MINING AND METALLURGY INSTITUTE BOR
UDK: 622

ISSN: 2334-8836 (Štampano izdanje)
ISSN: 2406-1395 (Online)

UDK: 681.3.06:008.4(045)=111
DOI: 10.5937/mmeb2401001M

Received: 16.04.2024.
Revised: 08.05.2024.
Accepted: 10.05.2024.

Original Scientific Paper
Engineering

Aleksandra Milosavljević[*1], *Marijana Pavlov-Kagadejev*[*2], *Predrag Stolić*[**3]

THE IMPORTANCE OF DESIGN OF EXPERIMENTS[***]

Orcid: 1) https://orcid.org/0000-0003-3841-7357; 2) https://orcid.org/0000-0003-1090-6351;
3) https://orcid.org/0000-0002-2574-4765

Abstract

Design of experiments (DOE) is very meaningful and applied in various investigations from science to the industry in order to optimize the process itself. There are several such techniques and each of them has its own advantages, so it is very important to know the basics of DOE. Besides that, every problem, technology, product etc. is unique, so the knowledge about those is crucial as the first step. The most relevant fact is the dependence among variables – input factors and output responses as well as mutual connection between factors. In order to demonstrate the usability and adaptability of DOE for various purposes, some examples are given in this paper.

Keywords: *design of experiments, DOE, statistics, software*

INTRODUCTION

Design of experiments (DOE) plays a key role in science as well as industry, medicine, engineering, etc. Without a well-designed experimental plan, it is difficult to draw the reliable conclusions and make progress in research and development. It is a methodology that enables scientists and engineers to study the relationship between the input and output variables. DOE is a part of statistics which consists of planning, conducting, analyzing and interpreting tests in order to obtain relations and rules between the process factors and their responses. Besides that, the experiments are used to test laws, theories, and hypotheses. Based on the results of experiments, we can confirm or deny the certain claims. The planned experiments help in process optimization and mathematical modelling as well in order to obtain the optimized values and predict some future trends.

The first principles of DOE were postulated by Ronald Fisher in the early 1920s, who first applied it in the field of agriculture [1]. He studied how various factors such as weather conditions, soil conditions, and similar, affect yields in agriculture. Although the DOE method was first used in an

[*] *Mining and Metallurgy Institute Bor, Alberta Ajnstajna 1, 19210 Bor, Serbia, e-mail: aleksandra.milosavljevic@irmbor.co.rs*
[**] *University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210 Bor, Serbia*
[***] *This work was financially supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, Contract on realization and financing of the scientific research work of the Mining and Metallurgy Institute Bor in 2024, Contract No.: 451-03-66/2024-03/ 200052 and University of Belgrade, Technical Faculty in Bor Contract No.: 451-03-65/2024-03/200131.*

agricultural context, the method has been applied successfully in the industry and all aspects of science since then. In the first time, the most common approach among all researchers was OVAT (One Variable at a Time) or OFAT (One Factor at a Time) [2]. This approach involves varying one factor over time while keeping other factors constant which requires resources and time and depends on the experience of operator. Those results are often unreliable, inefficient and may present a false process optimization. More efficient is to observe more factors at the same time in order to build the new or improved products or processes.

The fundamental principles in DOE are:

- factorial principle
- randomization
- replication
- blocking

The factorial principle shows how mu-ltiple factors (independent variables) influence a response (dependent variable). Randomization refers to the order of experiments, while replication is a rerun of complete experiments including setup. Replication increases the precision of experiment and also the signal-to-noise ratio when the noise originates from uncontrollable nuisance variables [3]. Blocking gives us an opportunity to restrict the randomization to one group of factors and later perform the other experiment with other factors. It is a method for increasing the precision removing the effect of known nuisance factors and batch-to-batch variability. So, the experiment is performed to samples of material from one batch, then to samples from another batch, and so on [3].

Over the years, the field of application of DOE has expanded and includes many areas [4 - 9]. Although foundations of DOE were posted very early, the expansion of application started from 2000. It could be noticed that DOE has the biggest role in medicine and engineering according to a number of publications which are given in Figure 1 [10].

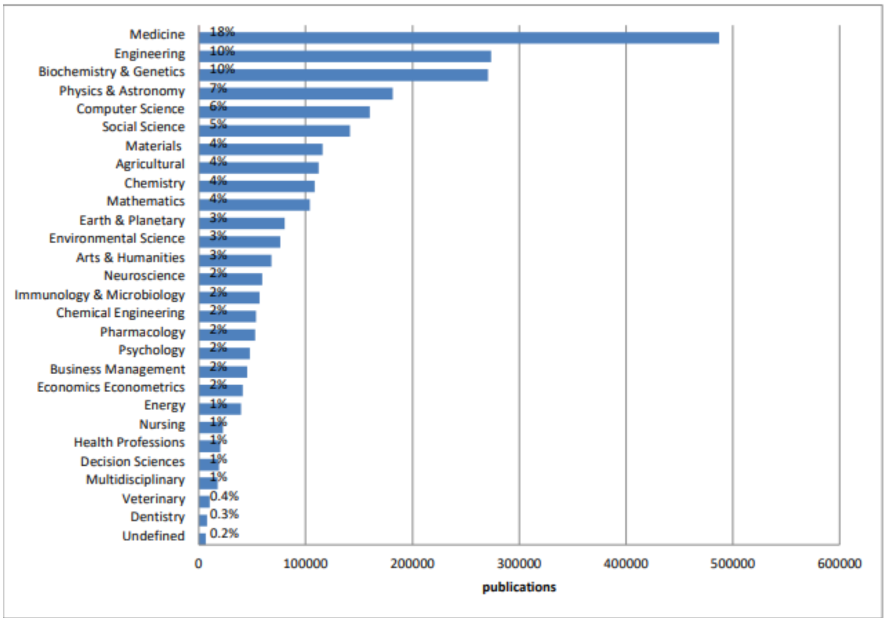


Figure 1 *An overview of the application of DOE to different scientific areas [10]*

THEORETICAL BACKGROUND OF DOE

Mathematics and statistics are found in the very theoretical basis of DOE, more precisely dispersion and regression analysis.

The basic task at the beginning is to determine the importance of the input factors (Xi) and dependence degree on the output responses (Yi) as it is shown in Figure 2.

Figure 2 *The beginning of DOE: input & output variables*

Different variants are possible: some input values - factors will significantly affect the output, while the influence of some factors will be completely negligible. Also, certain factors will have mutual dependence, so that must also be considered. Multifactorial experiment plans enable taking into account a large number of factors during research, for which there is no previous experience. After partial experiments, it is possible to eliminate those factors which are not significant from further consideration. This kind of case is the most complicated but is used when there are many unknown factors and can be described through a mathematical model:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ij} x_i x_j + \cdots$$

In the case of a three-factor experiment, the matrix of complete experiment (full factorial design) will be as it is shown in Table 1.

Table 1 *Matrix of full factorial design (three-factor)*

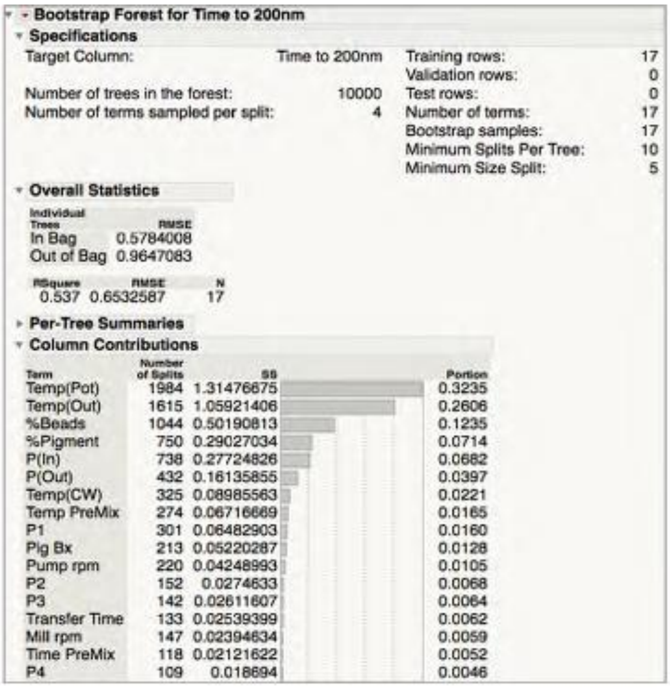
No	X0	X1	X2	X3	X1X2	X1X3	X2X3	X1X2X3
1	1	1	1	1	1	1	1	1
2	1	-1	1	1	-1	1	1	-1
3	1	1	-1	1	1	-1	-1	-1
4	1	-1	-1	1	-1	-1	-1	1
5	1	1	1	-1	1	1	-1	-1
6	1	-1	1	-1	-1	-1	-1	1
7	1	1	-1	-1	1	-1	1	1
8	1	-1	-1	-1	-1	1	1	-1

More matrices could be made from this matrix for fractional factorial design. Calculations can be very complex, but various software is used for these purposes and is constantly being updated, such as SPSS, Statistica, JMP, etc.

DOE EXAMPLES

Experimental design has been refined over the years and progressed from application the basic statistical models to the specialized software for these purposes. Some of them are free, while the more demanding ones are not available to everyone. Below are some examples of created DOE using different software.

The development of new products or establishment the existing technologies is always a challenge in terms of the required funds. In order to avoid unnecessary costs that would be incurred by performing a large number of experiments, it is necessary to reduce that number to a minimum. Therefore, an example of LCD manufacturer, which is trying to satisfy a customer demand, was given using DOE to perform only those experiments that are assumed to be usable. The company needs pigment particles to be milled down to less than 200 nm so that the milling time is as short as possible, up to 5 h [11]. Detailed data for this DOE example, which was performed thanks to the use of specialized software [11], are given in Figure 11.



Bootstrap Forest for Time to 200nm

Specifications

Target Column:	Time to 200nm	Training rows:	17
		Validation rows:	0
Number of trees in the forest:	10000	Test rows:	0
Number of terms sampled per split:	4	Number of terms:	17
		Bootstrap samples:	17
		Minimum Splits Per Tree:	10
		Minimum Size Split:	5

Overall Statistics

Individual Trees	RMSE
In Bag	0.5784008
Out of Bag	0.9647083

RSquare	RMSE	N
0.537	0.6532587	17

Per-Tree Summaries

Column Contributions

Term	Number of Splits	SS	Portion
Temp(Pot)	1984	1.31476675	0.3235
Temp(Out)	1615	1.05921406	0.2606
%Beads	1044	0.50190813	0.1235
%Pigment	750	0.29027034	0.0714
P(In)	738	0.27724826	0.0682
P(Out)	432	0.16135855	0.0397
Temp(CW)	325	0.08985563	0.0221
Temp PreMix	274	0.06716669	0.0165
P1	301	0.06482903	0.0160
Pig Bx	213	0.05220287	0.0128
Pump rpm	220	0.04248993	0.0105
P2	152	0.0274633	0.0068
P3	142	0.02611607	0.0064
Transfer Time	133	0.02539399	0.0062
Mill rpm	147	0.02394634	0.0059
Time PreMix	118	0.02121622	0.0052
P4	109	0.018694	0.0046

Figure 3 *Data from the 17 production runs [11]*

Within DOE, numerous calculations and graphs were provided, and the final results can be seen in Figure 4. As it can be seen in a part of diagram, the three confirmation runs showed milling times below the 5 h which was the assign at the first place [11]. In this way, DOE has proven to be very efficient in terms of minimizing the number of experiments and saving money.

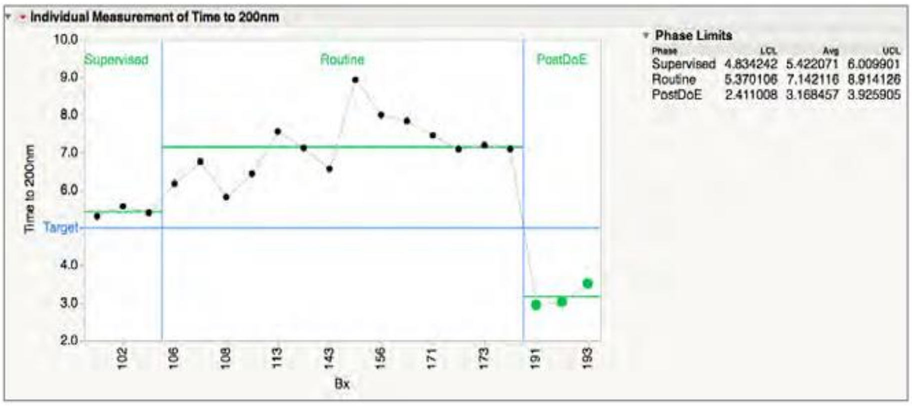
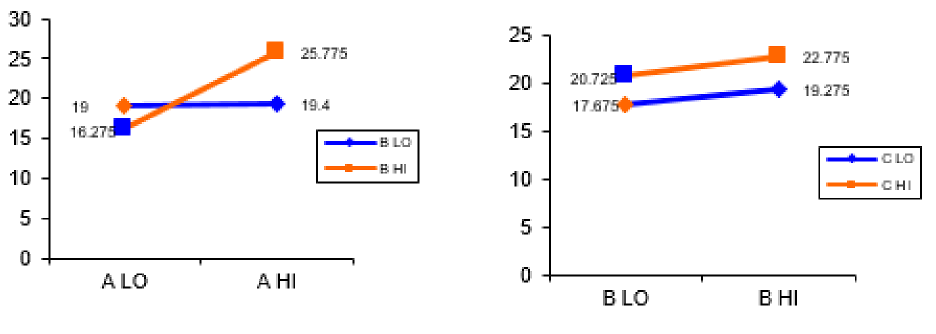


Figure 4 *Confirmation runs for DOE [11]*

Another example, different from the previous one, without details and calculations is given in Figure 5. The difference between these two examples is in complexity and availability. From this example, the three factors and their values can be seen with the high (1) and low (-1) value for each factor, and as it can be seen in Figure 5b and 5c, there is difference in results depending on which interaction was chosen for calculation [12].

	Run Order	Boost	Moist	Cycle	AxB	AxC	BxC	AxBxC	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Avg
1	6	45	6	11	1	1	1	-1	17.9	17.6				17.75
2	8	45	6	16	1	-1	-1	1	20.3	20.2				20.25
3	1	45	10	11	-1	1	-1	1	15	14.8				14.9
4	4	45	10	16	-1	-1	1	-1	18	17.3				17.65
5	2	90	6	11	-1	-1	1	1	17.5	17.7				17.6
6	5	90	6	16	-1	1	-1	-1	21.5	20.9				21.2
7	3	90	10	11	1	-1	-1	-1	24.2	23.1				23.65
8	7	90	10	16	1	1	1	1	27.6	28.2				27.9

a)



b) *c)*

Figure 5 *A DOE example: a) Factors & values; b) AxB interaction; c) BxC interaction [12]*

CONCLUSION

Experiment planning is an important aspect of any research. In that way it is important to know what we want to predict, which data are the input factors and which the outputs, what kind of relations exist between them. DOE is very powerful tool which can do more in less time.

Multifactorial experiments have a large number of factors for which sometimes there is no previous experience. With DOE, it is possible to minimize the number of experiments and to do modelling and prediction. So, DOE is useful not only for known, but also for unknown relations and factors. It helps to optimize the process and product quality as well.

DOE identifies factors which are significant and those which are not. With variation of factors and their interactions, it finds the best combination for requiring conditions.

REFERENCES

[1] Fisher R. A., The Design of Experiments, London, Oliver and Boyd, 1935.

[2] Antony J., A Strategic Methodology to the Use of Advanced Statistical Quality Improvement Techniques (PhD thesis), University of Portsmouth, UK, 1997.

[3] Telford J. K., Johns Hopkins APL Technical Digest, Vol. 27, 3 (2007) 224 – 232.

[4] Mager P. P., Medicinal Research Reviews, Vol.17, 5 (1997) 453 - 475

[4] Hibbert D. B., Journal of Chromatography B, Vol. 910, 1 December (2012) 2-13.

[5] Paulo F., Santos L., Materials Science and Engineering: C, Vol. 77, 1 August (2017) 1327-1340.

[6] Garud S. S., Karimi I. A., Kraft M., Computers & Chemical Engineering, Vol. 106, 2 November (2017) 71-95

[7] Schlueter A., Geyer P., Automation in Construction, Vol. 86, February (2018) 33-43.

[8] Puente-Massaguer E., Lecina M., Gòdia F. Applied Microbiology and Biotechnology, 104 (2020) 1569–1582

[9] Yu P., Low M. Y., Zhou W., Trends in Food Science & Technology, 71 (2018) 202–215.

[10] Durakovic B., Periodicals of Engineering and Natural Sciences, 5, 3 (2018) 421 – 439.

[11] https://www.jmp.com

[12] https://asq.org