

ON THE SEARCH FOR CONDITIONS FOR PRODUCTION PROCESS OPTIMIZATION

Małgorzata Szerszunowicz

Abstract. The experimental design methodology mainly deals with the estimation of a usually unknown response surface function. An important problem in the field of design of experiments is to determine the number of experimental trials, taking into account certain limitations resulting from the nature of the manufacturing process. The experimental design methodology also includes – as if a separate section – methods of searching for conditions for the optimization of processes and the optimal design of experiments.

The aim of this paper is to present a method of searching for conditions for the optimization of manufacturing process results. The proposed method uses non-classical statistical methods and is presented for selected empirical data.

Keywords: design of experiments, response surface function, optimal design.

JEL Classification: C99.

DOI: 10.15611/me.2014.10.07.

1. Introduction

The elements of the theory of the design of experiments were first introduced at the beginning of the 20th century by R.A. Fisher in agricultural experiments. The results obtained at that time made it possible to use the experimental design methodology in many areas of science, inter alia biology, medicine, chemistry, engineering and statistical quality control [Ryan 2007].

The design of experiments, which has applications in production companies, allows one to improve the technological and economic results of the production process. The theory of the optimization of design of experiments in particular, plays an important role in practice and is extensively described in specialist literature [Puckelsheim 1993; Silvey 1980; Ryan 2007].

Małgorzata Szerszunowicz

Department of Statistics, University of Economics in Katowice, ul. 1 Maja 50, 40-287 Katowice, Poland.

E-mail: malgorzata.szerszunowicz@ue.katowice.pl

The purpose of this paper is to present a method of searching for the optimal conditions for manufacturing processes and its application. The aim of this research is to limit the number of experimental trials of design of experiment, which contributes to decreasing the cost of the experiment.

2. The theoretical basis for design of experiments

The application of the methods of the design of experiments in statistical quality control leads to determining the factors which most significantly affect the variable characterizing of the investigated process, as well as allows to specify the values of the factors for which the result variable reaches the desired value or the smallest variability.

The use of the design of experiments while preparing the production process is connected with the detailed formulation of the problem and the proper selection of the factors and response variable characterizing the process under study. Moreover, one ought to determine the number of experimental trials and take into account the possible randomization restrictions. Then, after performing the experiment, one should analyze the results and formulate conclusions and recommendations for the process under study [Montgomery 1997].

The experiment is a sequence of n successive experimental trials which are a single result of the value of response variable Y , with the fixed values of X_1, X_2, \dots, X_m specified on sets X_1, X_2, \dots, X_m respectively. The experimental area is a set of points $x = (x_1, x_2, \dots, x_m)$, where $x_i \in X_i, i = 1, 2, \dots, m$. Then the set of pairs $P_n = \{x_j, p_j\}_{j=1}^n$ is a design of the experiment with n experimental trials, where $x_j = (x_{1j}, x_{2j}, \dots, x_{mj})$ and $p_j = \frac{n_j}{n}$, where n_j is the number of experimental trials in point x_j of the experimental area, and also $\sum_{j=1}^n n_j = n$, $\sum_{j=1}^n p_j = 1$ for $j = 1, 2, \dots, n$. The relationship between the set of factors and the response variable characterizing the process in the best way is presented in the form of the following statistical model [Wawrzynek 1993]:

$$Y(X_1, X_2, \dots, X_m) = y(X_1, X_2, \dots, X_m) + \varepsilon, \quad (1)$$

where $EY(X_1, X_2, \dots, X_m) = y(X_1, X_2, \dots, X_m)$, $E\varepsilon = 0$ and $V\varepsilon = \sigma^2$, where σ^2 is a constant value, independent of the values of the factors. This model

can be presented as a general linear model [Wawrzynek 2009] $Y = F\beta + \varepsilon$, where

$$Y^T = (Y_1 Y_2 \dots Y_n) \quad (2)$$

$$\varepsilon^T = (\varepsilon_1 \varepsilon_2 \dots \varepsilon_n) \quad (3)$$

$$\beta^T = [\beta_1 \beta_2 \dots \beta_k] \quad (4)$$

$$f^T(x) = (f_1(x) f_2(x) \dots f_k(x)) \quad (5)$$

$$F = \begin{bmatrix} f_1(x_1) & \dots & f_k(x_1) \\ \vdots & \ddots & \vdots \\ f_1(x_n) & \dots & f_k(x_n) \end{bmatrix} = [f(x_1) f(x_2) \dots f(x_n)]^T \quad (6)$$

and $f_i(x_j) \equiv x_{ij}$, for $i = 1, 2, \dots, k, j = 1, 2, \dots, n$. Then the response surface function is defined as formula $y = F\beta$. The realization of the classical design of experiments leads to the estimation of parameters of the response surface function defined as [Wawrzynek 2009]

$$y(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \beta_{12} x_1 x_2 + \dots + \beta_{12\dots m} x_1 x_2 \dots x_m. \quad (7)$$

The estimation of the parameters of the response surface function is usually carried out with the use of the least squares method. The use of this method depends on certain assumptions being fulfilled, which in the case of many different production processes cannot always be ensured [Montgomery 2001]. Then, in order to estimate the parameters of the response surface function, one may use the bootstrap method of the estimation of response surface function [Szerszunowicz 2014].

3. The theory of the optimization of design of experiments

The experimental design methodology is not limited only to issues of the analysis of variance or problems of the analysis of regression. An important issue in the theory of design of experiments is the optimization of the design of experiments. One may understand the theory of the optimization of the design of experiments in two ways: as the search for conditions for the optimization of product quality and as the optimal design of experiments.

The search for conditions for the optimization of product quality consists in describing the values of factors for which the response surface function reaches the appropriate value. The most frequently used methods of the optimization of product quality are [Wawrzynek 2009]:

- the Box–Wilson method,
- Evolutionary Operation (EVOP),
- Taguchi methodology.

The abovementioned methods require the following: the fulfillment of some assumptions and the use of some elements of differential calculus (the Box–Wilson method); numerous experimental trials with small changes in the levels of the process operating variables (EVOP); the determination of the appropriate orthogonal arrays (Taguchi methods). In practice, the realization of the abovementioned methods of optimization may be impossible or too time-consuming (prolonged duration of production processes), which significantly influences their economical results.

The optimal design of experiments consists in choosing the proper design of experiment which is the best one according to a certain optimality criterion. The most popular criteria are [Ryan 2007]:

- D-optimality criterion;
- G-optimality criterion;
- A-optimality criterion.

The theory of the optimal design of experiments is widely described in specialist literature by F. Puckelsheim [1993] and S.D. Silvey [1980], among others. Over the years, the optimality criteria have received criticism because of practical reasons [Ryan 2007].

4. The method of the search for conditions for process optimization and its application

Within specialist literature, there have been a few publications taking into account the design of experiment realization costs [Joiner, Campbell 1976; Draper, Stoneman 1968; Ryan 2007]. The present article discusses the issue of the search for conditions for the optimization of product quality with particular reference to the economic aspect of manufacturing processes.

4.1. The proposed method

Let us assume a design of experiment which takes into account m factors occurring at $n_i, i = 1, 2, \dots, m$ levels. Also, let the response surface func-

tion be given by the formula (7). An increased number of factors at many levels means a large number of experimental trials; the realization of trials in all the experimental area points is impossible because of either the costs of the experimental trials or the prolonged time of the experiment. In either case, the experimenter's task is to select from among the points of the experimental area of $n = \prod_{i=1}^m n_i$ the possible points which allow one to obtain the best results. The proposed algorithm for determining the points of the experimental area consists in defining a confidence area for the response surface function with the use of the estimation of confidence intervals for the parameters of the response surface functions. If the assumptions of the least squares method have been fulfilled, one may use the estimation of the confidence interval for parameter $\hat{\beta}_i$ as follows [Zieliński 2007]:

$$\left(\hat{\beta}_i - t_{\alpha, n-p} s \sqrt{v_{ii}}, \hat{\beta}_i + t_{\alpha, n-p} s \sqrt{v_{ii}} \right), \quad (8)$$

where

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

$t_{\alpha, n-p}$ – quantile of Student's t-distribution with $n - p$ degrees of freedom,

v_{ii} – appropriate element of matrix $(X^T X)^{-1}$.

However, if the use of the classical method of estimation of response surface functions is impossible or – in the case of a small number of observations – unreliable, one should use the bootstrap confidence intervals [Domański, Pruska 2000]. A bootstrap distribution of each $\hat{\beta}_i$ parameter is obtained as a result of the use of the bootstrap estimation of response surface function. Consequently, the confidence interval, with confidence level at $1 - \alpha$, is defined as follows [Kończak 2012]:

$$\left(\hat{\beta}_{i, \alpha/2}^*, \hat{\beta}_{i, 1-\alpha/2}^* \right), \quad (9)$$

where $\hat{\beta}_{i, \gamma}^*$ – quantile of γ order of the empirical distribution of $\hat{\beta}_i$ parameter.

On the basis of the confidence intervals for the response surface parameters, one may determine the upper $\tilde{y}_U(x)$ and lower $\tilde{y}_L(x)$ estimations of the response surface function. Then, from all the points of the experimental

area, the experimenter ought to select the points for which the values $|\tilde{y}(x) - \tilde{y}_U(x)|$ and $|\tilde{y}(x) - \tilde{y}_L(x)|$ are the lowest.

4.2. The use of the proposed algorithm

T.P. Ryan (2007), considers an experiment which includes two factors – Supplement (X_1) and Whey (X_2) – and the response variable Quality (Y). The experimental data are presented in Table 1.

Table 1. The experimental data

X_1	X_2	Y
Supplement	Whey	Quality
1	0	4,4
1	10	4,633
1	20	4,7
1	30	4,8
2	0	3,2
2	10	3,7
2	20	5,033
2	30	5,433

Source: [Ryan 2007].

For this experimental data, the purpose of the experimenter's research is to select the points of the experimental area in which the values of the estimation of the response surface function are the best according to the proposed algorithm. It has been established that it is possible to conduct only four experimental trials. In order to present the proposed approach, the classical and bootstrap methods have been used.

Let us assume that the assumptions of the least squares method are fulfilled. Then the estimated response surface function is given in the form of

$$\tilde{y}(x) = 4,57 - 0,63x_1 + 0,1267x_2 + 0,6767x_1x_2. \quad (10)$$

According to the proposed scheme, for each parameter of the response surface function the confidence interval (8) was computed. The results are presented in Table 2.

On the basis of the obtained results, the upper and lower bounds of the response surface function are defined respectively as follows:

$$\tilde{y}_U(x) = 5,171 + 0,5716x_1 + 0,5272x_2 + 1,478x_1x_2, \quad (11)$$

$$\tilde{y}_L(x) = 3,97 - 1,8316x_1 - 0,2739x_2 - 0,1244x_1x_2. \quad (12)$$

Table 2. The estimation of the response surface parameters with the appropriate confidence intervals (the classical method)

β_i	$\hat{\beta}_i$	$\hat{\beta}_{iL}$	$\hat{\beta}_{iU}$
β_0	4,57	3,969	5,171
β_1	-0,63	-1,832	0,572
β_2	0,127	-0,274	0,527
β_3	0,677	-0,124	1,478

Source: own elaboration.

Figure 1 shows the graphical representation of the obtained results.

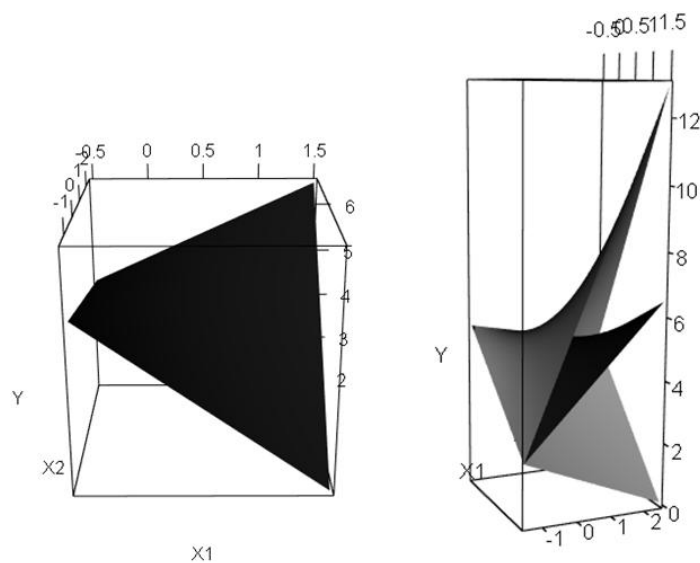


Fig. 1. The graphical representation of the response surface function (left) with the upper and lower bounds (right) obtained with the classical method

Source: own elaboration.

Then the values of $|\tilde{y}_i - \tilde{y}_{Li}|$, $|\tilde{y}_i - \tilde{y}_{Ui}|$ and $|y_i - \tilde{y}_i|$ were computed. The results are presented in Table 3.

Having analyzed the results, one can note that the design of the experiment should be defined as follows:

$$P_{MNK} = \{x_1, x_2, x_3, x_5\}.$$

Table 3. Values of appropriate estimations (the classical method)

x_i	$ \tilde{y}_i - \tilde{y}_{Li} $	$ \tilde{y}_i - \tilde{y}_{Ui} $	$ y_i - \tilde{y}_i $
x_1	0.2003	0.2003	0.04333
x_2	0.6008	0.6008	0.0633
x_3	1.00135	1.00135	0.0033
x_4	1.40189	1.40189	0.0233
x_5	0.6008	0.6008	0.0633
x_6	1.8024	1.8024	0.24
x_7	3.004	3.004	0.29
x_8	4.2057	4.2057	0.1133

Source: own elaboration.

Parallel considerations were made with the use of the bootstrap method. The estimated values of the response surface parameters with the corresponding confidence intervals are presented in Table 4.

Table 4. The estimation of the response surface parameters with the appropriate confidence intervals (the bootstrap method)

β_i	$\hat{\beta}_i^*$	$\hat{\beta}_{i0,025}^*$	$\hat{\beta}_{i0,975}^*$
β_0	4.572	4.459	4.6953
β_1	-0.634	-0.851	-0.415
β_2	0.125	-0.0037	0.251
β_3	0.6795	0.4993	0.8534

Source: own elaboration.

Then the response surface function can be expressed with the following equation:

$$\tilde{y}^*(\mathbf{x}) = 4,527 - 0,634x_1 + 0,125x_2 + 0,6795x_1x_2, \quad (13)$$

and the upper and lower bounds of the designed response surface function are as follows:

$$\tilde{y}_U^*(x) = 4,6953 - 0,415x_1 + 0,251x_2 + 0,853x_1x_2 \quad (14)$$

$$\tilde{y}_L^*(x) = 4,459 - 0,851x_1 - 0,0037x_2 - 0,4993x_1x_2. \quad (15)$$

Figure 2 shows the graphical representation of the obtained estimation of the response surface functions.

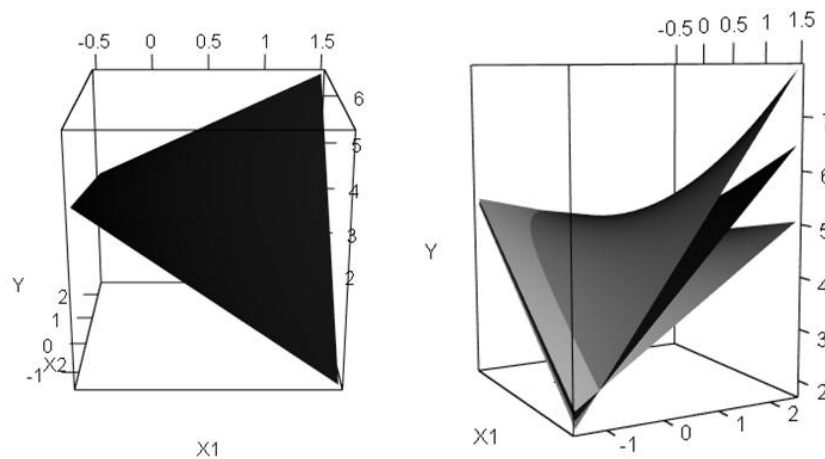


Fig. 2. The graphical representation of the response surface function (left) with the upper and lower bounds (right) obtained with the bootstrap method

Source: own elaboration.

Just like with the classical method, for every point of the experimental area the values of $|\tilde{y}_i^* - \tilde{y}_{L_i}^*|$, $|\tilde{y}_i^* - \tilde{y}_{U_i}^*|$ and $|y_i - \tilde{y}_i^*|$ were computed and presented in Table 5.

According to the proposed method of the optimization of product quality, the points in which one ought to conduct the next experimental trials were selected:

$$P_{MB} = \{x_1, x_2, x_3, x_5\}.$$

It is important to notice that this design of experiment involves the same points as the selected design with the classical approach.

Table 5. Values of appropriate estimations (the bootstrap method)

x_i	$ \tilde{y}_i^* - \tilde{y}_{Li}^* $	$ \tilde{y}_i^* - \tilde{y}_{Ui}^* $	$ y_i - \tilde{y}_i^* $
x_1	0.0147	0.1265	0.048
x_2	0.1135	0	0.0612
x_3	0.2417	0.1265	0.0033
x_4	0.3699	0.2529	0.0212
x_5	0.022	0.0818	0.0655
x_6	0.3306	0.2186	0.2386
x_7	0.639	0.5189	0.291
x_8	0.947	0.8193	0.1133

Source: own elaboration.

On the basis of the results presented in Table 2 and Table 4, one can state that the form of response surface functions (10) and (13) are similar, but the determined confidence intervals for the response surface parameters are narrower in the case where the bootstrap method of the estimation of the response surface function was used. Therefore, a more precise estimation of the response surface function can be expected when the bootstrap method is used. These suppositions can be confirmed through comparing the results presented in Table 3 and Table 5. There are comparable differences between the empirical and theoretical values, but the differences between the upper and lower estimations of the response surface function are also lower for the bootstrap method than for the least squares method.

Thanks to the use of the proposed scheme for searching for the conditions for the optimization of product quality, the number of experimental trials was reduced, and the relevant characteristics of the results of the manufacturing process were preserved.

5. Conclusions

The design of experiments is one of the tools of statistical quality control allowing one to improve the results of the production processes effectively. The theory of the optimization of the design of experiments in particular leads to defining the design of the experiment which is appropriate because of the optimal criterion or conditions of product quality. An im-

portant issue in terms of planning the results of production processes is taking into account the economic aspect of the design of experiments.

The proposed algorithm for searching for the conditions for the optimization of manufacturing processes allows one to determine the design of experiments for a determined number of experimental trials. Moreover, this paper presents the possibility of using the bootstrap method. In this approach, the estimated confidence area for a response surface function is characterized by greater precision in comparison to the classical approach. The possibilities of the presented method of searching for the optimal conditions for manufacturing processes result in reductions in experiment cost and duration, and also allow one to maintain the proper technological level of production processes.

Acknowledgements

The research was supported by the Polish Science Center grant DEC-2011/03/B/HS4/05630.

References

- Domański Cz., Pruska K. (2000). *Nieklasyczne metody statystyczne*. PWE. Warszawa.
- Draper N.R., Stoneman D.M. (1968). *Factor changes and Linear Trends in Eight-Run Two-Level Factorial Designs*. Technometrics. No. 10. Pp. 301-311.
- Joiner B.L., Campbell C. (1976). *Designing Experiments When Run Order is Important*. Technometrics. No 18. Pp. 249-258.
- Kończak G. (2012). *Wprowadzenie do symulacji komputerowych*. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach. Katowice.
- Montgomery D.C. (2001). *Design and Analysis of Experiments*. John Wiley & Sons Inc. New York.
- Montgomery D.C. (1997). *Introduction to statistical quality control*. John Wiley & Sons Inc. New York.
- Puckelsheim F. (1993). *Optimal Design of Experiments*. John Wiley & Sons Inc. New York.
- Ryan T.P. (2007). *Modern Experimental Design*. John Wiley & Sons. New Jersey.
- Silvey S.D. (1980). *Optimal Design*. Chapman and Hall. London.
- Szerszunowicz M. (2014). *On the Bootstrap Method of Estimation of Response Surface Function*. Acta Universitatis Lodziensis. Folia Oeconomica. No. 3 (302). Pp. 101-118.
- Wawrzynek J. (2009). *Planowanie eksperymentów zorientowane na doskonalenie jakości produktu*. Wydawnictwo Uniwersytetu Ekonomicznego. Wrocław.

Wawrzynek J. (1993). *Statystyczne planowanie eksperymentów w zagadnieniach regresji w warunkach małej próby*. Wydawnictwo Akademii Ekonomicznej. Wrocław.

Zieliński W. (2007). *Teoretyczne podstawy ekonometrycznych jednorównaniowych modeli liniowych*. Wydawnictwo SGGW. Warszawa.