sciendo

PRODUCTION ENGINEERING ARCHIVES 20 (2018) 43-48

PRODUCTION ENGINEERING ARCHIVES

ISSN 2353-5156 (print)
ISSN2353-7779 (online)

Exist since 4th quarter 2013
Available online at www.qpij.pl/production-engineering-archives

The outline of the expert system for the design of experiment

Renata Dwornicka[1], Jacek Pietraszek[1]

[1] Cracow University of Technology, Faculty of Mechanical Engineering, Al. Jana Pawła II 37, 31-864 Kraków, Poland
e-mail: renata.dwornicka@mech.pk.edu.pl

Article history
Received 03.07.2018
Accepted 20.09.2018
Available online 30.09.2018

Keywords
expert system
design of experiment
factorials
Taguchi robust design
RSM

Abstract
The design of experiment (DoE) is a methodology originated from early 1920s when Fisher's papers created the analysis of variance and first known experimental designs: latin squares. It is focused on a construction of empirical models based on measurements obtained from specifically structured and driven experiments. Its development resulted in the constitution of four distinctive branches recognized by the industry: factorials (full or fractional), Taguchi's robust design, Shainin's Red-X® and a response surface methodology (RSM). On one hand, the well-known success stories of this methodology implementations promise great benefits, while on other hand, the mathematical complexity of mathematical and statistical assumptions very often lead to improper use and wrong inferences. The possible solution to avoid such mistakes is the expert system supporting the design of experiments and subsequently the analysis of obtained data. The authors propose the outline of such system and provides the general analysis of the ontology and related inference rules.

DOI: 10.30657/pea.2018.20.09 JEL: L23, M11

1. Introduction

The methodology known as 'design of experiment' is very efficient tool for discovering new knowledge about technological processes as well as fine tuning them. Unfortunately, this tool set contains many methods which achieve their maximum efficiency for different problems and at different assumptions. In recent decades, software tools (e.g. Statistica, Statgraphics, Minitab, R) have freed engineers and researchers from tedious calculation work. However, it is still the user who is responsible for selection of the right experimental design and appropriate data analysis methods, especially so sophisticated as logistic regression models (Hosmer and Lemeshow, 2000; Hilbe, 2009), a categorical data analysis (Agresti, 2002), a principal component analysis PCA (Jolliffe, 2010), a cluster analysis (Everitt et al., 2011) or – last but not least – a multivariate statistical techniques at all (Izenman, 2008).

The rapid development in a specific field of an artificial intelligence – expert systems (Liebowitz, 1998; Jackson, 1999; Siler and Buckley, 2005) – gives a hope for further relief of users. The advisory expert system appears to be a proper tool to support users in these very difficult and risky decisions. In this paper, authors try to sketch outline of the rules set for such a system.

Another promising area – linguistic summarization (Niewiadomski, 2008) – appears to be desirable tool for a valid results interpretation and automatic creation of condensed, transparent, but not necessarily precise, interpretations of large result datasets, however such problem is beyond a scope of this paper.

In the next chapters, authors make a review of the most popular methods existing under the franchise of 'design of experiment' banner.

2. Review and analysis of popular methods

The authors chose several methods that were considered the most popular among DOE techniques: Latin squares, factorials, Taguchi robust design, Shainin's Red-X® and response surface methodology.

2.1. Latin squares - beginnings

The design of experiments foundations were created in 1918, when R. A. Fisher published his paper (Fisher, 1918) where the analysis of variance (ANOVA) (Gentle and Hardle, 2012) was mentioned first time as a method to split a combined effect into separated impacts of particular factors. Some years later, in 1925, Fisher found an answer (Fisher, 1925) to the question being a specific "inverse-problem": how to set a scheme of the experiment to make ANOVA maximally effective. The founded scheme is known as a "Latin square" (LS) because of its special construction diagram similar to a "magic square" filled with a Latin letters (Fig.1). This diagram should be interpreted as:

- the column of the diagram is a first factor and the number of the column is its level,
- the row of the diagram is a second factor and the number of the row is its level,
- the Latin letter set at a cross of the column and the row is a level of the third factor.

For the example (Fig.1): the third column, the second row and the letter 'B' – it means that the first factor is set at third level, the second factor – at second level and third factor – at second level (i.e. at letter 'B'). The Latin square constructed for *n* levels is 1/nth of the full experimental design i.e. it is a small fraction from the set of all possible combinations of factors.

A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A

Fig. 1. The sample latin square for three factors with four levels each

However Latin squares allow to serve any arbitrary number of levels, their impose some limitations:

- only three factors may be analyzed,
- the number of levels must be the same for all factors,
- no interactions between factors are allowed.

Later, in 1935 Fisher published a book (Fisher, 1935) where the term 'design of experiment' was explicitly used even in the title.

2.2. Factorials

The concept of Latin squares were generalized into 'general full factorial' (GFF) experimental design. The GFF is a combination of all levels of considered factor with all levels of all other factors. The GFF has very good statistical properties and a capacity to identify all linear effects and all possible interactions up to the highest possible order however it is the most expensive variant of an experimental research.

In 1935, F. Yates found a way out this impasse. He proposed (Yates, 1935) the recipe how to greatly reduce the number of the required experiments. He reversed the statement "if GFF is used then all effects and interactions are identifiable" into "if higher order interactions are removed from the model then only some combinations of factor treatments should be used". Such subset of GFF is known as the fractional factorial where the term 'fractional' is related to the fragmentary in contrary to the whole experimental design.

He might to make such radical decision about removing of the higher order interactions at a relatively low risk because they are very rarely observed in the real industrial processes, especially in machining. The only exceptions are chemical and termomechanical processes.

Additionally, Yates described the effective algorithm (Montgomery, 2008) to construct the fractional factorial in the specific case where all factors have only two level settings. It is very often met, especially when sensitivity of the technological process is investigated at preliminary research.

The first step is to determine the biggest two-level full factorial which size is still lower than assumed limit determined by the economy or deadline limitations. The founded experimental design is assumed as a core while interactions of sup-ported factors are used to generated the rest of the factors.

The following example describe this procedure for the searching of the smallest fractional factorial supporting only linear effects:

- the process with four factors (A, B, C, D) should be investigated,
- the economy limitations does not allow to make more than 10 experimental tests,
- two-level fractional factorial should be proposed.

The test number limitations (i.e. 10) imposes that the biggest two-level full factorial in this limit is 23 i.e. full factorial (Fig.2a) for three factors (acronyms according to well-known Yates's notation).

a)

A	B	C
-	-	-
+	-	-
-	+	-
+	+	-
-	-	+
+	-	+
-	+	+
+	-	+

b)

A	B	C	D
-	-	-	-
+	-	-	+
-	+	-	+
+	+	-	-
-	-	+	+
+	-	+	-
-	+	+	-
+	-	+	-

Fig. 2. Sample two-level full factorial for three factors (a) as a core for the two-level fractional factorial (b) with generator D=ABC

Three factors allows four interactions: three second-order i.e. AB, AC, BC and one third-order i.e. ABC. In the particular sample (Fig.2b) the interactions ABC was used as a generator for the fourth factor D.

Such experimental design may be used to determine the sensitivity of the process onto the investigated factors. The typical graphical tool used to present this sensitivity is well-known Pareto chart.

In 1946, Plackett and Burman showed (Plackett and Burman, 1946) a special kind of two-level experimental designs based on Hadamard matrices with linear complexity instead of a power complexity in Yates's designs.

2.3. Taguchi's robust design

Taguchi robust design (Phadke, 1989) is not only the specific variant of DoE, but also a general idea of a process or product design insensitive to disturbances. His main concept is to introduce two separate experimental design: one, controlling significant factors of the process or product (known as the internal array) and the second, (known as the external array), controlling most significant disturbances or weakly controlled factors (known as noise factors). The aim of the procedure is to find settings of the controlled factors which simultaneously optimize (maximize, minimize or stabilize) the process outcome and maximize a resistance of the process to the influence of noise factors.

The internal array may be a special orthogonal experimental design (prepared by Taguchi) with different number of levels for each of factors, or typical factorial – full or fractional – experimental design.

The external array is usually highly fractionalized two-level fractional factorial, where levels are related to extreme settings of noise factors e.g. the lowest or highest temperature, humidity, vibrations etc.

Both arrays are crossed giving the complete Taguchi robust design (Fig. 3).

			E	-	-	+	+
			D	-	+	-	+
A	B	C					
-	-	+					
+	-	-					
-	+	-					
+	+	+					

Fig. 3. Combinations of two crossed experimental designs in Taguchi robust design: the internal array (A,B,C) and the external array (D,E). The green box is a sample test with controlled factors (A,B,C) set to (+, –, –) and noise factors (D, E) set to (–, +)

Testing of disturbances may be conducted simultaneously with factors or separately. The first case appears when noise factors imposes production process directly e.g. raw materials instability, process drift etc. The second case is conducted when noise factors are shifted over time in relation to the main process e.g. test in a climatic chamber for impact of operating conditions.

Each of the internal array treatment should be replicated for all of the external array treatments. It means that number of the required tests is a multiplication of the internal array treatments and the external array treatments. To avoid relatively high level of cost, the external array is typically selected as highly fractionalized two-level factorial while the internal array is selected as fractional factorial (with more than two levels) or a specific design prepared by Taguchi and known as Taguchi orthogonal arrays (Phadke, 1989).

The analyzed output is not the measurements directly, but their specific transformation SNR (Signal-to-Noise Ratio) made with one of three available functions:

- *lower-the-better* (Eq.1) – where the sum is evaluated for the particular internal treatment i over all $(1\ldots k)$ external treatments j

$$SNR_i = -10\log\left[\frac{1}{n}\sum_{j=1}^{k} y_{i,j}^2\right] \tag{1}$$

- *greater-the-better* (Eq.2) – where the sum is evaluated for the particular internal treatment i over all $(1\ldots k)$ external treatments j

$$SNR_i = -10\log\left[\frac{1}{n}\sum_{j=1}^{k}\frac{1}{y_{i,j}^2}\right] \tag{2}$$

- *nominal-the-best* – where two variants are considered depending on the relation between the main outcome and its variance over disturbance (Eq.3): (a) linear relation (Eq.4) and (b) lack of the relation (Eq.5)

$$s_i^2 = \frac{\sum_{j=1}^{k}\left(y_{i,j} - \bar{y}_i\right)^2}{k-1} \tag{3}$$

$$SNR_i = 10\log\left(\frac{\bar{y}_i^2}{s_i^2}\right) \tag{4}$$

$$SNR_i = -10\log\left(s_i^2\right) \tag{5}$$

Their names describes the goal selected for the original output while the trans-formed SNR is always maximized. One SNR value is calculated for each of the internal array treatment over all of the related external arrays treatments (Phadke, 1989; Montgomery, 1997).

Instead of the classic factorials, the robust design optimizes not the sole out-come at random disturbances, but the outcome at the average existence of the extremely changing disturbances. It may lead to the results that are slightly suboptimal than obtained from factorials experiments, but are significantly more insensitive to the disturbances.

The limitations of the Taguchi's robust design is lack of the any interactions detection. The typical graphical tool used to select optimal settings is the margin-al means plot (Montgomery, 1997).

2.4. Shainin's Red-X®

Shainin's Red-X® methodology (Bhote, 1991) is not a single algorithm but a set of seven recipes what make it very similar to Six-Sigma package. It goal is not to optimize a process outcome itself but to minimize its variability i.e. to make process window narrower. Its basic foundation is the assumption that small number of factors (up to maximum three) are responsible for the main part of disturbance observed in the process.

The approach was originally introduced in Grumman Aircraft Engineering Corporation during construction of Lunar Expedition Module in Apollo mission, where Dorian Shainin was responsible for the production process quality. Later, Shainin moved to General Motors and implemented his procedure in the automotive industry.

The first step in Shainin's Red-X® is to set a measurement system i.e. to gain assurance that measurement procedure has a capacity to really obtain reliable data and not artifacts. In modern advanced factories, it is usually provided by a measurement system analysis (MSA) but it still very often is not met. In such situation, Shainin's system provides a MSA substitute i.e. ISO Plot® where 30 product samples are two times measured and obtained data are paired. The measurement are presented on the 2D plot and the horizontal dispersion is a process dispersions while the dispersion around slope (line 45 degree) is a measurement dispersion and reflects measurement process repeatability.

The optional second step is Multi-Vari® analysis performed with a specific control charts used to detect a location of the process disturbance: (a) raw material or machined part, (b) production site or tool and (c) production site environment.

Next step depends on the character of the analyzed object/process. If the investigated object is a product allowing reversible dismantling i.e. dismantling and reassembly, the ComponentSearch® algorithm is used. If the investigated object does not allow reversible dismantling i.e. the dismantling and the inspection is destructive, the PairedComparison® algorithm is used. At last, if a process (instead of a material object) is investigated then VariableSearch® (structurally very similar to ComponentSearch®) is used.

The result of this step is a set of the main sources of variability: Red-X® (the main source), Pink-X® (the second order source) and Pale Pink-X® (the third order source).

In the next step, FullFactorial®, the possible interactions between these main source are detected and evaluated. If interactions are not significant, the optimum settings are detected for each of factors independently. If any interaction is more significant than a linear effect (single factor), it means that an optimum pair or an optimum triplet should be determined at the same time i.e. their settings depend mutually on each other.

Finally, the B vs. C® algorithm, based on Tukey's range significance test (Siegel and Tukey, 1960), checks if the new optimum settings changes product/process so enough to be detected.

Optional step, based on dispersion plots of the measured process outcome versus Red-X® and Pink-X®, allows to determine reasonable and argued specification of raw materials and source parts.

The specific property of Shainin's Red-X® approach is to almost completely avoid of any explicit use of the statistics. The base tools for this approach are a millimeter paper, a setsquare, a pencil and a four-function calculator. The necessary statistics are hidden beyond "magic" number located explicitly in the some simple formulas. Due to its simplicity and a low cost, Shainin's approach is very popular in the industry.

2.5. Response Surface Methodology

The response surface methodology (RSM) was developed in 1951 by Box and Wilson (Box and Wilson, 1951). In contrary to the factorial approach, it allows to use continuous factors settings i.e. settings described by a numbers, not a labels. Instead of a fixed-effects model (like Latin square, Taguchi or factorials), it introduced a classic mathematical approximation formulas based usually on a second order polynomials. In 1958, Scheffé generalized this approach to a specific situation of mixtures where a sum of factors setting must to be constant (Scheffé, 1958).

Typically, a RSM model is identified by a least square method providing a maximum likelihood estimation if the noise term is Gaussian (Montgomery, 2008). This condition may be a priori checked very rarely and this assumption is typically tested a posteriori by a normality test of residuals (John, 1998).

Typical exploration tools are: the analysis of effects, the analysis of variance, the inspection of the pure error, 2D and 3D plots.

The RSM is very popular however authors practically observed many of improper use cases and inference mistakes. The typical error is ignoring a specific statistical assumptions related to a particular experimental design. It very often lead to an erroneous conclusions.

3. Selection rules

The scope of the investigation is theoretically a results of the assumed goal, but in the industrial practice, the most important limitations are economical i.e. available funds and imposed deadlines. They most strongly limit tools and methods which may be used.

The question addressed to the expert system may be formulated in two forms:

1. what are necessary resources (funds, machines, workers, time) for the defined scope of the investigation,
2. what is maximum available scope of the investigation for the defined limits of the resources (funds, machines, workers, time).

In the academy, the first approach is often met when a future research grant is considered. In the industrial practice, especially in production engineering, the second approach is the only allowed one.

3.1. Resource limitations

In the beginning, two basic properties of the single experimental test have to be considered: a cost C_{test} and a duration T_{test}.

In the industrial approach, the budget B_{limit} and deadline T_{limit} are strictly de-fined. The maximum number of tests limited by a budget is defined by an equation:

$$N_{budget} = \frac{B_{limit}}{C_{test}} \tag{6}$$

while the maximum number of tests possible to made on a single experimental unit is defined by a formula:

$$N_{time} = \frac{T_{limit}}{T_{test}} \tag{7}$$

If the N_{budget} is larger than N_{time}, it means that experimental tests may be distributed on many experimental units simultaneously however it requires to introduce a special blocking factor to take into account a systematic error i.e. drift or constant differences between experimental units.

3.2. Selection of the investigation aim

The observed outcome Y should be defined. The DoE investigation should select one of the four possible goals related to this outcome:
1. the identification of the most important factors (screening research),
2. the process stabilization i.e. narrowing of the process window,
3. the process optimization i.e. searching of the process optimum settings,
4. the process mapping i.e. construction of the process prediction model.

The possible rules are:
1. if the goal is screening then possible approaches are factorial (fractional, Plackett-Burman) or Shainin's,
2. if the goal is stabilization then possible approaches are Shainin's, fractional factorial or RSM,
3. if the goal is optimization then possible approaches are factorial, Taguchi or RSM,
4. if the goal is mapping then possible approaches are factorial or RSM.

In the industrial practice, the stabilization should be achieved first even before optimization.

3.3. Selection of investigated factors

It leads to selection of the factors related to the observed outcome. The settings nature of the factors determined the possible methodology:
1. if the factors are continuously set and the constant sum must be met, they should be investigated by mixture otherwise
2. if the factors are continuously set, they should be investigated by RSM, other-wise
3. it should be investigated by factorials, Taguchi's or Shainin's approach.

A special trick, an arbitrary discretization, may be used to transform factors from continuous into not-continuous set.

3.4. Possibility of interactions and the model selection

The *n*-order interactions is a mathematical term for a phenomenon where observed impact of the particular factor depends on (n–1) other factors settings. Some approaches *a priori* reject the existence of interactions. It means that such approaches should not be used for the investigation of the processes where investigations may be met:
1. if the interactions may be met, then do not use Taguchi, Plackett-Burman, linear factorials or linear RSM otherwise
2. use factorials with non-linear models, Shainin's or RSM with non-linear models.

3.5 Outcome transformations

The observed and measured outcome may related to one of three numbering scales defined by Stevens (Stevens, 1946):
- interval scale, where the value is not limited i.e. may be negative or positive,
- ratio scale, where the value is one-sided bound, typically it means that the value is positive or non-negative,
- absolute scale, where the value is two-sided bound.

In the most popular least squares method, the random noise term describing uncontrolled disturbances is assumed to be Gaussian i.e. normally distributed. It requires unlimited (at least theoretically) outcome, both in negative and positive side. If the outcome is one-side or two-side bound, it requires preliminary transformation:
1. if the outcome is one-side bound, then all values should be shifted into positives and next processed by a special transformation into the whole real numbers space, otherwise
2. if the outcome is two-sided bound, then all values should be shifted and scaled into interval [0, 1] and next processed by a special transformation into the whole real numbers space otherwise
3. the outcome should be left intact.

The typical transformation of the outcome related to the ratio scale is logarithmic transformation. It transforms zero bound into negative infinity, one into zero and larger values into positive values up to positive infinity.

The typical transformations of the outcome related to the absolute scale are logistic (Hosmer and Lemeshow, 2000) or probit transformations (Hosmer and Lemeshow, 2000).

4 Conclusions

The rationale for the construction of the expert system for the design of the experiment was presented. The history roots of DoE and its developing into four main branches was showed. Five groups of selection rules were defined: the resource limitations, the selection of the investigation aim, the selection of the investigated factors, the model selection according to existence of interactions and the outcome pre-processing transformations.

Further efforts will be focused on the testing of the working model in the industrial environment and building rules related to the analysis methods and results interpretations.

Reference

Agresti, A., 2002. *Categorical data analysis*. John Wiley & Sons, Hoboken.

Bhote, K., 1991. *World Class Quality*. AMACOM, New York.

Box, G.E.P., Wilson, K.B., 1951. *On the Experimental Attainment of Optimum Conditions*. Journal of the Royal Statistical Society B 13, 1-45.

Everitt, B.S., Landau, S., Leese, M., Stahl, D., 2011. *Cluster analysis 5th edition*. John Wiley & Sons, Hoboken.

Fisher, R., 1918. *Studies in Crop Variation. I. An examination of the yield of dressed grain from Broadbalk*. Journal of Agricultural Science 11, 107-135.

Fisher, R.A., 1925. *Statistical Methods for Research Workers*. Oliver & Boyd Press, Edinburgh.

Fisher, R.A., 1935. *The Design of Experiments*. Oliver & Boyd Press, Edinburgh.

Gentle, J.E., Hardle, W.K., 2012. *Handbook of Computational Statistics*. Springer-Verlag, Berlin-Heidelberg.

Hilbe, J.M., 2009. *Logistic regression models*. CRC Press, Boca Raton.

Hosmer, D.W., Lemeshow, S., 2000. *Applied Logistic Regression*. John Wiley & Sons, Hoboken.

Izenman, A.J., 2008. *Modern multivariate statistical techniques: regression, classification, and manifold learning*. Springer, New York.

Jackson, P., 1999. *Introduction to expert systems*. Addison-Wesley, Harlow, England.

John, P.W.M., 1998. *Statistical design and analysis of experiments*. SIAM, Philadelphia.

Jolliffe, I.T., 2010. *Principal component analysis*. Springer, New York.

Liebowitz, J. (ed.), 1998. *The Handbook of applied expert systems*. CRC Press, Boca Raton.

Montgomery, D.C., 1997. *Introduction to Statistical Quality Control*. John Wiley & Sons, Hoboken.

Montgomery, D.C., 2008. *Design and Analysis of Experiments*. John Wiley & Sons, Hoboken.

Niewiadomski, A., 2008. *Methods for the linguistic summarization of data: applications of fuzzy sets and their extensions*. AOW Exit, Warszawa, Poland.

Phadke, M.S., 1989. *Quality Engineering Using Robust Design*. Prentice Hall International, Inc., London.

Plackett, R.L., Burman, J.P., 1946. *The design of optimum multifactorial experiments*. Biometrika 33, 305-325.

Scheffé, H., 1958. *Experiments with Mixtures*. Journal of the Royal Statistical Society B 20, 344-360.

Siegel, S., Tukey, J.W., 1960. *A nonparametric sum of ranks procedure for relative spread in unpaired samples*. Journal of the American Statistical Association 55 (291), 429-445.

Siler, W., Buckley, J.J., 2005. *Fuzzy expert systems and fuzzy reasoning*. John Wiley & Sons, Hoboken.

Stevens, S.S., 1946. *On the theory of scales of measurements*. Science 103, 677-680.

Yates, F., 1935. *Complex experiments*. Journal of the Royal Statistical Society Suppl. 2, 181-247.

实验设计专家系统概述

關鍵詞
专业系统
实验设计
阶乘
Taguchi坚固的设计
RSM

摘要
实验设计（DoE）是一种起源于20世纪20年代早期的方法，当时Fisher的论文创建了方差分析和第一个已知的实验设计：拉丁方块。它侧重于基于从特定结构和驱动实验获得的测量结果的经验模型的构建。它的发展导致了该行业认可的四个独特分支的构成：阶乘（全部或分数），田口的稳健设计，Shainin的RedX®和响应表面方法（RSM）。一方面，这种方法实现的众所周知的成功案例有很大的好处，而另一方面，数学和统计假设的数学复杂性经常导致不正确的使用和错误的推论。避免此类错误的可能解决方案是支持实验设计的专家系统，并随后对获得的数据进行分析。作者提出了这种系统的概述，并提供了对本体和相关推理规则的一般分析。