}55v_2v_3 + 0{,}94v_1z_2 + 0{,}57v_2z_2 - 58{,}94v_1v_3\,(v_1 - v_3) \\
& + 0{,}70v_2z_1z_2 - 3{,}51v_1v_2z_1z_2 - 3{,}50v_2v_3z_1z_2
\end{aligned}
\tag{23}
$$

Table 3 shows the t-Student test for the fitted model.

Table 3 – M1-Model Test.

Label	Estimate	Std. Error	t value	p-value
v_1	5.99	0.3132	19.14	< 0.0001
v_2	12.51	0.3228	38.76	< 0.0001
v_3	−38.48	3.2182	−11.96	< 0.0001
$v_1 v_2$	−6.47	1.5781	−4.10	0.0003
$v_1 v_3$	87.34	6.2773	13.91	< 0.0001
$v_2 v_3$	78.55	5.8740	13.37	< 0.0001
$v_1 z_2$	0.94	0.2355	3.98	0.0005
$v_2 z_2$	0.57	0.2430	2.33	0.0275
$v_1 v_3 (v_1 - v_3)$	−58.94	6.4119	−9.19	< 0.0001
$v_2 z_1 z_2$	0.70	0.3218	2.17	0.0390
$v_1 v_2 z_1 z_2$	−3.51	1.4379	−2.44	0.0216
$v_2 v_3 z_1 z_2$	−3.50	1.3644	−2.56	0.0163

The lack-of-fit test was not significant, with a p-value of 0.9336.

The ME data modeling may become very complex when the region of definition of the components is very limited. When the interval of a variable is substantially smaller than the interval of other variables, there may be high level of multicollinearity. This collinearity may turn the estimators of model coefficients unstable and very inflated. In this situation, the model fitting may not be simple, as the columns of **W** (covariates matrix) may be nearly linearly dependent. As a result, **W′W** may be nearly singular, and the least-squares estimators of the parameters in **β** may have great standard deviations and be highly correlated, besides the fact that the estimates depend on the precise location of the experimental points. In this context, stepwise, forward and backward selection of variables may result in arbitrary selection of variables that belong to the model (Harrell, 2001, p. 64). Therefore, the selection of the "best" model arising out of a set of candidate models may be very complex. One alternative is to use criteria based on the information theory. Once all the terms that may be part of the MPE model are known, we may then use model selection criteria based on Information Theory.

For model selection, Akaike (1973) developed the Akaike Information Criterion, *AIC*. However, Burnham & Anderson (2002) recommend the use of *AIC* only when $n/p \geq 40$, being n the number of observations and p the number of parameters in the model. For this experiment, $n = 39$. Consequently, *AIC* is not recommendable.

For the selection of models in the event of responses with normal distribution and small samples ($n/p < 40$), Hurvich & Tsai (1989) developed the *AIC* criterion.

$$AIC_c = \ln\left(\hat{\sigma}_p^2\right) + \frac{n+p}{n-p-2} \tag{24}$$

where $\hat{\sigma}_p^2$ is the estimator of maximum likelihood of the error variance.

A Matlab[®] routine was prepared for the case of delay mix of MPE, for the calculation and storage of *AIC* values. Among all of the models, the one to display the smallest AIC_c value was chosen (Model M2) (see Dal Bello, 2010 and Dal Bello & Vieira, 2011)

$$\hat{y} = 6{,}11v_1 + 5{,}22v_2 + 32{,}11v_3 + +0{,}59z_2 + 7{,}35v_2^2 - 2{,}58v_2z_1z_2$$
$$- 110{,}24v_3^3 - 53{,}48v_1v_3(v_1 - v_3) + 3{,}31v_2^2z_1z_2 \tag{25}$$

Table 4 shows the *t*-Student test for the Model M2.

<div align="center">Table 4 – M2-Model Test.</div>

Label	Estimate	Std. Error	*t* value	*p*-value
v_1	6.11	0.2564	23.83	< 0.0001
v_2	5.22	1.1037	4.73	0.0001
v_3	32.11	1.7490	18.36	< 0.0001
z_2	0.59	0.1176	5.06	< 0.0001
v_2^2	7.35	1.2196	6.03	< 0.0001
$v_2z_1z_2$	−2.58	0.8205	−3.15	< 0.0001
v_3^3	−110.24	7.1329	−15.46	0.0037
$v_1v_3(v_1 - v_3)$	−53.48	5.7271	−9.34	< 0.0001
$v_2^2z_1z_2$	3.31	0.9799	3.37	0.0021

The lack-of-fit test was not significant, with a *p*-value of 0.9775.

As it will be seen below, the Model M2, presented in Equation (25), is better than the Model M1, presented in Equation (23).

Myers & Montgomery (2002) recommend the use of studentized residuals to check the following assumptions: normality, independence and constant variance. The studentized residuals (r_i) are defined as follows:

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2(1 - h_{ii})}} \tag{26}$$

where $e_i = y_i - \hat{y}_i$ and h_{ii} are elements of the hat matrix diagonals $\mathbf{H} = \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'$.

Figures 5 to 8 present the diagnosis plots used to check the adequacy of the fitted model. In the normal probability plot for the studentized residuals shown in Figure 5, we may observe that there is no indication that the normality assumption should not be accepted, as there are no points way off the alignment.

In order to check the independence assumption, there is the studentized residuals plot for the observations in the order in which the experiments were performed (see Fig. 6).

As the residuals from the plot shown in Figure 6 are randomly distributed and without any obvious trend to correlate them, there is no reason to suspect that the independence assumption is not valid.
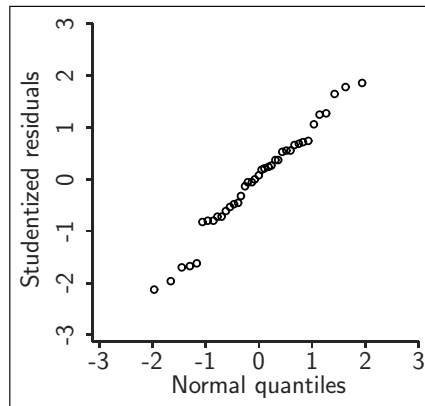
Figure 5 – Normal probability plot for the studentized residuals.
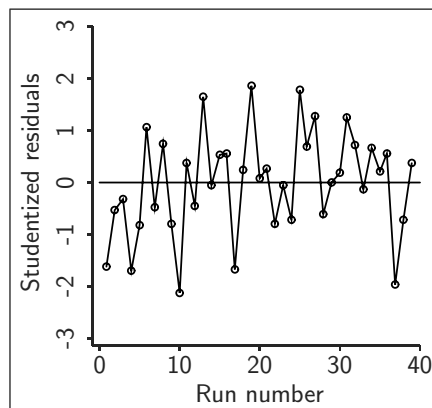
Figure 6 – Plot of studentized residuals versus run number.

In order to check the additivity of the model regarding the linear model, there is the plot of studentized residuals *versus* fitted values, shown in Figure 7.

As the residuals shown in the plot from Figure 7 are randomly distributed around zero, there is no reason to suspect that the additivity assumption should not be accepted.

In order to check the assumption of constant variance, the plot of absolute values of studentized residuals *versus* fitted values, which is shown in Figure 8 is used.

The plot shown in Figure 8 does not present any indication of growth of variance with the increase in the fitted value.

Therefore, the adequacy of Model M2 was checked. It is also worth pointing out that the same diagnosis plots (not shown here) were built for Models M1, which enabled the verification of adequacy of this model.
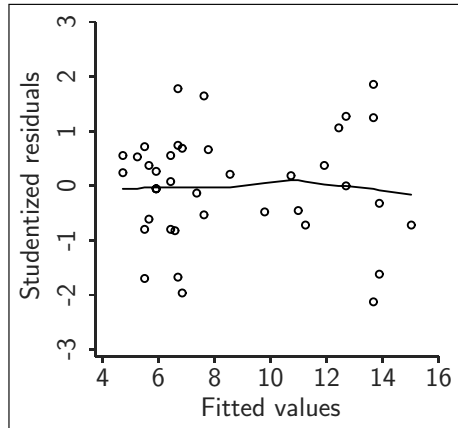
Figure 7 – Plot of studentized residuals *versus* adjusted value.
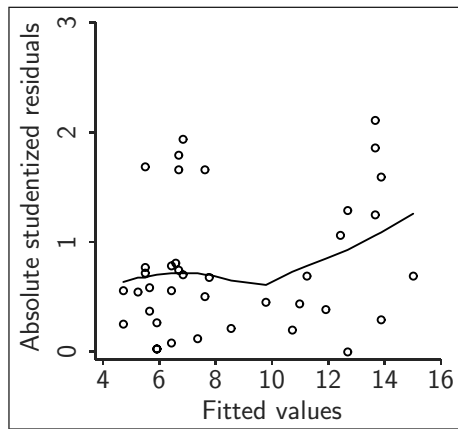

Figure 8 – Plot to check non-homogeneity of variance.

7   RESPONSE OPTIMIZATION

This section presents the response optimization taking Models M1 and M2 into account. We may observe that the Model M2 provides the lowest variance of a future response at the point of interest. As mentioned above, a response of 8 seconds is desirable in the delay mix experiment. Several formulations may result in prediction of new response equal to 8 seconds. Consequently, a desirable aim is to minimize the variance of a new response among the combinations of formulations and process variables that result in an expected response value of 8 seconds. Figures 9 to 12 present contour plots of response prediction and standard deviation of the mean response as a function of the L-pseudocomponents and the level of process variables ($z_1 = 1$ and $z_2 = 1$), considering the Models M1 and M2.

By analyzing the plots in Figures 9 to 12, we may observe that, when the 8-second contour curve is overlapped with the contour curves of standard deviation, a minimum standard deviation of
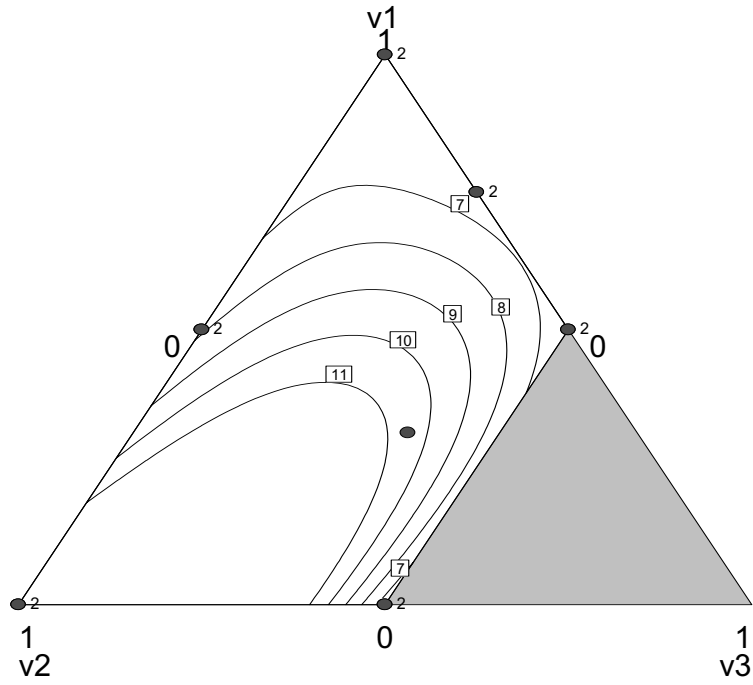
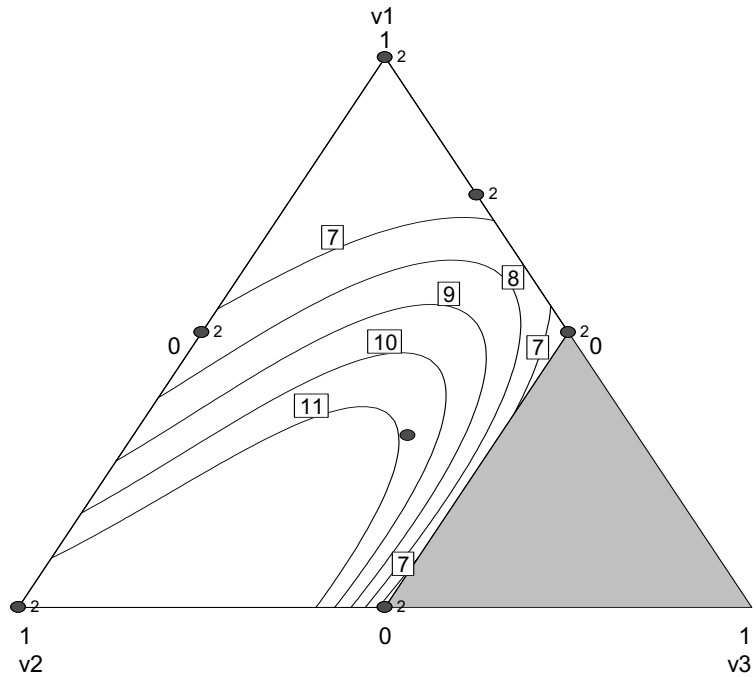**Figure 9** – Contour plots of response prediction for Model M1.

Figure 10 – Contour plots of response prediction for Model M2.
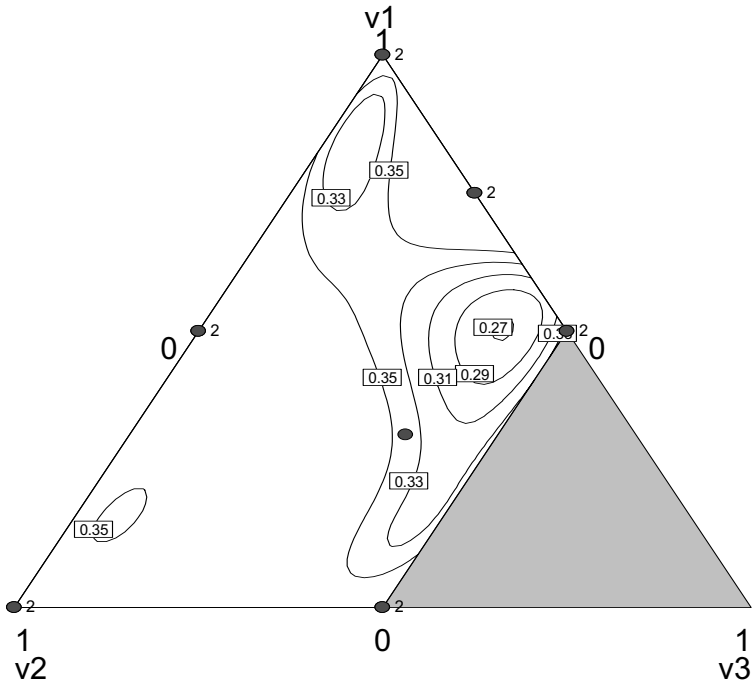
Figure 11 – Contour plots of standard deviation for Model M1.
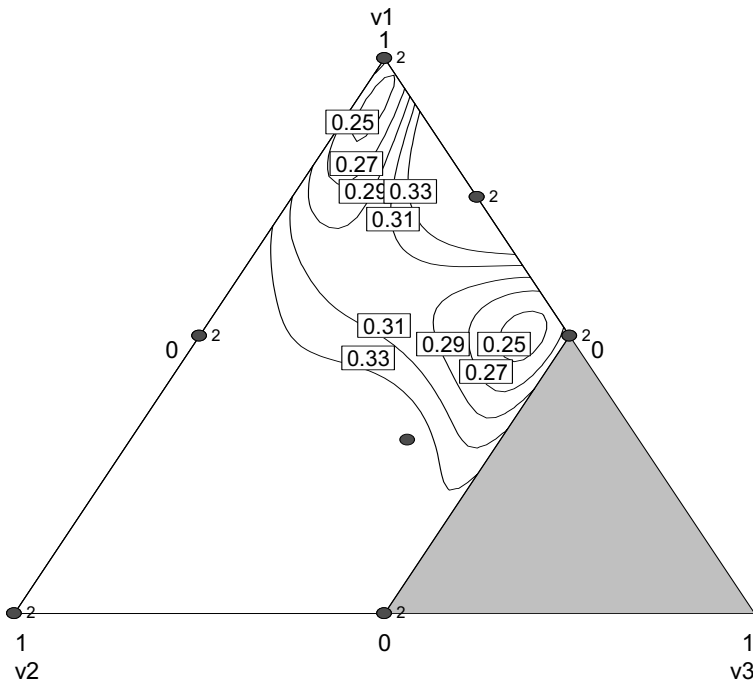

Figure 12 – Contour plots of standard deviation for Model M2.

the mean near 0.27 (Model M2) and 0.25 (Model M1) must be obtained. However, a stricter optimization procedure must be performed.

The estimation vector for the coefficients is $\hat{\boldsymbol{\beta}} = \left(\mathbf{W'W}\right)^{-1}\mathbf{W'y}$, the variance-covariance matrix is $\text{var}(\hat{\boldsymbol{\beta}}) = \sigma^2 \left(\mathbf{W'W}\right)^{-1}$. The prediction value for the response at point $\mathbf{w}$ ($\mathbf{w'}$ is a matrix line $\mathbf{W}$) is represented by $\hat{\mathbf{y}}(\mathbf{w})$ and its variance $\text{var}\left[\hat{y}(\mathbf{w})\right] = \sigma^2 \mathbf{w'}\left(\mathbf{W'W}\right)^{-1}\mathbf{w}$. Suppose any given point $\mathbf{w}_0$ in the space of the mixture components and the process variables, the model in the point may be represented as follows: $y(\mathbf{w}_0) = \mathbf{w'}_0\boldsymbol{\beta} + \boldsymbol{\varepsilon}$.

The estimation for a new response at this point is the same estimation of the mean:

$$E\left[\hat{y}(\mathbf{w}_0)\right] = \mathbf{w'}_0\hat{\boldsymbol{\beta}} \tag{27}$$

The variance for a new response at this point is then:

$$\text{var}\left[\hat{y}(\mathbf{w}_0)\right] = \text{var}\left(\mathbf{w'}_0\hat{\boldsymbol{\beta}} + \varepsilon\right) = \text{var}\left(\mathbf{w'}_0\hat{\boldsymbol{\beta}}\right) + \text{var}(\varepsilon) = \sigma^2\mathbf{w'}_0\left(\mathbf{WW}\right)^{-1}\mathbf{w_0} + \sigma^2 \tag{28}$$

or:

$$\text{var}\left[\hat{y}(\mathbf{w}_0)\right] = \sigma^2\left[\mathbf{w'}_0\left(\mathbf{W'W}\right)^{-1}\mathbf{w}_0 + 1\right] \tag{29}$$

The problem may then be formulated as follows:

$$\min \quad \text{var}\left[\hat{y}(\mathbf{w}_0)\right] = \sigma^2\left[\mathbf{w'}_0\left(\mathbf{W'W}\right)^{-1}\mathbf{w}_0 + 1\right]$$

$$\text{subject to:} \quad E\left[\hat{y}(\mathbf{w}_0)\right] = \mathbf{w'}_0\boldsymbol{\beta} = 8;$$

$$v_1 + v_2 + v_3 = 1;$$

$$0 \leq v_1 \leq 1;$$

$$0 \leq v_2 \leq 1;$$

$$0 \leq v_3 \leq 0,5;$$

$$z_1 = -1, \quad 1;$$

$$z_2 = -1, \quad 1.$$

Using a codified search routine in Matlab®, the following solutions were found:

Model M1:

$$v_1 = 0,5026; \quad v_2 = 0,0855; \quad v_3 = 0,4119; \quad z_1 = 1; \quad z_2 = 1;$$

$$\text{var}\left[\hat{y}(\mathbf{w}_0)\right] = 0,6626; \quad E\left[\hat{y}(\mathbf{w}_0)\right] = 8,0001.$$

Model M2:

$$v_1 = 0,4995; \quad v_2 = 0,0652; \quad v_3 = 0,4353; \quad z_1 = 1; \quad z_2 = 1;$$

$$\text{var}\left[\hat{y}(\mathbf{w}_0)\right] = 0,5757; \quad E\left[\hat{y}(\mathbf{w}_0)\right] = 8,0000.$$

Using Equation (4), the following final solution is achieved in terms of actual components:

Model M1:

$$x_1 = 0,7901; \quad x_2 = 0,1434; \quad x_3 = 0,0665; \quad z_1 = 1; \quad z_2 = 1.$$

Model M2:

$$x_1 = 0,7900; \quad x_2 = 0,1426; \quad x_3 = 0,0674; \quad z_1 = 1; \quad z_2 = 1.$$

By analyzing the solutions, we may conclude that the two models provide practically the same mix formulation at the optimal point. However, the Model M2 is chosen, since it provides better variance (0.5757) for the estimation of future responses.

8   CONCLUSIONS

A tutorial for mixture-process experiments was presented, as well as the mixture experiment with the inclusion of process variables. There is a wide range of combined models to represent the problems of the mixture-process type. In addition, the constraints in mixture experiments are very common, thus making the resulting experimental region a sub-region of the original region.

The data analysis in very restricted mixture regions may be very complex due to the very nature of the mix variables, with possible multicollinearity among some of the terms of the model concerned, which may result in very unstable and inflated coefficient estimators, resulting in arbitrary selection of important variables. With a view to reducing the effect of this type of problem, the use of pseudocomponents was chosen, instead of actual components.

At first, the backward model selection method was used and, then, a model selection method based on the Information Theory, which proved to be more efficient in the case under study.

Then a model was selected, with which was possible to determine the proportion of each of the three components of the chemical mix and the level of the two process variables, so that the prediction of the delay time has minimum variance and expected value of 8 seconds.

It is worth pointing out that the proposed modification of process in relation to the granulometry of the mix and the introduction of the orifice in the delay mechanism result in the reduction of the variance of a new response at the point of interest.

REFERENCES

[1]   AKAIKE H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle, in Mehra RK & Csaki F (editors), *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest.

[2]   BURNHAM KP & ANDERSON DR. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach*. Second edition, Springer, New York.

[3]   CHUNG PJ, GOLDFARB HB & MONTGOMERY DC. 2007. Optimal Designs for Mixture-Process Experiments with Control and Noise Variables. *Journal of Quality Technology*, **39**: 179–190.

[4]   CORNELL JA. 2000. Fitting a Slack-Variable Model to Mixture Data: Some Questions Raised. *Journal of Quality Technology*, **32**: 133–147.

[5] CORNELL JA. 2002. *Experiments with Mixtures: Designs, Models and the Analysis of Mixture Data*. 3rd edition, John Wiley & Sons, New York.

[6] CORNELL JA & GORMAN JW. 2003. Two New Mixture Models: Living With Collinearity but Removing its Influence. *Journal of Quality Technology*, **35**: 78–88.

[7] DAL BELLO LHA. 2010. Modelagem em Experimentos Mistura-Processo para Otimização de Processos Industriais. D.Sc. Thesis, Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

[8] DAL BELLO LHA & VIEIRA AFC. 2010. Optimization of a Product Performance Using Mixture Experiments. *Journal of Applied Statistics*, **37**: 105–117.

[9] DAL BELLO LHA & VIEIRA AFC. 2011 (to be published). Optimization of a Product Performance Using Mixture Experiments Including Process Variables. *Journal of Applied Statistics*.

[10] GOLDFARB HB, BORROR CM & MONTGOMERY DC. 2003. Mixture-Process Variable Experiments with Noise Variables. *Journal of Quality Technology*, **35**: 393–405.

[11] GOLDFARB HB, BORROR CM, MONTGOMERY DC & ANDERSON-COOK CM. 2004a. Three-Dimensional Variance Dispersion Graphics for Mixture-Process Experiments. *Journal of Quality Technology*, **36**: 109–124.

[12] GOLDFARB HB, BORROR CM, MONTGOMERY DC & ANDERSON-COOK CM. 2004b. Evaluating Mixture-Process with Control and Noise Variables. *Journal of Quality Technology*, **36**: 245–262.

[13] GOOS P & DONEV AN. 2006. The D-Optimal Design of Blocked Experiments with Mixture Components. *Journal of Quality Technology*, **38**: 319–332.

[14] GOOS P & DONEV AN. 2007. Tailor-Made Split-Plot Designs for Mixture and Process Variables. *Journal of Quality Technology*, **39**: 326–339.

[15] HARRELL FE JR. 2001. *Regression Modelling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer-Verlag, New York.

[16] HURVICH CM & TSAI C-L. 1989. Regression and Time Series Model Selection in Small Samples. *Biometrika*, **76**: 297–307.

[17] KHURI AI. 2005. Slack Variable Models Versus Scheffé's Mixture Models. *Journal of Applied Statistics*, **32**: 887–908.

[18] KOWALSKI SM, CORNELL JA & VINING GG. 2002. Split-Plot Designs and Estimation Methods for Mixture Experiments with Process Variables. *Technometrics*, **44**: 72–79.

[19] MYERS RH & MONTGOMERY DC. 2002. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 2nd edition. John Wiley & Sons, New York.

[20] PIEPEL GF. 2004. 50 Years of Mixture Experiment Research: 1955-2004, in *Response Surface Methodology and Related Topics*. A.I. Khuri, ed., World Scientific Publishing, Singapore, 283–327.

[21] PIEPEL GF. 2007. A Component Slope Linear Model for Mixture Experiments. *Quality Technology & Quantitative Management*, **4**(3): 331–343.

[22] PRESCOTT P. 2004. Modeling in Mixture Experiments Including Interactions with Process Variables. *Quality Technology & Quantitative Management*, **1**(1): 87–103.

[23] PRESCOTT P, DEAN AM, DRAPER NR & LEWIS SM. 2002. Mixture Experiments: Ill-Conditioning and Quadratic Model Specification. *Technometrics*, **44**: 260–268.

[24] SAHNI NS, PIEPEL GF & NÆS T. 2009. Product and Process Improvement Using Mixture-Process Variable Methods and Optimization Techniques. *Journal of Quality Technology*, **41**(2): 181–197.

[25] VIEIRA AFC & DAL BELLO LHA. 2006. Experimentos com Mistura para Otimização de Processos: Uma Aplicação com Respostas Não-Normais. *Pesquisa Operacional*, **26**(3): 605–623.