Design of Experiments Application, Concepts, Examples: State of the Art

Benjamin Durakovic
Industrial Engineering, International University of Sarajevo

Article Info	ABSTRACT
	Design of Experiments (DOE) is statistical tool deployed in various types of system, process and product design, development and optimization. It is multipurpose tool that can be used in various situations such as design for comparisons, variable screening, transfer function identification, optimization and robust design. This paper explores historical aspects of DOE and provides state of the art of its application, guides researchers how to conceptualize, plan and conduct experiments, and how to analyze and interpret data including examples. Also, this paper reveals that in past 20 years application of DOE have been grooving rapidly in manufacturing as well as non-manufacturing industries. It was most popular tool in scientific areas of medicine, engineering, biochemistry, physics, computer science and counts about 50% of its applications compared to all other scientific areas.

Corresponding Author:

Benjamin Durakovic
International University of Sarajevo
Hrasnicka cesta 15
7100 Sarajevo, Bosnia
Email: bdurakovic@ius.edu.ba

1. Introduction

Design of Experiments (DOE) mathematical methodology used for planning and conducting experiments as well as analyzing and interpreting data obtained from the experiments. It is a branch of applied statistics that is used for conducting scientific studies of a system, process or product in which input variables (Xs) were manipulated to investigate its effects on measured response variable (Y).

Over past two decades, DOE was a very useful tool traditionally used for improvement of product quality and reliability [1]. The usage of DOE has been expanded across many industries as part of decision-making process either along a new product development, manufacturing process and improvement. It is not used only in engineering areas it has been used in administration, marketing, hospitals, pharmaceutical [2], food industry [3], energy and architecture [4] [5], and chromatography [6]. DOE is applicable to physical processes as well as computer simulation models [7].

1.1. Historical perspective

One Factor At a Time (OFAT) was very popular scientific method dominated until early nineteen century. In this method one variable/factor is tested at a time while the other variables are constrained except the investigated one. Testing multiple variables at a time is better especially in cases where data must be analyzed

421

carefully. In the 1920s and 1930s **Ronald A. Fisher** conducted a research in agriculture with the aim of increasing yield of crop in the UK. Getting data and was challenging e.g. if he relayed on his traditional method ANOVA (*F*-test, means *Fisher* - test) he may plant a crop in spring and get results in fall which is too long for getting data. Finally, he came up with design of experiment and officially he was the first one who started using DOE. In 1935, he wrote a book on DOE, in which he explained how valid conclusion could be drawn from the experiment in presence of nuisance factors. He analyzed presence of nuisance factors with fluctuation of weather conditions (temperature, rainfall, soil condition). Credit for Response Surface Method (RSM) belongs to **George Box** who is also from the UK. He was concerned with experimental design procedures for process optimization. In 1550s, **W. Edwards Deming** was concerned with design of experiment as well as statistical methods. **Genichi Taguchi** was Japanese statistician concerned with quality improvement methods. He contributed to statistic by introducing Loss function and experiments extending with an "outer array" in DOE as an advanced method in the Six Sigma initiatives [8].

1.2. The main uses of DOE

Design of experiment is multipurpose tool that can be used in various situations for identification of important input factors (input variable) and how they are related to the outputs (response variable). Therefore, DOE mainly uses "hard tools" as it was reported in [9]. In addition, DOE is bassically regression analysis that can be used in various situations. Commonly used design types are the following [10]:

1. **Comparison** – this is one factor among multiple comparisons to select the best option that uses t–test, Z–test, or F–test.
2. **Variable screening** – these are usually two-level factorial designs intended to select important factors (variables) among many that affect performances of a system, process, or product.
3. **Transfer function identification** – if important input variables are identified, the relationship between the input variables and output variable can be used for further performance exploration of the system, process or product via transfer function.
4. **System Optimization** – the transfer function can be used for optimization by moving the experiment to optimum setting of the variables. On this way performances of the system, process or product can be improved.
5. **Robust design** – deals with reduction of variation in the system, process or product without elimination of its causes. Robust design was pioneered by Dr. Genichi Taguchi who made system robust against noise (environmental and uncontrollable factors are considered as noise). Generally, factors that cause product variation can be categorized in three main groups:
   - external/environmental (such as temperature, humidity and dust);
   - internal (wear of a machine and aging of materials);
   - Unit to unit variation (variations in material, processes and equipment).

1.3. DOE application in research and procedure

Even though DOE tool are not new techniques but its application has expended rapidly over the scientific areas including product/process quality improvement [11], product optimization [12] and services in the past two decades. Trainings and recent user-friendly commercial and non-commercial statistical software packages contributed significantly to DOE expansion in the research in this period. DOE application over the globe and across various scientific areas in the period from 1920 to 2018 is shown in Figure 1.
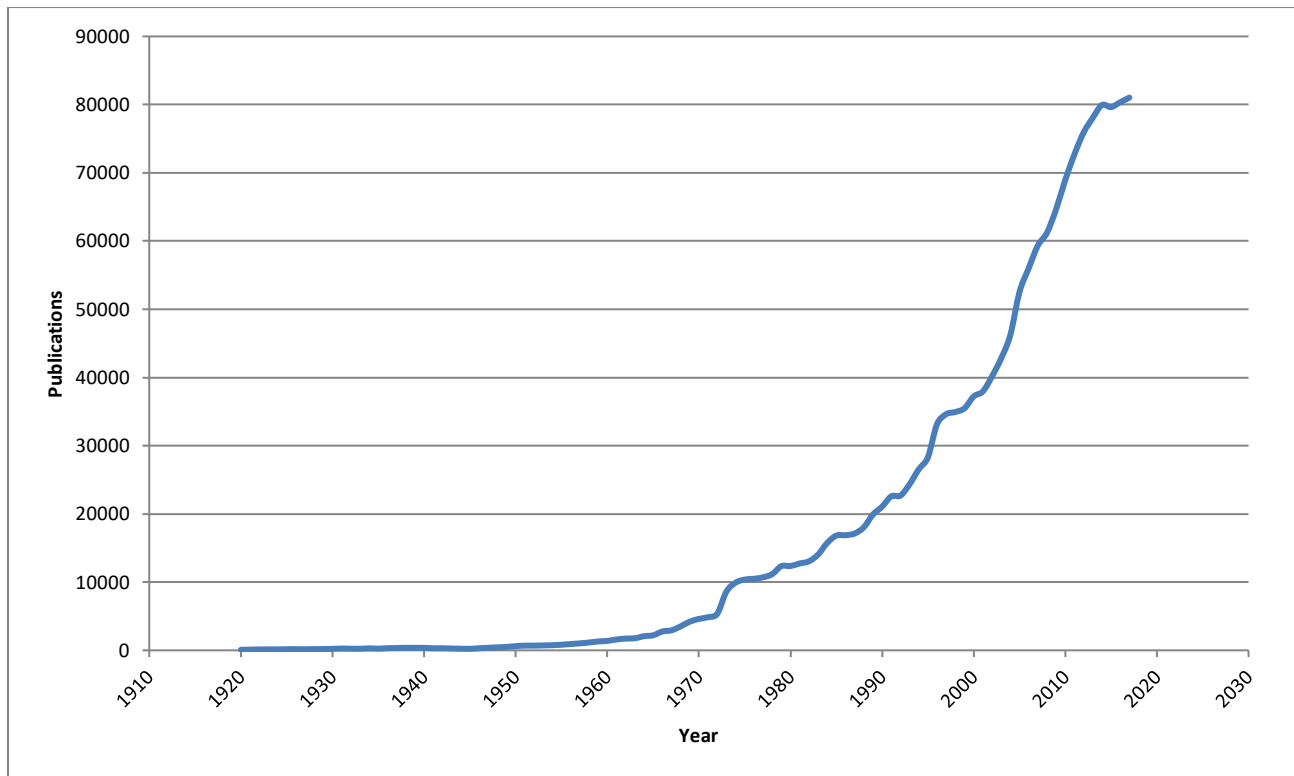
Figure 1. DOE application in scientific research [13] [1]

Application of DOE in research started in 1920s with Fisher's research in agriculture. Over four upcoming decades later its application in the research was negligible. Significant use of DOE in the research project was noticed in the late 1960s and 1970s. Thus, it took about 50 years for the DOE to achieve significant application in the research. Since in this period there were no software packages that would foster its application DOE had not signified a strong expansion. Thanks to edication and software development in 1990s and later, the use of DOE in research over various scientific areas has risen sharply. A linear model that represents a rapid increase in the use of DOE in the research projects is shown in Figure 2 and represented by a mathematical linear model.
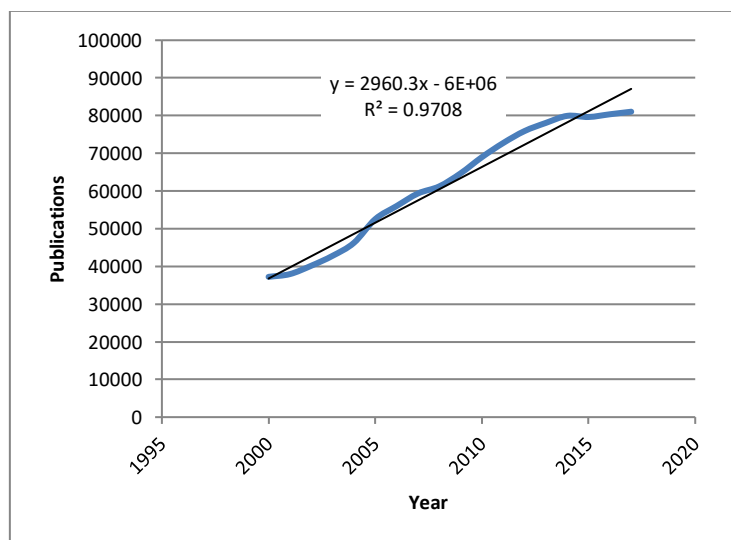

Figure 2. Progressive use of DOE as scientific method over past two decades [13][2]

---

[1] Data obtained from Scopus for search "design of experiments" OR "experimental design" OR "DOE" in Title – abstact – Key words
[2] Data obtained from Scopus for search "design of experiments" OR "experimental design" OR "DOE" in Title – abstact – Key words

DOE application in the future can be predicted with linear regression model, which is based on past date over 18 years. Therefore, it can be expected that DOE usage expansion will continue in the future including its application in the existing and new scientific areas. The state of the art of DOE application in certain scientific areas is shown in Figure 3.
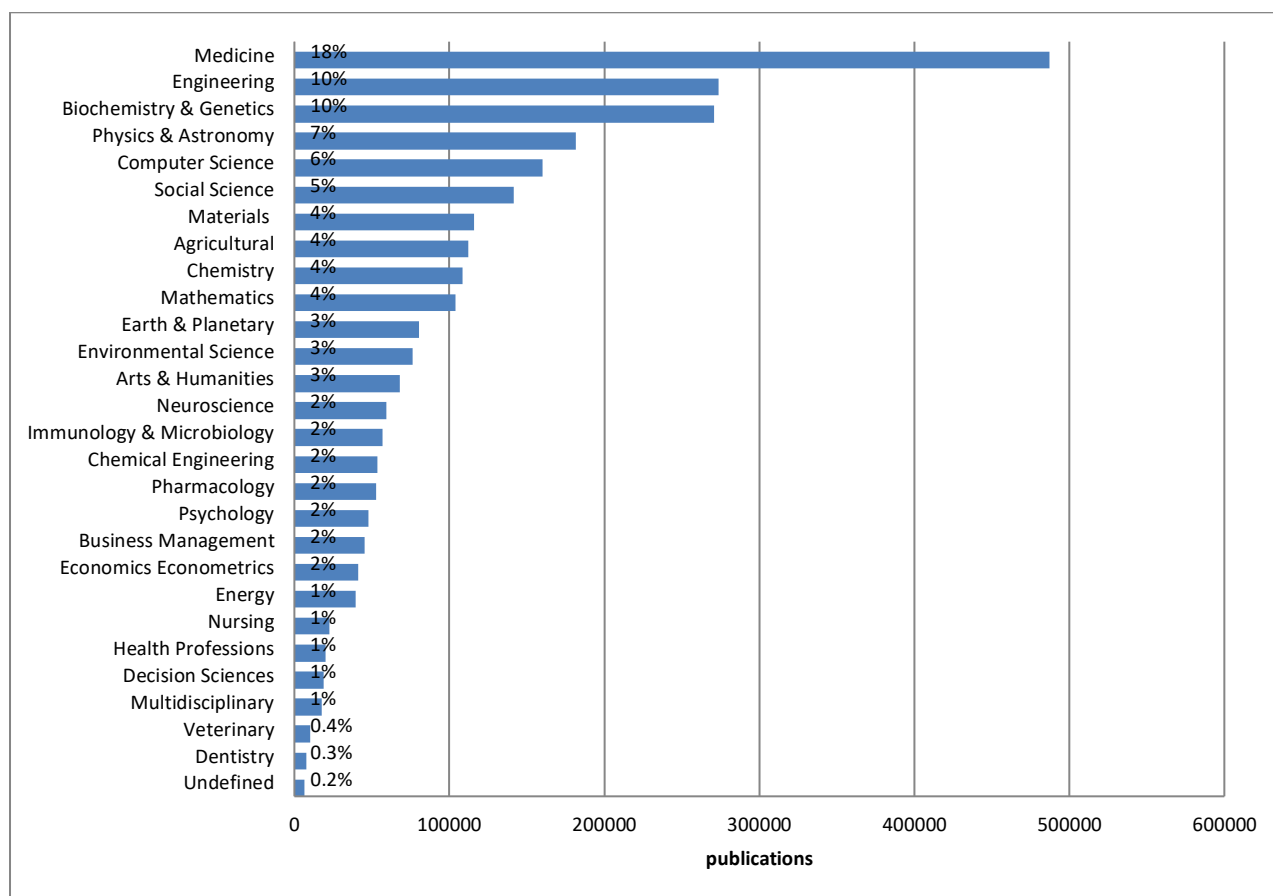
Figure 3. DOE application per scientific area [13][3]

DOE as scientific method was most popular in scientific areas of medicine, engineering, biochemistry, physics and computer science. Its application in these areas counts about 50% compared to all other scientific areas. Only medicine participates about 18%, while engineering and biochemistry together participate with 20%; and physics and computer together science participate with 13%.

General practical steps and guidelines for planning and conducting DOE are listed below:

1. **State the objectives** – it is a list of problems that are going to be investigated.
2. **Response variable definition** – this is measurable outcome of the experiment that is based on defined objectives.
3. **Determine factors and levels** – selection of independent variable (factors) that case change in the response variable. To identify factors that may affect the response variable fishbone diagram might be used.
4. D**etermine Experimental design type** – e. g. a screening design is needed for significant factors identification; or for optimization factor-response function is going to be planned, number of test samples determination.
5. **Perform experiment** using design matrix.

---

[3] Data obtained from Scopus for search "design of experiments" OR "experimental design" OR "DOE"   in Title – abstact – Key words

6. **Data analysis** using statistical methods such as regression and ANOVA.
7. **Practical conclusions and recommendations** – including graphical representation of the results and validation of the results.

Most difficult part of DOE is to plan experiment in term of selecting appropriate factors for testing (x–variables), what ranges of x's to select, how many replicates is supposed to be used, and is center point required?

1.4. DOE Software

DOE can be quickly designed and analyzed with the help of suitable statistical software. For this purpose, there are some commercial and freeware statistical packages. The well-known commercial packages include: Minitab, Statistica, SPSS, SAS, Design-Expert, Statgraphics, Prisma, etc. The well-known freeware packages are: R and Action for Microsoft Excel [14].

The most popular commercial packages Minitab and Statistica are equipped with user friendly interface and very good graphics output. Freeware package Action has suitable graphics output and utilizes R platform together with Excel

Also, DOE design and analysis can be done easily in Microsoft Excel, using the procedure and formulas described in the following paragraphs. To perform any DOE as it was described earlier, understanding of analysis of variance (ANOVA) and linear regression as statistical methods is required. Therefore, more details about these two statistical methods are provided below.

2.   Statistical Background

2.1. Analysis of variance (ANOVA)

In cases that there are more than two test samples ANOVA is used to determine whether there are statistically significant differences between the means the samples (treatments). In cases that experiment contains two samples only, then *t*-test is good enough to check whether there are statistically significant differences between the means of treatments. In this case it is tested hypothesis assuming that a least one mean treatment value (μ) differs from the others. Therefore, null and alternative hypotheses can be express as [15]:

$$H_0: \mu_1 = \mu_1 = ... = \mu_k = 0$$

$$H_1: \mu_j \neq 0 \text{ for at least one } j \text{ different than zero.}$$

(1)

The procedure of test involves an *analysis of variance (ANOVA)* and performing *F*-test. Observed value is calculated as the ratio between treatment mean squares (MS_{Tr}) and error mean squares MS_E (error variance):

$$F_o = \frac{MS_{Tr}}{MS_E} = \frac{\frac{SS_{Tr}}{a-1}}{\frac{SS_E}{a(n-1)}}$$

(2)

where, SS_{Tr} is sum of squares of treatment, SS_E is sum of squares of error, $(a-1)$ represents treatments degrees of freedom, $a(n-1)$ represents error degrees of freedom, a is number of treatments (number of samples), n is number of observation for particular treatment. Total sum of squares (SS_T) is addition of sum of squares of treatment and sum of squares of error and it is calculated as:

$$SS_T = SS_{Tr} + SS_E; \quad or \quad \sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij} - \bar{\bar{y}}\right)^2 = n\sum_{i=1}^{a}\left(\bar{y}_i - \bar{\bar{y}}\right)^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_i\right)^2$$

(3)

whre, y_{ij} is j-th observation taken in treatment i, $\bar{\bar{y}}$ is an overall average (grand mean) of all observation for each material group i, $\bar{y}_i$ is an average value of observation in treatment i, a is number of treatments (groups), n is number of observations per each treatment. All of them are shown in Table 1 as well.

Table 1. General data matrix for a single-factor ANOVA

Material group	Observations				Averages
1	y_{11}	y_{12}	$\cdots$	y_{1n}	$\bar{y}_1$
2	y_{21}	y_{22}	$\cdots$	y_2	$\bar{y}_2$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	$\cdots$	y_{an}	$\bar{y}_a$
					$\bar{\bar{y}} = \dfrac{\sum_{i=1}^{a} \bar{y}_i}{a}$

Having calculated F_o, H_0 can be accepted or rejected in the following cases:

$$H_0 \text{is} \begin{cases} \text{Rejected if } F_o > F_{\alpha,(a-1),a(n-1)} \\ \text{Accepted if } F_o < F_{\alpha,(a-1),a(n-1)} \end{cases} \tag{4}$$

H_0 is going to be rejected if observed value of F_o is grater than its critical value $F_{\alpha,(a-1),a(n-1)}$. The critical value is taken from corresponding statistical table for significance level α, the degrees of freedom for the numerator $(a - 1)$, and degrees of freedom for the denominator $a(n - 1)$. Also, H_0 is going to be accepted if observed value is lower than its critical value ($F_o < F_{\alpha,(a-1),a(n-1)}$). If p-value approach for the statistics F_o is used than H_0 is going to be rejected if $p < \alpha$, and will be accepted if $p > \alpha$.

2.2. Linear regression

A general multiple linear regression model with k repressor variables is express as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{5}$$

where, y is response variable also called dependant variable, β_j, β_j, $j = 0. 1, 2, \ldots, k$, are regression coefficients or parameters and represent expected change in response variable per unit change in input variable, x_j, is regressesor variable or input variable or independent variable, ε is random error, which supposed to be normally distributed with N $(0, \sigma^2)$ for given values of x and y.

Parameter estimate in multiple linear regression models is done using least squares method. In case that there are multiple observations (n) on the response variable $y_1, y_2, \ldots y_n$, and that there are observation at each input variable x_{ij}, ($i = 1, 2, \ldots, n$) than it can be represented as matrix in Table 2 an written as [16]:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \tag{6}$$

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i; \quad i = 1, 2, \ldots, n \tag{7}$$

Table 2. Multiple linear regression model data

y	x_1	x_2	...	x_k
y_1	x_{11}	x_{12}	...	x_{1k}
y_2	x_{21}	x_{22}	...	x_{2k}
...	...	...	...	...
y_n	X_{n1}	X_{n2}	...	X_{nk}

Referring to Table 2, it can be formulated using the general multivariate linear regression model in matrix form as follows:

$$Y_{nx1} = X_{nx(k+1)}\beta_{(k+1)x1} + \varepsilon_{nx1} \tag{8}$$

where, Y denotes an (*n x 1*) vector of dependent variable observation, X is a *n x (k+1)* matrix of independent variables, β is a *((k+1) x 1)* vector of unobserved parameters of interest, and ε is an (*n x 1*) random errors vector. Subscript *n* is a number of observations, subscript *k* is a number of independent variables. The general matrix form of the model becomes as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{9}$$

After slowing the function that minimizes sum of squared errors ε^2, the least squared estimator of β can be calculated from the following:

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{10}$$

Significance test of regression model is to determine existence of the significant relationship between input variable and output variable. In this case it is tested hypothesis assuming that a least one β parameter differs from the others.

H₀: $\beta_1 = \beta_1 = ... = \beta_k = 0$

H₁: $\beta_j \neq 0$ for at least one *j* different than zero. $\tag{11}$

The procedure of test involves an *analysis of variance (ANOVA)* and performing *F*-test. Observed value is calculated as the ratio between regression mean squares (MS_R) and error mean squares MS_E (error variance):

$$F_o = \frac{MS_R}{MS_E} = \frac{\frac{SS_R}{k}}{\frac{SS_E}{n-k-1}} \tag{12}$$

whre, SS_R regression sum of squares, SS_E is error sum of squares, *k* is degree of freedom for the numerator, (*n – k – 1*) is degree of freedom for the denominator. Total sum of squares is addition of regression sum of squares and error sum of squares and it is calculated as:

$$SS_T = SS_R + SS_E = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{13}$$

whre, $\hat{y}_i$ is the predicted value for the i_{th} test, $\bar{y}$ is the mean of all response variable observations, y_i is the i_{th} observed value of the response variable.

$$H_0 \text{is} \begin{cases} \text{Rejected if } F_o > F_{\alpha,k,n-k-1} \\ \text{Accepted if } F_o < F_{\alpha,k,n-k-1} \end{cases} \tag{14}$$

H_0 is going to be rejected if observed value of F_o is grater than its critical value $F_{\alpha,k,n-k-1}$. The critical value is taken from corresponding statistical table for significance level α, the degrees of fridom for the numerator k, and degrees of freedom for the denominator $(n - k - 1)$. Also, H_0 is going to be accepted if observed value is lower than its critical value $(F_o < F_{\alpha,k,n-k-1})$. If p-value approach for the statistics F_o is used than H_0 is going to be rejected if $p < \alpha$, and will be accepted if $p > \alpha$.

Goodness of fit of a model is measured by coefficient of determination R^2, which is a measure of how well real data points are approximated with regression model. It measures how amount of reductions of variability in response variable caused using input variables in regression model.

$$R^2 = \frac{SS_R}{SS_T} \tag{15}$$

Higher value of R^2 means better fit. If $R^2 = 1$ it means that there is perfect fit. However, having a higher value dos not meat that there is a good fit and that regression model is good one because adding a new variable to the model (either the variable is significant or not) will increase R^2 value, which will lead to poor prediction. To solve this, an adjusted R^2 is introduced, which will not always increase with adding a new variable.

$$R^2_{adj} = \frac{SS_R}{SS_T} \tag{16}$$

If non-significant variable is added to the model R^2_{adj} will decrease and vice versa.

3. Factorial Design

Factorial experiments can be design with one, two, three and more factors. Experiments with only one factor are often called simple comparative experiments. In these cases, t–test or ANOVA were used for analysis.

Factorial experiments with two factors (*A* and *B*) usually include two level factorial designs for identification of factor effects on the response variable by investigating all possible combinations of the factor levels. The factor effect is defined as change in the response variable by changing the level of the factor.

Factorial experiments with multiple factors (*A*, *B*, …, *K*), with two levels ("low" and "high") the complexity of experimentation might be a problem. The number of possible combinations goes up with the number of factors, for instance a 2-level design with 8 factor has 256 combination which very set such type of experiments and analyze data. Multiple factor experiment requires a lot of resources, materials and it is time consuming and expensive. Additional problem with multiple factorial design is to maintain experimental conditions unchanged during a huge number of experiments.

Trying to overcome the problems with multiple factor factorial designs and depending from a case to case, it is possible to be designed as Full Factorial Design 2^k or Fractional Factorial Design 2^{k-p}. In this case number 2 represents number of levels, while k is number of factors and p is the fraction size of the full factorial used. More details about full factorial and fractional factorial design with examples is provided below [17].

3.1. Full Factorial Design

A full factorial design is convenient for a low number of factors if the resources are available. Conceptual approach for DOE is explained for two 2^2 and three 2^3 factors as well as general 2^k factorial design, in which k represents number of factors while number 2 represents number of levels. Uppercase letters A, B, C… are usually used for *factor* designation while lowercase letters are used treatments. Each factor has two levels *low* (–) and *high* (+). Number of combinations for 2^2 is four, for 2^3 is eight and so on. Each combination is called *treatment* which is represented with a lowercase letter. The number of test units for each treatment is called the

number of *replicates*. For example, if three units / samples were tested at each treatment, the number of replicates is three.

Design matrix for a 2^2 factorial design is given in Table 3:

Table 3. design matrix and signs for effects in 2^2 factorial design

	A	B	AB
−1	−1	−1	+1
a	+1	−1	−1
b	−1	+1	−1
ab	+1	+1	+1

where, A and B represent factors while AB represents interaction between A and B factor.

This design is an orthogonal. The following linear regression model is used for the analysis the analysis:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \qquad \textbf{(17)}$$

Where, x_1 and x_2 represent coded variables of factors A and B respectively, while $x_1 x_2$ cross–product and represents AB interaction, $x_j = -1$ and $x_j = +1$ represent *low* and *high* levels of factors.

$$ss = \frac{1}{n2^2}(contrast)^2 \qquad \textbf{(18)}$$

Whre, *contrast* is obtained from Table 3 as sum of products of signs in corresponding column and the runs listed in corresponding rows.

3.2. General 2^k Factorial Design

Experimental data analysis was done using Design of Experiment (DOE) – full factor factorial design. Generalized case of a 2^k factorial design is introduced and applied in this study, where k is number of factors at two levels. Statistical model includes k main effects, $\binom{k}{2}$ two–factor interaction, $\binom{k}{3}$ three–factor interactions, …, $\binom{k}{k}$ one k–factor interaction. The procedure for a 2^k factorial design was the following:

1. estimated factor effect – effects are estimated and their magnitudes were examined with the aim of important factors identification;

2. initial model formulation – full model is included that takes in account all main effects and interactions;

3. statistical testing – ANOVA is used to test significance of main effects and interactions;

4. model refinement – non significant factors from initial model are removed;

5. residual analysis – to check adequacy of the model and assumptions;

6. result interpretation – graphical analysis of the results such as main effects, interactions etc.

General form of ANOVA mentioned in step 3 is shown in the following table [17]:

Table 4. ANOVA for a general 2^k design

Source of variation	Sum of squares	Degrees of freedom	Mean Square	Fo
_k_ main effects				
_A_	SS_A	1	$MS_A = \dfrac{SS_A}{a-1}$	$F_o = \dfrac{MS_A}{MS_E}$
_B_	SS_B	1		
…	…	…		
_K_	SS_K	1		
$\binom{k}{2}$ **two–factor interaction**				
_AB_	SS_{AB}	1	$MS_{AB} = \dfrac{SS_{AB}}{(a-1)(b-1)}$	$F_o = \dfrac{MS_{AB}}{MS_E}$
_AC_	SS_{AC}	1		
…	…	…		
_JK_	SS_{JK}	1		
$\binom{k}{3}$ **three–factor interactions**				
ABC	SS_{ABC}	1	…	…
ABD	SS_{ABD}	1		
…	…	…		
IJK	SS_{IJK}	1		
…	…	…		
$\binom{k}{k}$ **_k_–factor interaction**				
ABC…K	$SS_{ABC…K}$	1	…	…
Error	SS_E	$2^k(n-1)$		
Total	SS_T	$n2^k-1$		

To estimate sum of squares and effects, contrasts associated with the effect were determined first.

$$Contrast_{AB…K} = (a \pm 1)(b \pm 1) … (k \pm 1) \qquad (19)$$

where, *a, b, …, k* are experiment treatment combinations.

After computation of the contrast, the **effects** and the **sum of squares** can be estimated as follows:

$$A, B … K = \frac{2}{n2^k}(Contrast_{AB…K}) \qquad (20)$$

$$SS_{A,B…K} = \frac{1}{n2^k}(Contrast_{AB…K})^2 \qquad (21)$$

where, *n* represents number of replicates.

Total sum of squares is calculated using the following equation:

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \left(y_{ijk} - \bar{y}_{...} \right)^2 = SS_A + SS_B + \cdots + SS_E \tag{22}$$

where, y_{ijk} is an observation for factor A at i-th level ($i = 1, 2, \ldots a$) and for B factor at j-th level ($j = 1, 2, \ldots b$) and for k-th replicate ($k = 1, 2, \ldots n$).

Sum of squares of error can be obtained from using the following equation:

$$SS_E = SS_T - SS_M \tag{23}$$

where, SS_M is model sum of squares and it is equal: $SS_M = SS_A + SS_B + \cdots + SS_K$.

3.3. Fractional Factorial Design

If the number of factors is increased, then the number of test units and treatment combinations (runs) is going to be increased. For example, for 5 factors, 32 units and treatments are needed ($2^5=32$). Since a low number of main effects and lower order interactions are significant to the response variable, and usually higher order interactions are not significant to the response variable, than fractional factorial designs are introduced. Therefore, fractional factorial designs take in account only a low number of main effects and lower order interactions. The higher order interactions are neglected due to its negligible effects on the response variable. For instance, $2^3 = 8$, and has 8 treatment combination and 8 test units is required. For some reason he cannot afford all 8 combinations he decided to run one half factorial design $2^{3-1} = 4$. Therefore, instead of 8 test units he will test only 4. Now he needs to determine which four treatment combinations to test. As the first step, he complete 2^3 full factorial design matrix shown in Table 5 and selected those with positive values in *ABC* column as it shown in

Table 6.

Table 5. Full 2^3 factorial design matrix

	I	A	B	AB	C	AC	BC	ABC
(1)	1	-1	-1	1	-1	1	1	-1
a	1	1	-1	-1	-1	-1	1	1
b	1	-1	1	-1	-1	1	-1	1
ab	1	1	1	1	-1	-1	-1	-1
c	1	-1	-1	1	1	-1	-1	1
ac	1	1	-1	-1	1	1	-1	-1
bc	1	-1	1	-1	1	-1	1	-1
abc	1	1	1	1	1	1	1	1

Table 6. The 2^{3-1} factorial design factorial design matrix with the defining relations $I = ABC$ and $I = -ABC$

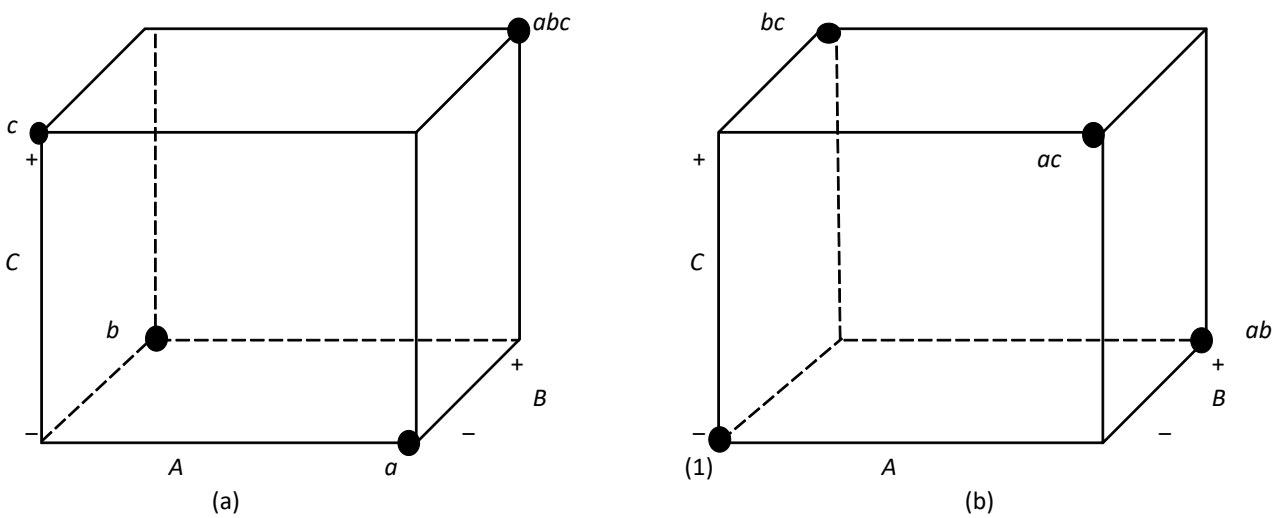|	I	A	B	AB	C	AC	BC	ABC
a	1	1	-1	-1	-1	-1	1	1
b	1	-1	1	-1	-1	1	-1	1
c	1	-1	-1	1	1	-1	-1	1
abc	1	1	1	1	1	1	1	1
(1)	1	-1	-1	1	-1	1	1	-1
ab	1	1	1	1	-1	-1	-1	-1
ac	1	1	-1	-1	1	1	-1	-1
bc	1	-1	1	-1	1	-1	1	-1

Figure 4. The 2^{3-1} factorial design factorial design orthogonal representation with defining relations: (a) principal fraction $I = ABC$; (c) alternative fraction $I = -ABC$

In this example ABC is called the generator. It noticed that some columns in

Table 6 have the same values (they are the same) and they are colored in the same color ($I = ABC$, $A = BC$, $B = AC$, $C = AB$), where I is called *defining relation*, which may be aliased with several effects. The alias can be determined by using the defining relation $I = ABC$ by multiplying any column by the defining relation yields the aliases such as: $A \cdot I = A \cdot ABC = A^2BC$. Since square of any column is equal to identity I, then it can be written as: $A = BC$. The other aliases may be obtained, $B \cdot I = B \cdot ABC = AB^2C$ and $C \cdot I = C \cdot ABC = ABC^2$, and replacing squared member by I it is obtained $B = AC$, $C = AB$. Therefore, when A, B and C is estimated it is actuality estimated $A + BC$, $B + AC$, $C + AB$.

Resolution a two-level fractional factorial design 2^{3-1} is called a **resolution III design** and it is equal to the number of letters in the shortest word in the defining relation in which Roman numbers in subscript are used to define resolution (e.g. 2_{III}^{3-1} design for $I = ABC$ *and* $I = -ABC$; 2_{IV}^{4-1} design for $I = ABCD$ *and* $I = -ABCD$ etc.).

Designs with a resolution less than III are never used. Fractional factorial designs with the highest available resolution with their defining relations is shown in Table 7.

Table 7. The 2^{3-1} factorial design factorial design matrix with the defining relations $I = ABC$ and $I = -ABC$

Number of factors	Fraction	Resolution	Runs (treatments)	Defining relation
3	2^{3-1}	III	4	$I = ABC$
4	2^{4-1}	IV	8	$I = ABCD$
5	2^{5-1}	V	16	$I = ABCDE$
6	2^{6-1}	VI	32	$I = ABCDEF$
7	2^{7-1}	VII	64	$I = ABCDEFG$

4. Application

4.1. The 2^4 Full Factorial Design

DOE a very suitable tool used for process performance and product quality improvement [18], for solar technologies [19] etc. In this example DOE is applied on injection-molding process with the aim of improving product quality such as excessive flash. Factors considered as affecting for flash formation are: pack pressure (*A*), pack time (*B*), injection speed (*C*), and screw RPM (*D*), while clamping pressure, injection pressure and melting temperature were under control. Each factor affecting flash formation is considered at low and high levels and shown in Table 8.

Table 8. Factors contributing to flash formation

FACTORS	low (–1)	High +1
Pack pressure (*A*) in bar	10	30
Pack time (*B*), in sec	1	5
Injection speed (*C*) in mm/sec	12	50
Screw RPM (*D*) in rpm	100	200

The design matrix for 2^4 factorial design and single replicate response data for flash size in millimeters is shown in Table 9.

Table 9. Design matrix for flash formation

Run	Run label	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	BCD	ACD	ABD	ABCD	Flash (mm)
1	(1)	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1	0.22
2	a	1	-1	-1	-1	-1	-1	-1	1	1	1	1	-1	1	1	-1	6.18
3	b	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1	0
4	ab	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	1	1	-1	1	5.91
5	c	-1	-1	1	-1	1	-1	1	-1	1	-1	1	1	1	-1	-1	6.6
6	ac	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	6.05
7	bc	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	-1	1	1	1	6.76
8	abc	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1	8.65
9	d	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1	0.46
10	ad	1	-1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	5.06
11	bd	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	0.55
12	abd	1	1	-1	1	1	-1	1	-1	1	-1	-1	-1	-1	1	-1	4.84
13	cd	-1	-1	1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1	11.55
14	acd	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1	9.9
15	bcd	-1	1	1	1	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	9.9
16	abcd	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	9.9

Applying Equations (19) and Equation (20) in the design matrix (Table 9) main effects are calculated and shown in Table 10.

Table 10. Factor effect estimate and sum of squares

Model term	Effects	Coefficients = (effect / 2)	SS	Percent contribution
A	2.5575	1.27	26.16323	0.12126637
B	0.061875	0.03	0.015314	7.0981E-05
C	5.76125	2.88	132.768	0.6153788
D	1.47125	0.73	8.658306	0.04013119
AB	0.466125	0.23	0.86909	0.00402823
AC	-2.6345	-1.31	27.76236	0.12867835
AD	-0.748	-0.374	2.238016	0.01037319
BC	0.218625	0.10	0.191188	0.00088615
BD	-0.507375	-0.25	1.029718	0.00477273
CD	1.82325	0.91	13.29696	0.06163133
ABC	0.556875	0.27	1.240439	0.00574943
BCD	-0.598125	-0.29	1.431014	0.00663274
ACD	2.22045E-16	1.11E-16	1.97E-31	9.1409E-34
ABD	-0.130625	-0.06	0.068252	0.00031635
ABCD	-0.067375	-0.03	0.018158	8.416E-05
Mean[4]		5.78		

In this example there is a problem to apply F-test. Since there is only one replicate of the response variable F-test is not possible to get done because error degrees of freedom is zero ($MS_E = SS_E/0$; – division by zero is undefined). Potential solution can be a graphical solution such as Normal probability diagram of effects [5], or dropping entire factors from the model. Therefore based on Table 10, a graphical solution is used to represent main effects and identification of those significant, which is shown in Figure 5.
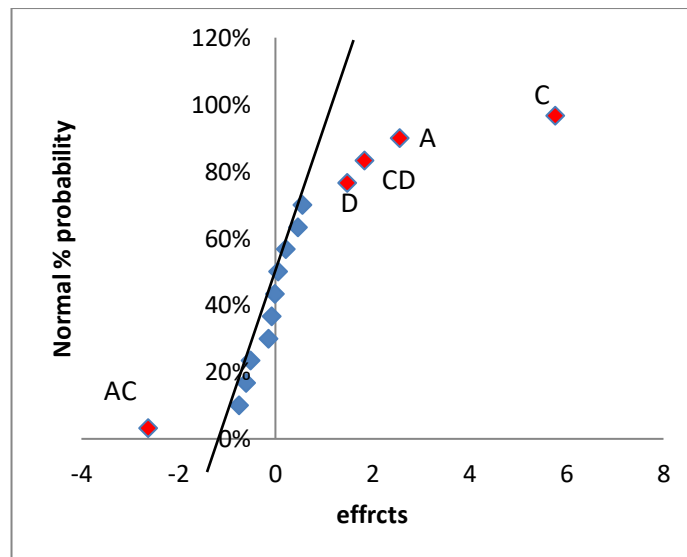

Figure 5. Normal probability plot of the effects

---

[4] Mean response, calculated as average flash value.

From normal probability diagram of effects it is observed that *A*, *C*, *D*, are significant factors and *AC*, and *CD* are significant interactions. Plots of the main effects of factor *A*, *C*, *D* as well as *AC* and *CD* interactions are shown in Figure 6.
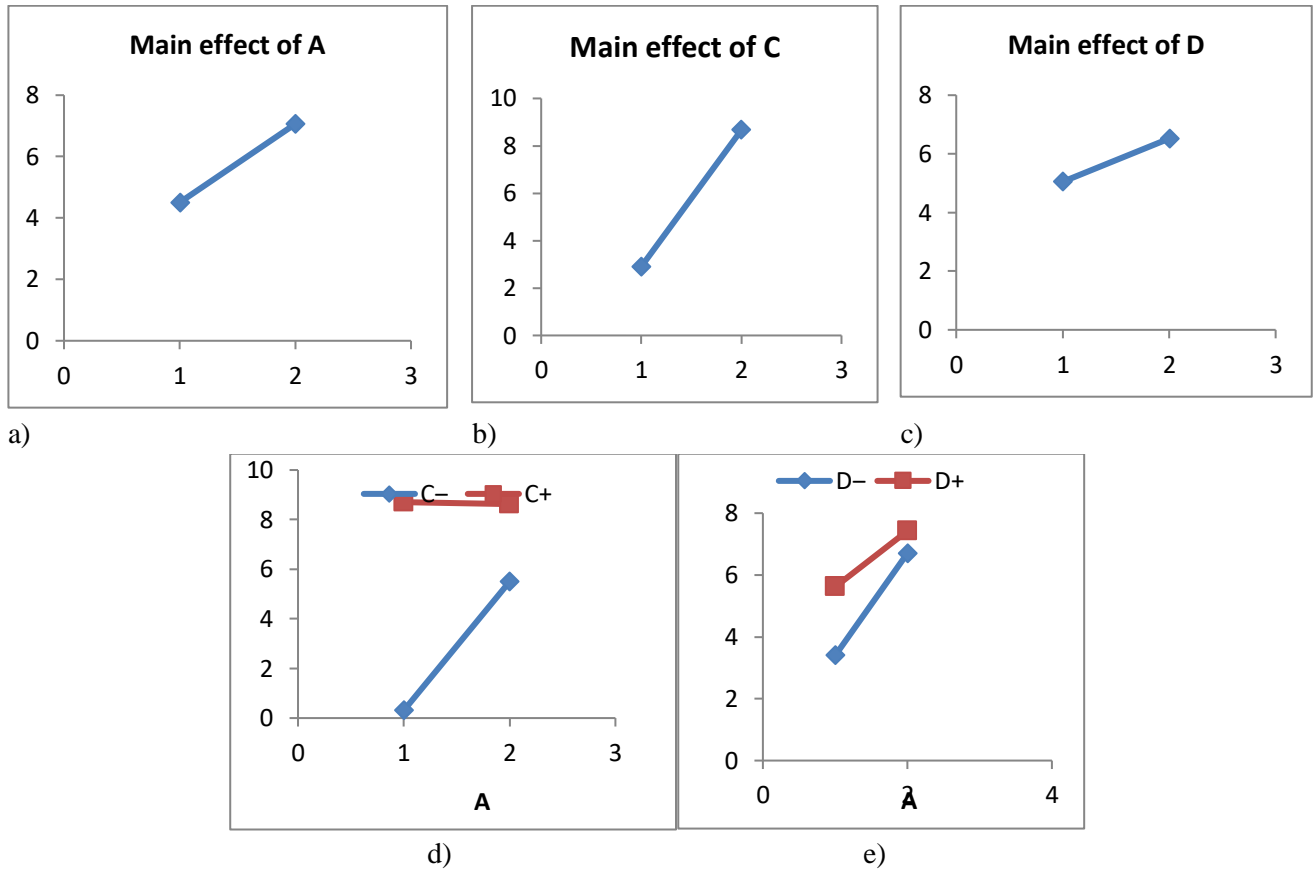

Figure 6. Main effects and interactions

Refereeing to Figure 6 it is observed that all three main effects are positive indicating that the molding process supposed to be performed on low level to minimize flash size. Alongside main effects there are significant interactions. Since main effect are involved with significant interactions than they have no much meaning. Screw RPM and injection speed have little meaning at high pack pressure but has better performances in reducing flash size at low level. Therefore, the best setup for getting better product quality (reducing flash size), is obtained when all factors *A*, *C*, *D* are at the low level.

Since factor *B* is not significant it is dropped from the model. All associated interactions with *B* factor are dropped from as well. Therefore, the design becomes a 2^3 factorial design with two replicates so called *hidden replication* (still has 16 responses). In this the mean case sum of squares of error (MS$_E$) can be calculated because degrees of freedom are different than zero ($2^k(n–1) = 2^3(2–1) = 8$). ANOVA result for 2^3 factorial design with two replicates is shown in Table 11.

Table 11. Factor effect estimate and sum of squares

	SS	DF	MSS	Percent contribution	F_0	F_{cr} = $F_{\alpha=5\%,v1,v2}$	$F_0 > F_{\alpha=5\%,v1,v2}$
A*	26.16	1	26.16	0.1212	29.47	5.32	Sign.
C*	132.76	1	132.76	0.6153	149.57	5.32	Sign.
D*	8.65	1	8.65	0.0401	9.75	5.32	Sign.
AC*	27.76	1	27.76	0.1286	31.27	5.32	Sign.
CD*	13.29	1	13.29	0.0616	14.97	5.32	Sign.
Model (SS$_M$)	208.64						
Error (SS$_E$)	7.10	8	0.88				
Total (SS$_T$)	215.75	15	14.38				

Where, total sum of squares (SS$_T$) is obtained using Equation (22), model sum of squares is obtained as SS$_M$ = Σ SS$_{A,B,C...K}$ = SS$_A$ + SS$_C$ + SS$_D$ + SS$_{AD}$ + SS$_{CD}$, while error sum of squares is obtained as SS$_E$ = SS$_T$ − SS$_M$.

Regression model that represent **predicted/fitted** value for flash formation is given below, and includes only significant factor and interactions.

$$\hat{y} = 5.78 + 1.3x_1 + 2.9x_3 + 0.7x_4 - 1.3x_1x_3 + 0.9x_3x_4 \qquad (24)$$

Based on Equation (24), it is possible to perform diagnostic check if it is applied to the residuals.

Table 12. Factor effect estimate and sum of squares

Run label	y	$\hat{y}$	$e = y - \hat{y}$
(1)	0.22	-0.427625	0.647625
a	6.1875	4.764375	1.423125
b	0	-0.427625	0.427625
ab	5.9125	4.764375	1.148125
c	6.6	7.968125	-1.368125
ac	6.05	7.891125	-1.841125
bc	6.765	7.968125	-1.203125
abc	8.657	7.891125	0.765875
d	0.462	1.043625	-0.581625
ad	5.06	6.235625	-1.175625
bd	0.55	1.043625	-0.493625
abd	4.84	6.235625	-1.395625
cd	11.55	9.439375	2.110625
acd	9.9	9.362375	0.537625
bcd	9.9	9.439375	0.460625
abcd	9.9	9.362375	0.537625

Normal probability plot of residuals and residuals versus predicted value of flash size is shown in Figure 7.
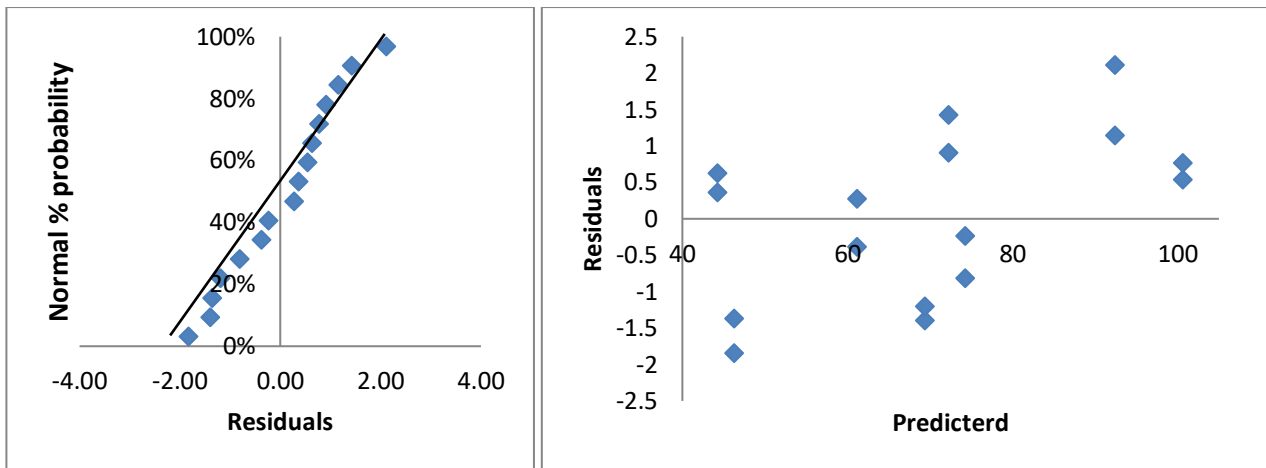
Figure 7. Model adequacy checking

Residuals versus predicted value plot shows randomly dispersed data around the horizontal axis (predicted values), therefore the variance has no tendency to change along predicted value axis indicating that linear regression model is appropriate for the data.

4.2. The 2^{4-1} Fractional Factorial Design

Applying the 2^{4-1} fractional factorial design on the previous example with the highest possible resolution (IV) and using defining relation $I=ABCD$ it is obtained the 2^3 design, which is shown in Table 13.

Table 13. The 2_{IV}^{4-1} factorial design with the defining relation $I = ABCD$

		Basic Design						**Flash**				**Residuals**
RUN	Run label	A	B	C	D=ABC	AB	AC	AD	y	$\hat{y}$	$\Sigma(y_i-\overline{y})^2$	$e = y - \hat{y}$
	alias	BCD	ACD	ABD	ABC	BC	BD	CD				
1	(1)	-1	-1	-1	-1	1	1	1	0.22	-0.8975	30.4704	1.1175
2	ab	1	1	-1	-1	1	-1	-1	5.9	4.7375	0.0256	1.1625
3	ac	1	-1	1	-1	-1	1	-1	6.05	7.2125	0.0961	-1.1625
4	bc	-1	1	1	-1	-1	-1	1	6.7	7.8175	0.9216	-1.1175
5	ad	1	-1	-1	1	-1	-1	1	5.1	6.2625	0.4096	-1.1625
6	bd	-1	1	-1	1	-1	1	-1	0.55	1.6675	26.9361	-1.1175
7	cd	-1	-1	1	1	1	-1	-1	11.5	10.3825	33.1776	1.1175
8	abcd	1	1	1	1	1	1	1	9.9	8.7375	17.3056	1.1625
	mean								5.74			

Referring to Table 13 factors A, B and C, represent basic design factors that are aliased with three factor interactions BCD, ACD and ABD respectively, while factor D is aliased with ABC three factor interaction. The two factor interactions AB, AC and AD are re aliased with the other two factor interactions BC, BD and CD respectively. Therefore, the relationships are: A = BCD, B = ACD, C = ABD, D = ABC, AB = BC, AC = BD and AD = CD. Using Equation (20) and multiplying corresponding column with measured values (flash) and summing up, a **contrast** is obtained for the certain column. Using Eq. (21), contrast is divided by $n2^k$ and main effects are obtained for number of replicates $n = 1$, and number of factors $k = 3$ and shown in Table 14.

Table 14. Estimate of effects and aliases

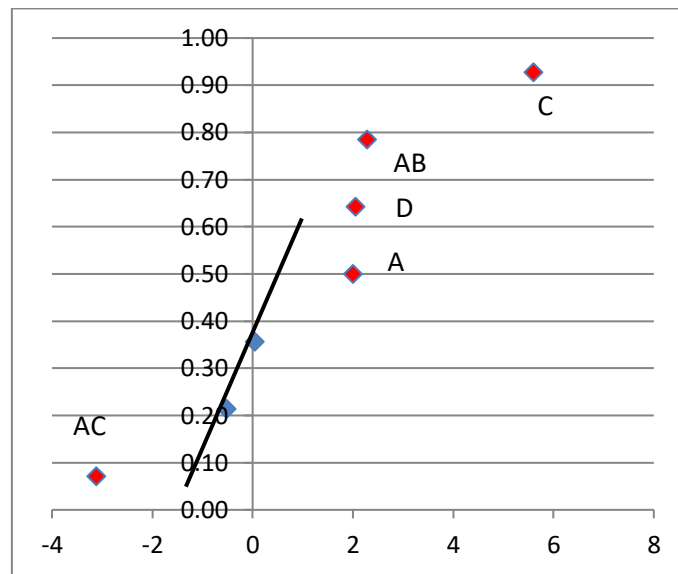| Effect estimate	Aliased structure
[A] = 1.995	[A] → A + BCD
[B] = 0.045	[B] → B + ACD
[C] = 5.595	[C] → C + ABD
[D] = 2.045	[D] → D + ABC
[AB] = 2.28	[AB] → AB + BC
[AC] = -3.12	[AC] → AC + BD
[AD] = -0.52	[AD] → AD + CD

Plot of main effects is shown in Figure 8.

Figure 8. Normal Probability plot of effects

From normal probability diagram of effects, it is observed that *A*, *C,* and *D*, are significant factors and *AC*, and *AB* are significant interactions. Based on the above analysis it is possible to predict flash size using the following regression model:

$$\hat{y} = 5.74 + 1.99x_1 + 5.59x_3 + 2.04x_4 + 2.28x_1x_2 - 3.12x_1x_3 \qquad (25)$$

5.   Conclusion

Thanks to software development, the use of DOE as scientific tool has increased rapidly in past 20 years in manufacturing and non–manufacturing industries over the world. It was most popular tool used by scientist in medicine (with 18%), engineering and biochemistry (with 20%), physics and computer science (with 13%), providing about 50% participation of these scientific areas compared to the all other scientific areas. The trend of use of DOE is rapidly growing and it is expected to slow down in the near future for current scientific areas, but also will expanded over new scientific areas and have rapid grow there.

Implementing valid and efficient factorial experiments provide quantitative data that can be used as support for decision making during system, process and product design or improvement. Full factorial design and fractional factorial design with examples of application of DOE in product quality improvement is presented with sep–by–step procedures and result interpretation. This can be a very useful guide to experimenter how to design and conduct experiments, and how to analyze and interpret data.

Refferences

[1] B. Duraković and H. Bašić, "Continuous Quality Improvement in Textile Processing by Statistical Process Control Tools: A Case Study of Medium-Sized Company," *Periodicals of Engineering and Natural Sciences,* vol. 1, no. 1, pp. 36-46, 2013.

[2] F. Paulo and L. Santos, "Design of experiments for microencapsulation applications: A review," *Materials Science and Engineering: C,* vol. 77, no. August, pp. 1327-1340, 2017.

[3] P. Yu, M. Y. Low and W. Zhou, "Design of experiments and regression modelling in food flavour and sensory analysis: A review," *Trends in Food Science & Technology,* vol. 71, no. January, pp. 202-215, 2018.

[4] A. Schlueter and P. Geyer, "Linking BIM and Design of Experiments to balance architectural and technical design factors for energy performance," *Automation in Construction,* vol. 86, no. February , pp. 33-43, 2018.

[5] B. Durakovic and M. Torlak, "Simulation and experimental validation of phase change material and water used as heat storage medium in window applications," *J. of Mater. and Environ. Sci.,* vol. 8, no. 5, pp. 1837-1746, 2017.

[6] D. B. Hibbert, "Experimental design in chromatography: A tutorial review," *Journal of Chromatography B,* vol. 910, no. Dec, pp. 2-13, 2012.

[7] S. S. Garud, I. A. Karimi and M. Kraft, "Design of computer experiments: A review," *Computers & Chemical Engineering,* vol. 106, no. November, pp. 71-95, 2017.

[8] J. K. Telford, "A Brief Introduction to Design of Experiments," *Johns Hopkins APL Technical Digest,* vol. 27, no. 3, pp. 224-232, 2007.

[9] B. Durakovic, H. Bašić and H. Muhič, "The Interrelationships between quality managment practicies and their effects on innovation," in *Trends in the Development of Machinery and Associated Technology*, Budapest, 2014.

[10] H. Guo and A. Mettas, "Design of Experiments and Data Analysis," in *2012 Annual Reliability and Maintainability Symposium*.

[11] B. Duraković and H. Bašić, "Textile Cutting Process Optimization Model Based On Six Sigma Methodology In A Medium-Sized Company," *Journal of Trends in the Development of Machinery and Associated Technology,* vol. 16, no. 1, pp. 107-110, 2012.

[12] B. Durakovic and M. Torlak, "Experimental and numerical study of a PCM window model as a thermal energy storage unit," *International Journal of Low-Carbon Technologies,* vol. 12, no. 3, p. 272–280, 2017.

[13] "Scopus," 2017. [Online]. Available: www.scopus.com. [Accessed September 2017].

[14] D. Granato, V. M. d. A. Calado and B. Jarvis, "Observations on the use of statistical methods in Food Science and Technology," *Food Research International,* vol. 55, p. 137–149, 2014.

[15] D. C. Montgomery, Statistical Quality Control, John Wiley, 2009.

[16] D. C. Montgomery and G. C. Runger, Applied Statistics and Probability for Engineers, Willey, 2005.

[17] D. C. Montgomery, Design and Analysis of Experiments, 5th edition, John Wiley and Sons, Inc. New York, 2001., 8 ed., John Wiley & Sons, Inc, 2013.

[18] H. Bašić, B. Duraković and A. Softić, "Six Sigma Model Testing In Optimizing Medium-Sized Company Production Process," *Journal of Trends in the Development of Machinery and Associated Technology,* vol. 16, no. 1, pp. 103-106, 2012.

[19] V. Okatia, A. Behzadmehra and S. Farsad, "Analysis of a solar desalinator (humidification–dehumidification cycle) including a compound system consisting of a solar humidifier and subsurface condenser using DoE," *Desalination,* vol. 397, no. November, pp. 9-21, 2016.