ISSN 1330-3651
UDC/UDK 658.522:519.233

IMPROVEMENT OF EXPERIMENTAL PLANS USING MULTI-CRITERIA APPROACH

Hrvoje Cajner, Nikola Šakić, Predrag Ćosić

Original scientific paper
The process of designing experimental plans is based on the choice of the number and placement of design points. The question which should be answered is how to choose the design points and number of replications to achieve a more precise model. Due to the limitation of research resources like time and costs of performing certain design points, it is necessary to take these constraints into account when designing experiments. The mentioned time consumption, cost and model precision are defined as new criteria. The scope of the research is to define the procedure of maximizing model precision using multi-criteria approach. The study refers to the research and development of a new procedure which will be a support in the process of improvement of experimental plans.

Keywords: *augmented design, design of experiments, optimality criteria, optimal design*

Poboljšanje planova pokusa višekriterijskim pristupom

Izvorni znanstveni članak
Proces planiranja i oblikovanja eksperimenata u osnovi se bazira na odabiru broja i rasporeda stanja pokusa (eksperimentalnih točaka). Cilj kvalitetnog procesa planiranja eksperimenata je postizanje optimalnog oblika plana pokusa obzirom na maksimizaciju preciznosti konačnog modela. Zbog ograničenosti resursa poput utroška vremena i direktnih troškova procesa izvođenja eksperimenta, potrebno je ta ograničenja uzeti u obzir i kako pri izradi eksperimenata, tako i poboljšanju postojećih modela. Navedena ograničenja u resursima (vrijeme i trošak) bit će definirani kao novi kriteriji. U radu će biti definiran postupak modifikacije postojećeg plana pokusa uzimajući novouvedene kriterije maksimizirajući preciznost rezultirajućeg modela. Novi pristup modifikaciji i poboljšanju plana pokusa moguće je koristiti kao podršku klasičnom pristupu oblikovanju eksperimenata.

Ključne riječi: *kriteriji optimalnosti, modificirani plan pokusa, optimalni plan pokusa, planiranje pokusa*

1 Introduction

Design of experiments (in further text DOE) is an unavoidable methodology for improving the quality of products and processes. This methodology is an important part of improving products or processes, and the need for choosing appropriate and more precise experimental design is crucial.

Sometimes the problem occurs when the experimental region is unknown and the experimenter is not satisfied with precision level of resulting model. That means that the design should be modified by adding more experimental points. The paper explains the new-developed criteria which guide the experimenter in the process of adding the design points. One of the new-developed criteria (incremental precision criterion) gives full information about the possible candidate points in a way of contribution to the model precision. This criterion is related with other significant characteristics of models such as costs and time consumption of certain design (experimental) points.

Furthermore, direct costs can become a limiting factor. With the combination of time limit (time consumption) it could cause several problems related with experiment performance. When the experiment is in progress, it is very difficult to carry out the modification of the initial model and there is a need for the use of an efficient method of model correction. Problems that could arise from the lack of the mentioned resources can lead to the main side effects:

1) Exceeding the research deadline,
2) Exceeding available financial resources,
3) Incompleteness of the model.

The first two side effects are understandable and there is no need for detailed discussion. But the third point states out the most undesirable characteristic when performing the experimental research. The result of incomplete design model is a lack of experimental points which requires the afterwards performing of the missing data. The afterwards testing spends a certain amount of time and financial resources so it is crucial to choose the points that are most appropriate. If there is no intention for improvement of accuracy (quality) assessment of response values and there is no intention of performing additional experiments the solution could be found in methods discussed in the literature sources [1] and [2].

The following sections describe the process of improving the experimental model regarding the precision (quality of response values) that applies to cases with incomplete data.

2 The concept of multi-criteria approach in modification of experimental plans

Previous mentioned criteria (costs, time consumption) will be combined with an incremental precision criterion (which will be defined in the following chapters) forming the desirability function. Each of the mentioned criteria is defined in the form of the function over the experimental area, bounded by the value of factor levels. Since this is a problem of designing of experimental plan in the specific area of response surface methodology, the factors in the experiment are independent, quantitative, continuous variables. The factors in designed experimental plans being continuous variables, in the same way the functions of the above mentioned criteria will also be continuous functions. From this basic characteristic of continuity it can be accepted that the distribution of costs and time consumption can be approximated by an appropriate form of regression function.

2.1 Definition of cost and time consumption functions

Because of similar characteristics of the cost and time consumption functions the definition will be presented only for the one term (cost function). The total cost of experimentation is composed of a constant and a variable part. The constant part of the cost of experimentation takes into account general costs related to the preparation of the experiment, and it is not affected by the number of experimental samples and the number of replications. In other words, the modification of the shape of experiment will not affect the constant part. Contrary to the constant part of costs, the variable part is associated with the direct costs of preparation and conducting of experimental samples and is strictly related with the performance of design points. This means that by adding or modifying of design points (experimental points) the change in overall cost will occur.

The following flow chart (Fig. 1) represents the structure of algorithm which defines the function of the cost of the experiment.

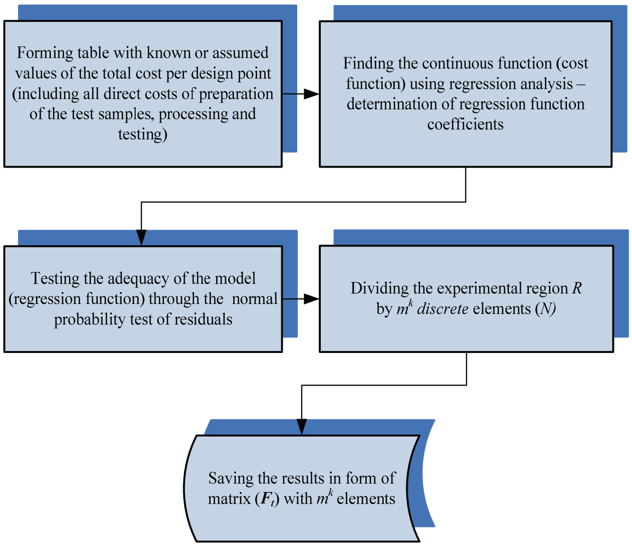

Figure 1 Definition of cost function

The cost function can have different shapes. The researcher is one who may have information about the explicit form of the function, which is unlikely, or estimate it by one of the existing methods such as regression analysis, nonlinear estimation [3]. The cost function is estimated like most of the response functions in RSM (Response surface methodology) by equation:

$$f_t\left(x_1, x_2 \ldots, x_k\right) = C_{T0} + \sum_{i=1}^{k} C_{Ti} x_i + \sum_{i=1}^{k} C_{Tii} x_i^2 + \sum_{i} \sum_{<j} C_{Tij} x_i x_j \quad (1)$$

where C_{T0} is intercept, C_{Ti} and C_{Tii} represent the increase of cost due to the change of main factors x_i and C_{Tij} is the coefficient which is related to the change in cost due to possible interaction of factors. In perspective of cost function, the interaction would mean that there is some variation in the cost due to the complexity of the performance of some combinations of factor values, and is not related to the individual increment of cost as a result of the main effects.

The resulting cost function (case with 2 factors) is defined through experimental area and strictly applies only to that area (Fig. 2). By picking any combination of factor levels (design points) it is possible to obtain the estimated value of the related cost.

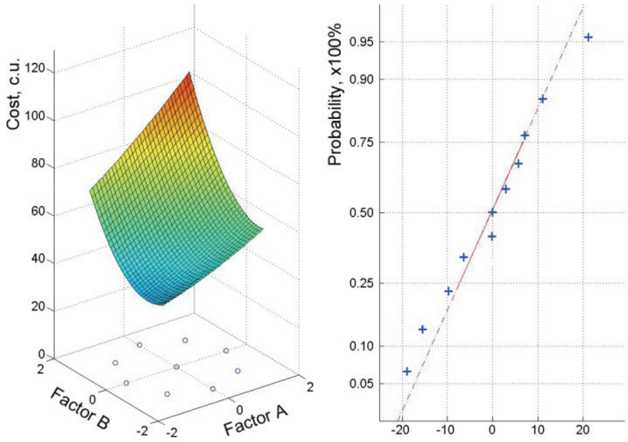

Figure 2 Graphical representation of cost function and normal probability plot of residuals (case with 2 factors)

By means of the transformation into a discrete form, the function is transformed into the matrix (array) containing discrete values of the cost. Number of discrete values depends on the density of the mesh grid (size of *N*). The transformation of cost function into matrix (for the case of two factors) can be written:

$$\boldsymbol{F}_t = \left[f_t(z_{ji}) \right], \quad (2)$$

where $f_t(z_{ji})$ is the value of cost function at a specific design point with coordinates $T(x_{1(i)}, x_{2(j)})$.

The function of time consumption is defined analogous to the cost function, so there is no need for a detailed description. The matrix form for the function of time consumption is expressed in equation 3.

$$\boldsymbol{F}_v = \left[f_v(z_{ji}) \right], \quad (3)$$

where $f_v(z_{ji})$ is the value of time consumption function at specific design point with coordinates $T(x_{1(i)}, x_{2(j)})$.

2.2 Incremental precision criterion

One of the basic facts which describe the adequacy of model is the model precision. Precision can be formulated throughout the error which comes out from design point performance. The precision is expressed by a standard error based on estimated variance and number of repetitions. Moreover, standard error is used for defining the confidence interval of a mean response which gives precise information of possible variation of the observed (measured) phenomena. The main expressions related to the above mentioned terms are thoroughly described in the literature [3].

Precision of model regarding whole design region could be defined as a portion of design space bounded by upper and lower limits of confidence interval. Reviewing the equation which describes confidence interval of mean responses, the conclusion is that the confidence intervals variability is only related to the value of $s_{\hat{y}(x)}$ (standard error) and confidence level. So the precision of model $Q(\zeta_\mathrm{d})$ could be expressed by the value of predicted variance over the design region (equation 4), of course as a relative measure which is suitable for comparing different designs.

$$Q(\zeta_\mathrm{d}) = s^2 \cdot \int_{\mathcal{N}} \boldsymbol{x}'^{(m)} (\boldsymbol{X}(\zeta_\mathrm{d})' \boldsymbol{X}(\zeta_\mathrm{d}))^{-1} \boldsymbol{x}^{(m)} \mathrm{d}x \,, \tag{4}$$

Because of assumption of $s^2 = 1$ and development of the computer algorithm, which is necessary to verify results and better understanding of computing process, the equations will be expressed in matrix form:

$$Q(\zeta_\mathrm{d}) = \sum_{i=1}^{m^k} \left\{ \boldsymbol{x}_i{'}^{(m)} (\boldsymbol{X}(\zeta_\mathrm{d})' \boldsymbol{X}(\zeta_\mathrm{d}))^{-1} \boldsymbol{x}_i^{(m)} \right\} \,, \tag{5}$$

where m^k = N is number of discrete elements which represent the design region. It can be noticed that smaller values of Q indicate higher accuracy, thus larger values of Q indicate lower accuracy. By means of transforming the equations from continuous form to discrete (matrix form) the function of model precision (function of model precision matrix) can be expressed by equation:

$$\boldsymbol{F}(\zeta_\mathrm{d}) = \boldsymbol{x}'^{(m)} (\boldsymbol{X}(\zeta_\mathrm{d})' \boldsymbol{X}(\zeta_\mathrm{d}))^{-1} \boldsymbol{x}^{(m)}. \tag{6}$$

A short definition of incremental precision of model can be defined through the amount of gained precision as a result of alternately adding new design points to the initial model (candidate points). If notation of design matrix for initial design is $\boldsymbol{X}(\zeta_\mathrm{p})$ and for the augmented design is $\boldsymbol{X}_\mathrm{aug}(\zeta_\mathrm{p})$ the process of incremental precision can be derived:

$$\boldsymbol{X}_\mathrm{aug}(\zeta_\mathrm{p}) = \begin{bmatrix} \boldsymbol{X}(\zeta_\mathrm{p}) \\ \hline\hline x_\mathrm{aug}{'}^{(m)} \end{bmatrix} \tag{7}$$

where $x_\mathrm{aug}{'}^{(m)}$ is a row aided to the design matrix and reflects the aided point in one iteration, and furthermore the function of model precision matrix is

$$\boldsymbol{F}_\mathrm{aug}(\zeta_\mathrm{p}) = \boldsymbol{x}'^{(m)} (\boldsymbol{X}_\mathrm{aug}(\zeta_\mathrm{p})' \boldsymbol{X}_\mathrm{aug}(\zeta_\mathrm{p}))^{-1} \boldsymbol{x}^{(m)}. \tag{8}$$

Model precision of augmented design is given by generic form shown in equation:

$$Q_\mathrm{aug}(\zeta_\mathrm{p}) = \sum\nolimits_N \boldsymbol{F}_\mathrm{aug}(\zeta_\mathrm{p}) = \sum_{i=1}^{m^k} \left\{ \boldsymbol{x}_i{'}^{(m)} (\boldsymbol{X}_\mathrm{aug}(\zeta_\mathrm{p})' \boldsymbol{X}_\mathrm{aug}(\zeta_\mathrm{p}))^{-1} \boldsymbol{x}_i^{(m)} \right\} \tag{9}$$

Using equation (9) and model precision matrix for initial design:

$$Q(\zeta_\mathrm{p}) = \sum_{i=1}^{m^k} \left\{ \boldsymbol{x}_i{'}^{(m)} (\boldsymbol{X}(\zeta_\mathrm{p})' \boldsymbol{X}(\zeta_\mathrm{p}))^{-1} \boldsymbol{x}_i^{(m)} \right\}, \tag{10}$$

and by subtracting them the incremental precision at specific location $x_\mathrm{aug}{'}^{(m)}$ can be expressed:

$$\Delta q = \sum\nolimits_N \left[\boldsymbol{F}(\zeta_\mathrm{p}) - \boldsymbol{F}_\mathrm{aug}(\zeta_\mathrm{p}) \right] = Q(\zeta_\mathrm{p}) - Q_\mathrm{aug}(\zeta_\mathrm{p}). \tag{11}$$

In the case of two factors, the matrix $\Delta\boldsymbol{Q}$ (equation 12) consists of elements Δq_{ij} whose values indicate the change in the model precision by adding specific design points $\mathrm{T}(x_{1(i)}, x_{2(j)})$ which is represented by $x_\mathrm{aug}{'}^{(m)}$ and correspond with indices in matrix. For example the element with index (1,1) reflects the incremental precision of model at a point which has coordinates: T(minimum level of first factor, minimum level of factor 2).

$$\underset{(m\times m)}{\Delta\boldsymbol{Q}} = \left[\Delta q_{ji} \right] = \begin{bmatrix} \Delta q_{11} & \Delta q_{1i} & \dots & \Delta q_{1m} \\ \Delta q_{j1} & \Delta q_{ji} & \dots & \Delta q_{jm} \\ \dots & \dots & \dots & \dots \\ \Delta q_{m1} & \Delta q_{mi} & \dots & \Delta q_{mm} \end{bmatrix}. \tag{12}$$

The $\Delta\boldsymbol{Q}$ can be called the incremental precision matrix which is fulfilled after evaluation of model precision for all points in the design region. The element in the matrix with a higher value means which design point is the best for the initial design upgrading, regarding the increase of precision.

2.3 Definition of desirability function

The desirability function is a result of addition of functions: cost, time consumption and function of incremental precision. Since continuous functions are transformed into matrix form of size m^k, desirability functions will also be expressed through the matrix $\boldsymbol{F}_\mathrm{p}$. The shape of desirability function depends on the shape of the individual functions and their weighting factors. Weighting factors are defined throughout Likert scale and are the result of experimenter choice. The experimenter (researcher) picks the weighting factor depending on importance of the mentioned criteria (cost, time consumption and model precision). Finally the desirability function is calculated in discrete form and is expressed:

$$\boldsymbol{F}_\mathrm{p} = k_\mathrm{t} \cdot \left(\boldsymbol{1} - \frac{\boldsymbol{F}_\mathrm{t}}{\max(\boldsymbol{F}_\mathrm{t})} \right) + k_\mathrm{v} \cdot \left(\boldsymbol{1} - \frac{\boldsymbol{F}_\mathrm{v}}{\max(\boldsymbol{F}_\mathrm{v})} \right) + k_\mathrm{pr} \cdot \left(\frac{\Delta\boldsymbol{Q}}{\max(\Delta\boldsymbol{Q})} \right) \,, \tag{13}$$

where k_t, k_v and k_pr are weighting factors. In order to obtain equality of criteria which are expressed in different units, it is necessary to transform them in a relative, dimensionless form. Values of particular function criteria will vary between 0 and 1. The matrix $\boldsymbol{F}_\mathrm{p}$ should be also standardized by the expression:

$$\boldsymbol{F}_{\mathrm{p}}^{\mathrm{r}} = \frac{\boldsymbol{F}_{\mathrm{p}}}{\max(\boldsymbol{F}_{\mathrm{p}})} \tag{14}$$

The elements of $\boldsymbol{F}_{\mathrm{p}}^{\mathrm{r}}$ matrix represent values which indicate which points are more suitable for including in the model. The points with the highest values of desirability are first candidates for modification of the initial experimental model.

The modification of design is a process of iteration and is graphically represented in Fig. 3. The iteration process is realized as computer algorithm and is developed in software MATLAB.

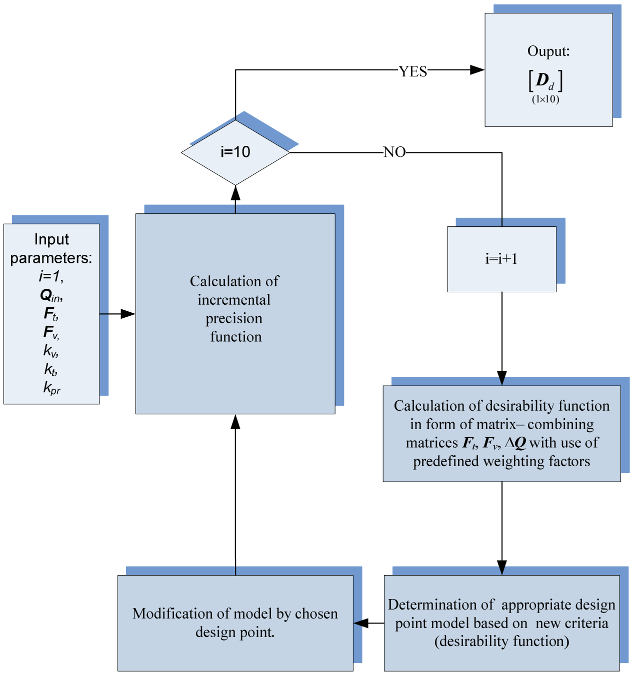

Figure 3 Process of modification of design model based on desirability function

The process of iteration in the algorithm is limited to ten steps. More precisely, ten steps mean that the initial experimental model is modified by adding nine design points. The algorithm was tested on problems with two or three factors and it was found that after the sixth iteration the accuracy of model was significantly improved, while other indicators of goodness of the model were maintained at acceptable values. The result is a set of modified experimental models which are stored in array:

$$\underset{(1\times10)}{[\boldsymbol{D}_d]} = \left[\underset{(N\times k)}{[\boldsymbol{D}(\zeta_p)]} \quad \underset{((N+1)\times k)}{[\boldsymbol{D}(\zeta_d)]} \quad \dots \quad \underset{((N+10)\times k)}{[\boldsymbol{D}(\zeta_{10})]} \right] \tag{15}$$

The next step is to select the appropriate modified model from the given set of solutions. The optimal model is the one which has the largest difference in the contribution to model precision, comparing to the adjacent model (Eq. 16).

$$\Delta G = \left| Q(\zeta_{d-1}) - Q(\zeta_d) \right| \tag{16}$$

Final choice of optimal design is on the experimenter and the results derived from the iteration process are only a suggestion.

One of the main characteristics which are affected by modification of experimental model is the loss of orthogonality. In a process of adding design points it is necessary to track the value of orthogonality, so when interpreting results (especially effects values) that fact must be taken into account. Nevertheless, if the selected design has acceptable loss of orthogonality and adequate values of other known optimality criteria such as G-optimality and D-optimality [7], [3], it can be indicated as optimal design.

3 Application of proposed procedure - example

The application of the developed procedure will be shown on the example elaborated in literature source [6]. The research is related to examination of life prediction of damaged polyethylene gas pipes. The initial (chosen) DOE model was a typical central composite design (CCD) with three factors. After the experiment was conducted the main problem occurred. Some of the samples (design points) had an extremely high value of response (life prediction time) so the whole research exceeded the predefined deadline. Because of the mentioned problem (lack of resulting data for one design point), the DOE model was incomplete. The formulation of the response surface in the case of incompleteness results contains a relatively large estimated variance, which means that the accuracy of estimation of response values is reduced. Certainly this is not a desirable feature so there is a need for increasing of model precision by means of the proposed procedure.

The initial design in coded values is represented in Tab. 1.

Table 1 Initial DOE model

Point num.	Length of cut (x_1)	Depth of cut (x_2)	Pressure (x_3)	Time consumption / h
1	–1	–1	–1	14066,40
2	–1	–1	1	19,64
3	–1	1	–1	181,21
4	–1	1	1	0,36
5	1	–1	–1	13323,17
6	1	–1	1	20,40
7	1	1	–1	95,88
8	1	1	1	0,01
9	–1,682	0	0	221,65
10	1,682	0	0	18,84
11	0	–1,682	0	8608,47
12	0	1,682	0	0,01
13	0	0	–1,682	**11903***
14	0	0	1,682	0,01
15(C)	0	0	0	24,71
16(C)	0	0	0	16,52
17(C)	0	0	0	32,48
18(C)	0	0	0	20,50
19(C)	0	0	0	23,11

As shown in results, the missing data is at design point 13. The time consumption function is equal to the response function and is defined:

$$f_v^{0,2} = 28,57 - 0,042x_1 - 11,18x_2 - 2,12x_3 + 0,237x_2x_3 + 3,71\times10^{-4}x_1^2 + 2,34x_2^2 + 0,059x_3^2 - 1,05\times10^{-6}x_1^3 - 0,209x_2^3 \quad , \qquad (17)$$

whilst the cost function is defined as constant because of equality of costs per design point. In this case the result of procedure will not be affected by the cost function. Matrices of standardized criteria functions are placed in relation (Eq. 13) with chosen weighting factors $k_{pr} = 5$, $k_t = 2$ and $k_v = 9$.

The resulting desirability function is modified in the process of iteration. After each iteration new desirability function is created and the model is augmented with one design point. The results of the process of iteration are a set of augmented designs. Fig. 4 shows the change of the criteria values by each iteration.

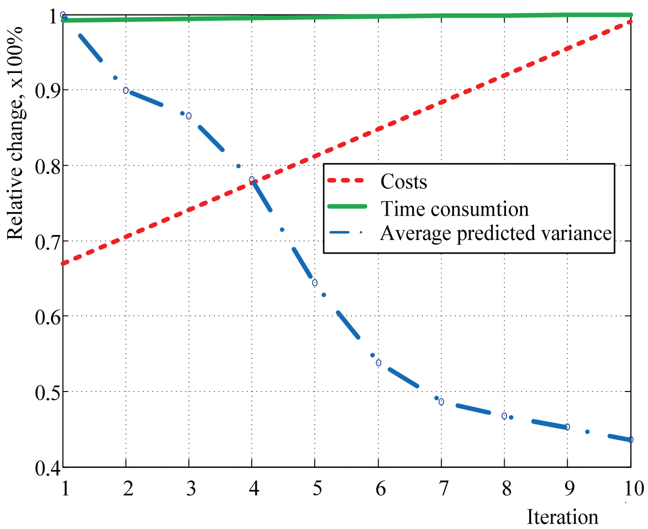

Figure 4 Relative change in criteria after ten iterations

It is clear that the augmented design after sixth iteration shows the reduction of average predicted variance near to 50 %, which is a significant gain of the model precision. Furthermore, the most important criterium, time consumption, has a slight increase by 1 % which is negligible. That means that the design is augmented with several points which have a relatively small value of time consumption. The costs have a higher value (up to 25 % of initial design) which is understandable and not that significant for the experiment because of the previously chosen level of importance ($k_t = 2$).

Table 2 Set of points for augmentation of initial design (coded values)

Point num.	Length of cut (x_1)	Depth of cut (x_2)	Pressure (x_3)
1.	1,68	1,68	1,68
2.	–1,68	1,68	1,68
3.	1,20	1,68	–1,68
4.	–1,68	1,68	–1,68
5.	0	0,24	–1,68

Looking at the other characteristics of design such as orthogonality, G-optimality, D-optimality there is no significant loss, so the particular design could be declared as optimal design after the sixth iteration. Finally the optimal design is results of augmentation of initial design by means of adding 5 extra points. These 5 points should be conducted under the same conditions as were those in initial design.

For the purpose of improvement of the model, which was incomplete, this procedure gives an optimal solution which is acceptable from the standpoint of the resources expenditure (especially time consumption). Using this procedure the relatively high estimated variance could be improved, keeping in mind the duration of the experimental research and the set deadlines.

4 Conclusion

Experiments with high costs and large time consumption (especially in the variable part) are most suitable for the application of the described procedure, which could result in substantial resource savings. If there is a need for improvement of model, the proposed procedure provides afterward modification of the model by increasing the quality assessment of the observed phenomena on the principle of maximizing the ratio of benefit vs. costs. The procedure was tested on example of incomplete data, but it can be used on any case when improvement of quality of results is needed. Furthermore, it is possible to answer the question how the values of some known optimality criteria change, indicating possible minor losses or improvements of model characteristics. The proposed procedure is developed as a computer algorithm and can have a significant part in the process of designing of experiments.

5 References

[1] Akhtar, M. Response surface Designs Robust to Missing Observations, doctorate, Southampton University, UK, 1985.
[2] Akram, M. Central composite designs robust to three missing observations, doctorate, The Islamia University Bahawalpur, Pakistan, 2002.
[3] Cox, D. R.; Reid, N. The Theory of the Design of Experiments, Chapman&Hall, Washington D.C., 2000.
[4] Myers, Raymond, H.; Montgomery, Douglas, C. Response Surface Methodology, Process and Product Optimization Using Designed Experiments, John Wiley & Sons, 1995.
[5] MATLAB 6.5, MathWorks, Licensed Software, Faculty of Mechanical Engineering and Naval Architecture University of Zagreb.
[6] Tonković, Z. Procjena preostalog vremena uporabe oštećenih polietilenskih plinovodnih cijevi, doctorate, Josip Juraj Strossmayer University of Osijek, Faculty of Mechanical Engineering Slavonski Brod, 2008.
[7] Kiefer, J.; Wolfowitz, J. Optimal design in regression problems. // Annual of Mathematical Statistics, 30, (1959), pp. 271-294.

Author's addresses

Hrvoje Cajner
University of Zagreb
Faculty of Mechanical Engineering and Naval Architecture
Department of Industrial Engineering
Ivana Lučića 1
10 000 Zagreb, Croatia
e-mail: hrvoje.cajner@fsb.hr

Nikola Šakić
University of Zagreb
Faculty of Mechanical Engineering and Naval Architecture
Department of Industrial Engineering
Ivana Lučića 1
10 000 Zagreb, Croatia

Predrag Ćosić
University of Zagreb
Faculty of Mechanical Engineering and Naval Architecture
Department of Industrial Engineering
Ivana Lučića 1
10 000 Zagreb, Croatia