Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 9, Issue, 07, pp.53495-53502, July, 2017

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

FACTORIAL PLANNING AND RESPONSE SURFACE FOR THE DEFINITION OF EXPERIMENTAL PARAMETERS IN SCIENTIFIC RESEARCH, A REVIEW

Claudiany S. Leite Lima, Danielle C. Barros, Davi T. Reis, Denis S. Passos, Naiane G. Silva, Pedro G. B. Bonfim, Vanessa M. Mattje, Gessiel Newton Sheidt and *Augustus Caeser Franke Portella

Bioprocess Engineering and Biotechnology Division, Federal University of Tocantins (UFT), Gurupi, TO, Brazil

ARTICLE INFO

Article History:
Received 26th April, 2017
Received in revised form
30th May, 2017
Accepted 06th June, 2017
Published online 22nd July, 2017

Key words:

Trials, Experiments, Optimization.

ABSTRACT

In order to obtain the optimization of a given experiment, we need to carry out a triage, using either full or fractional factorial planning. These are statistical procedures that seek to minimize the work required. This eliminates the variables called factors, which are not significant in the experiment. Factorial planning basically consists of carrying out a survey of the factors of the proposed experiment and evaluating the effects they exert on each other and on the final result. The Response Surface Methodology (RSM) consists of a collection of mathematical and statistical techniques based on the fit of a polynomial equation to the experimental data, which should describe the behavior of a dataset in order to make statistical predictions. The objective of the present work is to synthesize the theoretical and practical knowledge of the methodology of Factorial Planning and RSM as statistical tools for evaluation and optimization of parameters involved in an experimental scientific research project.

Citation: Claudiany S. Leite Lima, Danielle C. Barros, Davi T. Reis *et al.* 2017. "Factorial planning and response surface for the definition of experimental parameters in scientific research, A review", *International Journal of Current Research*, 9, (07), 53495-53502.

INTRODUCTION

As a result of the needs of modern society, scientific research has made great strides in all spheres of science, promoting a growing range of data and information, however, for the proper exploitation and understanding of information, it is of great value to use Of statistical tools (Cunico *et al*., 2008). Usually, the optimization of experimental variables is performed through procedures that analyze the effect of one variable at a time, called univariate, having a disadvantage in time spent for optimization and the lack of evaluation about the interactions among the variables that affect the process being studied. These disadvantages result in an inefficient optimization, preventing the rapid establishment of true optimums, which are achieved by the application of multivariate procedures (Rodrigues and Iemma, 2005). Among the various options of statistical evaluation methodologies, the Response Surface Methodology (RSM) and Factorial Planning are highlighted. Considering the varieties of experimental planning, experimental systems delineated in factorial schemes are the ones that provide more information about the system with a smaller number of experiments, and may involve combinations between levels of two or more factors, remembering Which are the independent variables or predictors, which had their levels fixed a priori, according to the interest of the researcher (Bezerra *et al*., 2008). Another advantage of Factorial Planning is that all variables are studied simultaneously. In addition, many scientists do not have an experimental planning methodology that is both simple and useful, and most of the time, they have difficulties choosing a physico-mathematical model that effectively represents the phenomena to be analyzed (Calado, 2003). For Bingöl *et al*. (2012), the Response Surface Methodology is a compilation of mathematical and statistical methods suitable for optimizing, developing and improving processes. The Response Surface Methodology is a collection of mathematical and statistical techniques based on the fit of a polynomial equation to the experimental data, which should describe the behavior of a dataset in order to make statistical predictions. It can be well applied when a response or set of responses of interest are influenced by several variables. The objective is to simultaneously optimize the levels of these variables to obtain the best performance of the system (Bezerra *et al*., 2008). Before applying the RSM methodology, it is necessary to first choose an experimental design that defines the experiments that must be performed in the experimental region under study. There are some experimental matrices for this purpose. Experimental designs for first-order models (eg, factorial designs) can be used when the data set has no curvature (Hanrahan; Lu, 2006). However, in order to approximate a response function to experimental data that cannot be

Corresponding author:* **Augustus Caeser Franke Portella,
Bioprocess Engineering and Biotechnology Division, Federal University of Tocantins (UFT), Gurupi, TO, Brazil.

described by linear functions, experimental designs should be used for quadratic response surfaces, such as three-factorial, Box-Behnken, central compound and Doehlert designs (Bezerra *et al*., 2008). The response surface is a methodology widely used for analysis and fermentation processes, modeling and optimization of heterogeneous photo-hay processes (Desai *et al*., 2008) and little used as statistical methods of experiments using adsorption (Ranjan; Mishra; Hasan, 2011; Geyikçi *et al*., 2012). In view of the above, the objective of the present work is to synthesize the theoretical and practical knowledge of the methodology of Factorial Planning and RSM as statistical tools for the evaluation and optimization of parameters involved in an experimental scientific research project.

Factorial Planning

Experiment planning was initiated by Ronald A. Fisher between the 1920s and 1930s, where he performed data analysis and statistics at the Experimental Agricultural Station in London. In addition, Fisher and other authors contributed for the first time to the development and use of the ANOVA technique, which means Analysis of Variance, since it was considered as a primitive tool for statistical analysis of experimental design (Cochran, 1947). Its importance is in developing new processes and improving those in use. Thus, proper planning in addition to improving processes, reduces variability of results, as well as reduces analysis time and costs. Experiment analyzes and planning techniques are used to improve the properties of manufacturing processes, to reduce testing, to optimize the use of company resources, such as materials, availability of equipment, Employees, among others (Montgomery, 2005).

Factorial Planning Detail

Factorial Planning basically consists of carrying out a survey of the factors of the proposed experiment and evaluating the effects they exert on each other and on the final result. This is done so that the factors vary with at least two different levels (n), that is, for k factors and n> = 2 levels, we will have: Nk experiments, for two levels we have 2k. Exemplifying; We have an experiment that we wish to optimize containing 4 factors we will have the following number of experiments: 24 or 16 experiments. This result will be the number of tests performed and to evaluate the uncertainties involved, these will be done in duplicate, that is, twice each test, totaling 32 experiments, always performed in a random way to avoid systematic errors in the experimental procedures (Filho *et al*., 2000). Factorial planning has different properties used (Martinez; Calil, 2003):

- Direct the research;
- Specify the size of the sample to be selected;
- Allow multiple comparisons, and thus facilitate the development and critique f models;

Provide highly efficient parameter estimators (parameter estimators with small variance).When planning an experiment it is necessary to stipulate a sequence of experimental data collections to later achieve a possible objective. Within the experimental methodologies available in the literature, Factorial Planning is the most feasible when it is sought to study the effects of two or more influencing variables, where all possible combinations of levels of each variable are investigated in each attempt or replicate (Button, 2005). This type of planning is represented by 2k, k means the number of factors evaluated in 2 levels. In case of having 5 variables, 25 = 32, it means that 32 experiments should automatically be performed. This type of planning is particularly useful in the early stages of an experimental work when many variables have to be investigated. When it comes to planning with k>4, the effects of high orders are almost always not significant, with which it is possible to obtain information of the most important effects with a smaller number of experiments, and to obtain, in most cases, the same conclusions that would be obtained with a complete factorial. The models that have these characteristics are known as fractional factorial planning, having as an example, 25-1, which would only result in 16 experiments (Rodrigues and Iemma, 2005). In this type of Factorial Planning, in which levels are usually coded with the (+) and (-) signals, the assignment to the upper or lower levels occurs arbitrarily and does not interfere with the realization of the experiments or interpretation of the results. These signals also allow the schematization of these variables in the form of planning matrices, as well as determining, through calculations, their influence and their interactions in the system (Calado, 2003). In the specialized databases, the factorial schemes are not considered experimental designs, but the design of treatments. Faced with this, each combination is a treatment. In this context we can have factorial treatment schemes in completely randomized experimental designs, in Latin squares, in randomized blocks among others (Cunico *et al*., 2008). The factorial scheme is considered complete when all possible combinations, among all levels of each factor, are present. However, in other cases we have an incomplete factorial scheme (Rodrigues and Iemma, 2005).Performing complete factorial experiments means estimating the effects of the factors and the effects of the interactions between them with equal precision. According to Montgomery, Runger and Calado (2000), interaction is the failure of one factor to produce the same effect on response, under different levels of the other factor. These authors also state that there is interaction when the difference in response between levels of one factor is not the same at all levels of the other factors. It is recommended to use full factorials when investigating a small number of factors, say up to five, since the larger the number of factors, the greater the number of treatments.

In a completely randomized design with two factors, each with two levels, example factor A = Temperature, with levels '1' and '2' and B = pH, with levels '1' and '2', we have the following combinations, Tests or treatments:

$$\begin{matrix} A1B1 & A1B2 \\ A2B1 & A2B2 \end{matrix}$$

Advantages and Disadvantages of Using Factorial Planning of Experiments

Factorial assays allow time and resource savings but, in particular, allow broader conclusions about the factors including the study of the interaction between them and greater precision for the estimates of the main effects of the factors. The disadvantage is related to the rapid increase in the number of treatment as the number of factors or the number of factor levels increases. In the case of the randomized block design, increasing block size entails a loss of efficiency if there is an increase in heterogeneity within the block (Lima; Abreu, 2001).

Among the several advantages of using Factorial Planning, the following stand out (Button, 2005):

- Reduction of the number of tests without prejudice to the quality of information;
- Simultaneous study of several variables, separating their effects;
- determining the reliability of results;
- Conducting the research in stages, in an interactive process of adding new essays;
- Selection of variables that influence a process with a reduced number of tests;
- Representation of the process studied through mathematical expressions;
- Elaboration of conclusions based on qualitative results.

Response Surface Models

When working with scientific research for product development or improvement in the laboratory, we are faced with diverse numbers, that is, experiments with several parameters that need to be analyzed and organized for verification of significance. The scientific methodology adapted to the industry by Shewhart (1931) in its PDCA cycle better defines the way in which planning and experimental analysis is performed (Schissatti *et al*., 1998). In this famous cycle the variability of information is analyzed in four aspects to inquire, to plan, to do and to act. Planning and experimenting, ascertaining the results and conducting the new knowledge to the production line should be regular steps applied in the company, demonstrating living and constant activity in continuous movement (Duncan, 1986). When we speak of response surface, we are referring to very useful mathematical tools when we are interested in optimizing a process in which we have the influence of several factors on a response variable. Therefore, response surface models can be explored to determine optimal working conditions or the sensitivity of the response variable to changes in the levels of the factors of interest (Action, 2016).

To work with a response surface model we follow three guidelines:

1. Sampling: where you have defined the number of tests that will be executed, already thinking about the model that will be programmed.
2. Modeling and Hypothesis Testing: here we fit the models and analyze.
3. Optimization: in this phase the optimum configuration of the levels of the factors of interest are obtained, in the intervals considered, and where the need to perform the experiment again considering newlevels for the factors is verified.

For surface response tasks the experimental design models are of the second order type, among them we have a central composite design, Box-Behnken designs and optimum designs. Once the sampling phase is followed, the following techniques can be used: regression models, parameter estimation (MQ), statistical inference techniques, ANOVA and lack of fit. Finally, the optimization of the now adjusted model follows some of the following techniques, calculating maximum or minimum or minimum of functions, contours, graphical visualization, numerical methods, steepest ascent or descent and Monte Carlo simulation (Action, 2016).

Model of Second Order of Response Surface

To sum up, when it comes to response surface the primitive idea of the techniques is to consider that there is always a relationship between the variables x_1, x_2, ..., x_n and y, a relation that is unknown but we can approximate the function by a polynomial relation, so that:

I. $y = f(x_1, x_2, \dots, x_n, \beta)$

II. $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \beta_{11} x_1^2 + \cdots + \beta_{nn} x_n^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \cdots + \beta_{(n-1)n} x_{n-1} x_n + \varepsilon$

To simplify, we will define the equation below as the response surface of the given curve:

III. $E(y) = f(x_1, \dots, x_n, \beta)$

The second-order response surface model is represented by equation 2, where s means the random error involved in the model and β are coefficients of the regression model. This model is suitable for experimental data systems with abundant curvature, which does not mean that this model fits well in all cases that contain curvature. In some cases it becomes necessary to use a more complex model, in rarer cases it may be necessary to use cubic terms until, to fit correctly.

When we design a response surface experiment, we must be able to provide estimates of the "p" parameters in the model, so that:

$$p = \frac{(k+1)(k+2)}{2}$$

For example, for k = 2 we have p = 6, so:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2$$

It is desired that a response surface experiment has characteristics such as, good fit to the data, provide enough information to perform a lack of fit test and to estimate the pure error, allow for easy data amplification, be robust To the presence of outliers in the data and in the control of factor levels, has a reasonable cost, allows the experiments to be performed in blocks and has constant prediction variance.

Table 1. Gain in a set of pistons at different emission and collection levels. Where T is the emission time DI is the dose of ions and G is the gain

Observation	T	DI	G
1	195	4.0	1004
2	255	4.0	1636
3	195	4.6	852
4	255	4.6	1506
5	225	4.2	1272
6	225	4.1	1270
7	225	4.6	1269
8	195	4.3	903
9	255	4.3	1555
10	225	4.0	1260
11	225	4.7	1146
12	225	4.0	1276
13	225	4.7	1225
14	230	4.3	1321

Example of Second Order Model Application

The gain of a transistor consists of the difference between the emitter and the collector. The variable Gain (in hFE) can be controlled in the ion deposition process by means of the variables Time of emission (in minutes) and Dose of ions (x_{10}^{14}). The data are shown in Table 1. Our objective is to evaluate the linear relationship between Gain of the transistors and the covariates Emission time and Dose of ions.

The linear regression presented the following results:

Table 2. ANOVA analysis

Factors	Degrees of freedom	Sum of Squares	Mean of Square	F ratio	p-value
T	1	630967.864	630967.864	517.143	1.342E-10
DI	1	20998.234	20998.234	17.210	0.00162
Error	11	13421.115	1220.101		

Analyzing Table 2 we observed that the variables T (time) and DI (Dose of ions) are significant for the model since the p-value for the two coefficients were smaller than $\alpha = 5\%$. Table 3 presents a descriptive analysis of the residuals of the model, showing quartiles, maximum and minimum, mean and median values.

Table 3. Exploratory analysis (residues)

Minimum	1Q	Median	Average	3Q	Maximum
-44.58	-26.35	-3.266	1.27E-16	26	63.2

In Table 4 we have the estimates of the coefficients related to the input and intercept variables. We observed that the coefficient of the time variable is positive, we conclude that in the interval of the analysis, increasing the time also increases the gain of the transistor. In contrast, the dose variable has a negative coefficient, that is, an increase in it causes a decrease in the gain of the transistor. The table also presents the p-values for each coefficient, in which the null hypothesis where the coefficient is not significant. The null hypothesis, because the calculated values were smaller than $\alpha = 5\%$, which shows that the intercept and the input variables are important for the model.

Table 4. Coefficients

Predict	Estimate	Standard desviation	T-Stat.	P-Value
Intercept	-520.0766	192.107	-2.707	0.0203
T	10.781	0.474	22.729	1.349E-10
DI	-152.148	36.675	-4.148	0.00162

The quality of the adjustment is shown in Table 5, where we can observe the adjusted R^2 value. In this case, we have that about 97% of the variability of the data is explained by the adjusted regression model. The Table 6 shows the confidence intervals for each parameter.

Table 5. Descriptive measurement of fit quality

Standard Desviation	Degreesoffreedon	R²	R Adjusted
34.929	11	0.979	0.9761

Table 6. Confidenceinterval

	2.50%	97.50%
(Intercept)	-942.901	-97.251
T	9.737	11.825
DI	-232.870	-71.426

The Table 7 shows the predicted or adjusted values and the respective confidence intervals and standard deviations for each observation. In Table 8 we have the expected value, the confidence interval and the standard deviation for the level 200 of the explanatory variable time.

Table 7. Prediction interval

T	DI	G	Adjusted Values	Lower Limit	Upper Limit	Standard Deviation
195	4.0	1004	973.653	927.051	1020.255	21.173
255	4.0	1636	1620.523	1574.522	1666.523	20.899
195	4.6	852	882.364	839.083	925.645	19.664
255	4.6	1506	1529.233	1486.413	1572.053	19.454
225	4.2	1272	1266.658	1243.315	1290.001	10.605
225	4.1	1270	1281.873	1253.787	1309.959	12.760
225	4.6	1269	1205.799	1176.260	1235.337	13.420
195	4.3	903	928.008	890.113	965.904	17.217
255	4.3	1555	1574.878	1537.617	1612.139	16.929
225	4.0	1260	1297.088	1262.985	1331.191	15.494
225	4.7	1146	1190.584	1154.804	1226.364	16.256
225	4.3	1276	1251.443	1230.675	1272.211	9.435
225	4.7	1225	1187.541	1150.428	1224.654	16.862
230	4.3	1321	1305.349	1284.028	1326.670	9.687

Table 8. Forecastinterval

T	DI	Adjusted Values	Lower Limit	Upper Limit	Standard Deviation
200	4.3	981.914	897.992	1065.837	34.9299

In Tables 9 and 10 we have the residue analysis. The Diagnostic Analysis, we find the analysis of some types of residues beyond the points of influence, DFf its, DFbetas and Cook distance. Table 10 shows the calculations used to obtain

Table 9. Diagnostic analysis, summary table

Obs	T	DI	Waste	Studentizad Waste	Standard Waste	Leverage	DFFITS	DFBETA	D-COOK
1	195	4.0	30.346	1.103	1.092	0.367	0.840	-0.573	0.480
2	255	4.0	15.476	0.534	0.552	0.358	0.399	0.268	0.238
3	195	4.6	-30.364	-1.057	-1.051	0.316	-0.720	0.526	0.413
4	255	4.6	-23.233	-0.786	-0.800	0.310	-0.527	-0.382	0.310
5	225	4.2	5.3414	0.153	0.160	0.092	0.0488	-0.0008	0.029
6	225	4.1	-11.873	-0.350	-0.365	0.133	-0.137	0.0020	0.082
7	225	4.6	63.200	2.316	1.959	0.147	0.963	-0.0103	0.470
8	195	4.3	-25.009	-0.809	-0.822	0.242	-0.458	0.3838	0.269
9	255	4.3	-19.878	-0.632	-0.650	0.234	-0.350	-0.2910	0.208
10	225	4.0	-37.088	-1.209	-1.184	0.196	-0.598	0.0078	0.338
11	225	4.7	-44.584	-1.526	-1.442	0.215	-0.802	0.0068	0.437
12	225	4.3	24.556	0.713	0.730	0.072	0.200	-0.0036	0.118
13	225	4.7	37.458	1.256	1.224	0.2331	0.6928	-0.0054	0.389
14	230	4.3	15.650	0.449	0.466	0.0769	0.129	0.0294	0.077

the results of Table 9. Table 11 is the results obtained from the application of the Bonferroni test for the verification of atypical points.

Table 10. Criterion

Diagnosis	Formula	Value
hii (Leverage)	(2*(p+1))/n	0.43
DFFITS	2* raíz (p/n)	0.93
DCOOK	1	1.00
DFBETA	2/raíz(n)	0.53
Standard Waste	(-2,2)	2.00
StudentizadWaste	(-2,2)	2.00

Table 11. Outliers

Observation	T-value	p-Value	Bonferroni index
7	2.3161	0.04305	0.602713

In Table 12, we have the Anderson-Darling, Shapiro-Wilk, Kolmogorov Smirnov, Ryan-Joiner tests to verify if the residues have a normal distribution, as observed p-values are all higher than 0.05, so we do not reject the Hypothesis of data normality.

Table 12. Normalitytest

Test	Statistics	p-Value
Anderson-Darling	0.307	0.518
Shapiro-Wilk	0.951	0.584
Kolmogorov-Smirnov	0.160	0.423
Ryan-Joiner	0.979	0.619

In Tables 13 and 14 we have the homoscedasticity tests with a confidence level of 0.05, all the tests have p-values higher than the level of significance, so we do not reject the hypothesis of homoscedasticity of the residues. We note that by adopting a significance level of 5% we have that by the Durbin-Watson test in Table 16 that the residues are independent.

Table 14. Homoscedasticity test of Goldfeld Quandt

Variable	Statistic	DF1	DF2	P-Value
T	10.781	0.474	22.729	1.34912E-10
DI	-152.148	36.675	-4.148	0.001620499

The Figure 1 shows the dispersion of outliers at x and the results of the graphs shown in Figure 2, we know that no DFFIT S and D - COOK is in module greater than 1, so we have no observation of the example is an influential point.

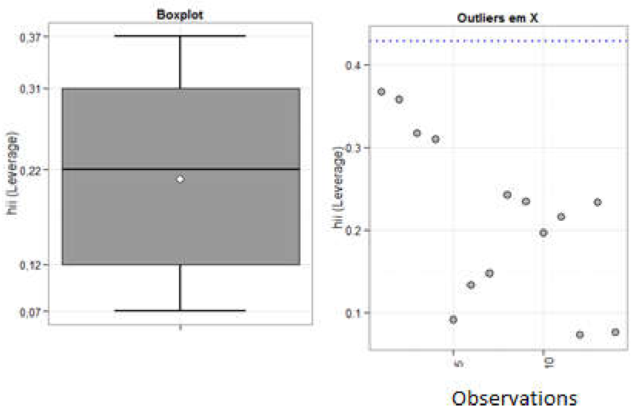

Figure 1. Outliers(x)

Table 13. Homoscedasticity test of Breusch Pagan

Statistic	DF	p-Value
0.432972	1	0.510

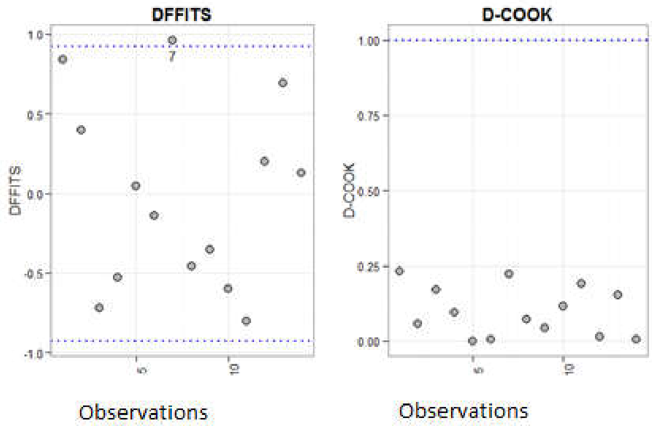

Figure 2. Point of influence

The Figures 4, 5, 6 and 7 show graphs with the response surface analysis. Where DI is the dose of ions, T is the emission time and G is the gain.

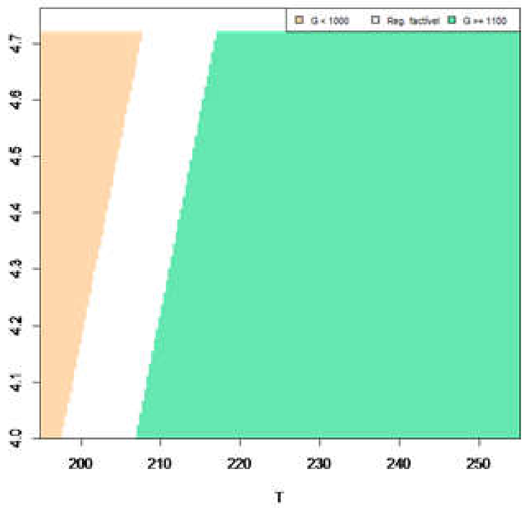

Figure 4. Graph of the feasible region, in this case shown in white above

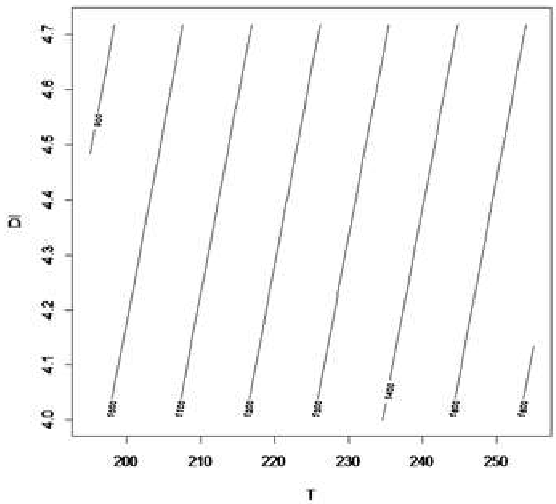

Figure 5. Line outline

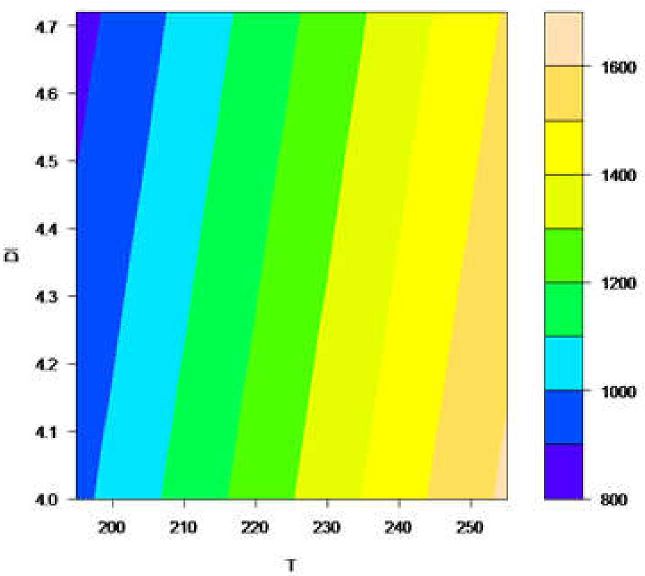

Figure 6. Contour of areas

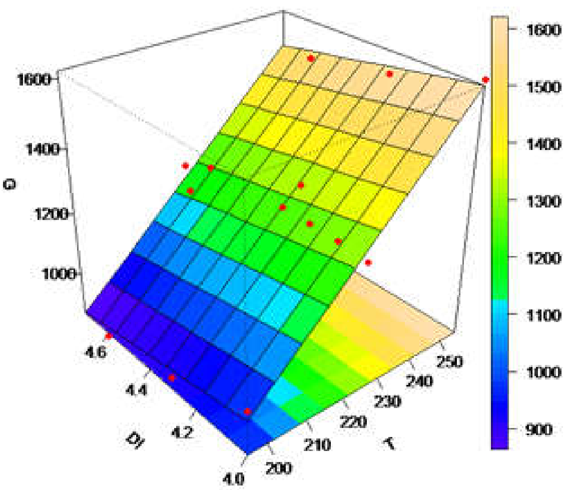

Figure 7. 3D representation of the response surface

Response Surface Methods

The response surface method is based on the simultaneous variation of several factors that are independent variables previously selected for their influence on process properties (dependent variables or responses). By using mathematical and statistical techniques, experimental results indicate a combination of factor levels within an optimal region (Grizotto *et al*., 2005). It is a technique that aims at the optimization of a variable response that is influenced by several factors, based on factorial planning. It has two distinct stages that are characterized as modeling and displacement, and are repeated as many times as necessary, aiming to reach an optimal region of the surface investigated (Comparini *et al*., 2009). According to Santiago (2016), this experimental strategy allows reducing the number of trials, ensuring statistical support for discussion of the results. This statistical method is recommended for experiments with high costs and difficult to stabilize treatments to reduce experimental error.

Method of Maximum Ascending Slope (Descending)

Initially the problem is interpreted so that the objective is clearly defined, as well as the factors that influence it. First, only a first order model is necessary in most cases, the relationship between objective and factors identified as important is established by a multiple regression (Mondim, 2014). Typically, the starting point of the experiment is far from the optimal process operating condition which is desirable. In this case, the goal will be to approach the optimum as quickly and economically as possible. When the hypothesis of applying the model and the linearity of the process are verified, one must walk towards the response surface that supposedly has the optimum (maximum or minimum). In order to do this, one must look for the maximum improvement in the response: by the maximum ascending slope method if we look for a maximum point, or by the maximum slope method if we look for a minimum (Action, 2016, Mondim, 2014). The maximum ascending (descending) method is used when the goal is to maximize (minimize) the response. In this method, one walks in sequence along the direction of maximum ascending slope, that is, in the direction and direction in which the maximum increase of the response variable occurs. If the response minimization goal moves sequentially along the direction of maximum downward slope, that is, in the direction and direction in which the maximum response variable occurs (Action, 2016).

Optimization Method

Optimizing a response means finding the best set of solutions for the independent variables x1, x2, ...,xk satisfying a given condition of y. The first derivative of this function is null if a function has a maximum or a minimum point. When the function is composed of a set of non-dependent variables, the partial derivatives of the function with respect to each of the non-dependent variables must also be null, so that there is a maximum or minimum point (Mondim, 2014). Most of the time, the form of the relationship between the response variable and the process factors is not known. Therefore, the first step of the response surface method consists of determining an equation that roughly represents the relationship between the response y and the process factors x1, x2, ...,xk (Action, 2016). When considering a small region of the response surface, away from the optimum point, there is almost no curvature on the surface. In this situation, the first order model is used. In the vicinity of the optimum, even for small regions of the surface, the curvature is generally more pronounced, and it is necessary to use a higher-order polynomial to represent the relationship between the response and the process factors x1, x2, ...,xk . Usually, a second-order model is employed (Action, 2016).

Example of Practical Application of Factorial Planning

In order to provide a better understanding of the application of the Factorial Planning methodology, an example of factorial 22 will be addressed, where the effect of pH and temperature on the activity of the inulinases enzyme was evaluated. Twelve trials were performed (4 factorials + 4 axial points + 4 repetitions at the central point).

Results and Discussion of the Example

The scheme used was a complete factorial 22, including the 4 axial points and 4 central points for evaluation of the pure error. Table 17 presents the planning with the 12 tests carried out, the coded values, the originals and the results of the enzymatic activities obtained.

Table 17. Factorial Planning, coded and original values of the study variables (pH and temperature) and enzymatic activity

Assays	pH	Temperature	pH	Temperature ° C	Activity (U / mL)
1	-1.0	-1.0	3.6	36	272
2	1.0	-1.0	6.4	36	83
3	-1.0	1.0	3.6	64	457
4	1.0	1.0	6.4	64	16
5	-1.4	0.0	3.0	50	360
6	1.4	0.0	7.0	50	83
7	0.0	-1.4	5.0	30	132
8	0.0	1.4	5.0	70	328
9	0.0	0.0	5.0	50	396
10	0.0	0.0	5.0	50	412

Note that depending on pH and temperature conditions, enzymatic activities may range from 16 to 457 U / mL. With these results it is possible to elaborate the 2nd order coded model that relates the enzymatic activity as a function of pH and temperature through software.

Table 18. Variance Analysis Regression

Variable	Coefficient	Standard error	T calculated	p-value
Average	393.00	22.19	17.71	0.00000
X_1	-127.72	15.69	-8.14	0.0002
X12	-90.44	17.55	-4.91	0.0021
X_2	49.40	15.69	3.15	0.0199
X_2^2	-86.19	17.55	-4.91	0.0027
$X_1.x_2$	63.00	22.19	2.84	0.0296
$Y1=393-127.72x_1-90.44x_1^2+49.40x_2-86.19x_2^2-63x_1x_2$				

Table 18 shows the analysis of variance (ANOVA) of the results obtained, with the coefficient of determination equal to 0.96 and the value of FCALC 5.9 times greater than the value of FT AB. It is then possible to construct the response surface and contour curves, which are respectively shown in Figure 9.

Table 19. Analysis of variance for the activity of inulinases

Factors	Degrees of freedom	Sum of Squares	Mean of Square	F ratio	p-value
Regression	5	249153.5	49830.7	25.3	0.00058
Residues	6	11820.8	1970.1		
Lack of Adjustment	3	10966.8	3655.6	12.8	0.03224
Error	3	854	284.7		
Total	11	260974.3			

$*R^2=95.52\%$ $F_{5,6;0.05}=4.4$

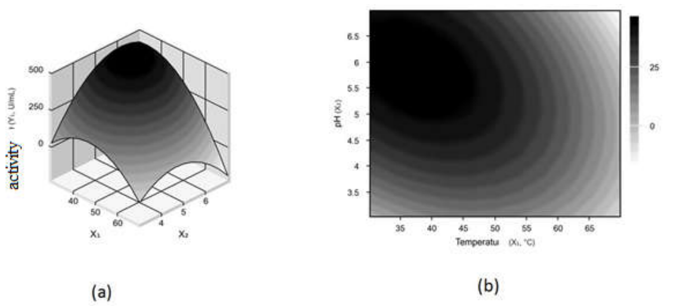

Figure 9. (a) Response surface (X_1 = Temperature and X_2 = pH) and (b) contour curves for inulinases activity

By analyzing the response surface and contour curves, we can verify the existence of an optimal region for enzymatic activity where a combination range of pH (3 to 5) and temperature (50 to 70 °C) is found. This methodology provides an adequate information for the number of tests performed. Evidencing a temperature and pH condition will be fixed for the reaction, however this optimal range range of the variables is much more interesting than just a point value since it provides information on the "robustness" of the process. That is, what is the temperature variation (+/- 10 ºC) that can be admitted around 60 ºC (optimal value) and pH 4 (+/- 1) that keeps the process in optimized condition.

Conclusion

The use of tools that allow the development of more accurate and viable data and thus to determine which variables influence the system can be useful in experiments. In general, we are only asked to do the propagation of uncertainties and to evaluate how reliable they obtained result is or to only make a mathematical analysis of the equations involved, obtaining importance of each greatness. In order to obtain the optimization of a given experiment, we need to carry out a triage, using either full or fractional factorial planning. These are statistical procedures that seek to minimize the work required. This eliminates the variables called factors, which are not significant in the experiment. The Factorial Planning determines what factors have significant effects on the response and also as the effect of one factor varies with the levels of other factors. In addition, it allows establishing and quantifying the correlations between the different factors. In view of the above, it is verified that without the use of factorial planning of experiments, important interactions between factors cannot be detected and the maximum optimization of the system may take longer to reach. This was evidenced in this work, confirming that multivariate systems, based on Factorial Planning of Experiments, allow the recognition of true optimal conditions from a small number of experiments.

REFERENCES

Action, N. P. *Método de Superfície de Resposta*. 2016. Método de Superfície de Resposta. Disponível em: <http://www.portalaction.com.br/planejamento-de-experimento/62-metodo-de-superficie-de-resposta>. Acessoem:24/05/2017. Citado 4vezesnaspáginas5,6,15e16.

Bezerra, M. A. *et al*. Response surfacemethodology (rsm) as a tool for optimization in analyticalchemistry. *Talanta*, Elsevier, v. 76, n. 5, p. 965–977, 2008. Citado 2 vezes nas páginas, 2e3.

Bingol, D. *et al*. 2012. Comparison of the results of response surface methodology and artificial neural network for the biosorption of lead usingblackcumin. *Bioresourcetechnology*, Elsevier, v. 112, p. 111–115, Citado na página2.

Button, S. T. 2005. Metodologia para planejamento experimental e análise de resultados. *São Paulo*, Citado 2 vezes nas páginas 4e5.

Calado, V. 2003. *Planejamento de experimentos usando o Statistica*. [S.l.]: Editora E-papers, Citado 2 vezes nas páginas 2e4.

Cochran, W. G. 1947. Some consequences when the assumptions for theanalysisofvariance are notsatisfied. *Biometrics*, *JSTOR*, v. 3, n. 1, p. 22–38, Citado na página3.

Comparini, A. *et al*. 2009. Metodologia de superfície de resposta: Uma introdução no software statistica. *São Carlos*, v. 1001, p. 14, Citado na página15.

Cunico, M. *et al*. 2008. Planejamento fatorial: Uma ferramenta estatística valiosa para a definição de parâmetros experimentais empregados na pesquisa científica. *Visão Acadêmica*, v. 9, n. 1, Citado 2 vezes nas páginas2e4.

Desai, K. M. *et al*. 2008. Comparisonof artificial neural network (ann) and response surface methodology (rsm) in fermentation media optimization: case study of fermentative production of scleroglucan. *Biochemical Engineering Journal*, *Elsevier*, v. 41, n. 3, p. 266–273, Citado na página 3.

Duncan, A. J. 1986. *Quality control and industrial statistics*. 5th ed. ed. Homewood, Ill: Irwin, ISBN 978-0-256-03535-3. Citado na página5.

Filho, J. d. P. A. *et al*. 2000. Atividades experimentais:: do método à prática construtivista. Florianópolis, SC, Citado na página3.

Geyikçi, F. *et al*. 2012. Modellingof lead adsorptionfrom industrial sludgeleachateonredmud by using rsmandann. *Chemical Engineering Journal*, Elsevier, v. 183, p. 53–59, Citado napágina3.

Grizotto, R. K. *et al*. 2005. Otimização via metodologia de superfície de respostas dos parâmetros tecnológicos para produção de fruta estruturada e desidratada a partir de polpa concentrada de mamão. *Ciência e Tecnologia de Alimentos*, SciELO Brasil, v. 25, n. 1, p. 158–164, Citado napágina15.

Hanrahan, G.; LU, K. Application of factorial and response surface methodology in modern experimental design and optimization. *CriticalReviews in AnalyticalChemistry*, Taylor & Francis, v. 36, n. 3-4, p. 141–151, 2006. Citado na página2.

Lima, P. and Abreu, A. d. 2001. *Estatística experimental: ensaios balanceados*. [S.l.]: Lavras: Universidade Federal de Lavras, Citado na página5.

Martinez, M. and Calil, C. 2003. Statistical design and orthogonal polynomial model to estimate the tensile fatigue strength of wood enfinger joints. *International Journal of fatigue*, Elsevier, v. 25, n. 3, p. 237–243, Citado na página3.

Mondim, A. F. d. S. 2014. *Metodologias de superfície de resposta: uma investigação no âmbito da indústria cerâmica e do vido*. Dissertação (Mestrado) — Universidade Aberta, Citado na página15.

Montgomery, D. 2005. *Diseño y análisis de experimentos*. LimusaWiley, ISBN 9789681861568. Disponível em:< https://books.google.com.br/books?id=TJFoAAAACAAJ>. Citado na página3.

Montgomery, D. C.; Runger, G. C.; Calado, V. 2000. *Estatística aplicada e probabilidade para engenheiros*. [S.l.]: Grupo Gen-LTC, Citado na página4.

Ranjan, D.; Mishra, D.; Hasan, S. 2011. Bioadsorption of arsenic: an artificial neural networks and response surface methodological approach. *Industrial & Engineering Chemistry Research*, ACS Publications, v. 50, n. 17, p. 9852–9863, Citado na página 3.

Rodrigues, M. I.; Iemma, A. F. *Planejamento de experimentos e otimização de processos: uma estratégia sequencial de planejamentos* . [S.l.]: Casa do Pão Editora, 2005. Citado 2 vezes nas páginas2e4.

Santiago, H. *Simulação de perdas por evaporação na pulverização aérea*. Tese (Doutorado) — Universidade Federal de Viçosa, 2016. Disponível em: <http://www.locus.ufv.br/bitstream/handle/123456789/10390/texto%20completo.pdf>. Acesso em: 20/05/2017. Citado na página15.

Schissatti, M. L. *et al*. 1998. Uma metodologia de implantação de cartas de shewhart para o controle de processos. Florianópolis, SC, Citado na página5.

Shewhart, W. A. *Economic control of quality of manufactured product*. [S.l.]: ASQ Quality Press, 1931. Citado na página5.

****