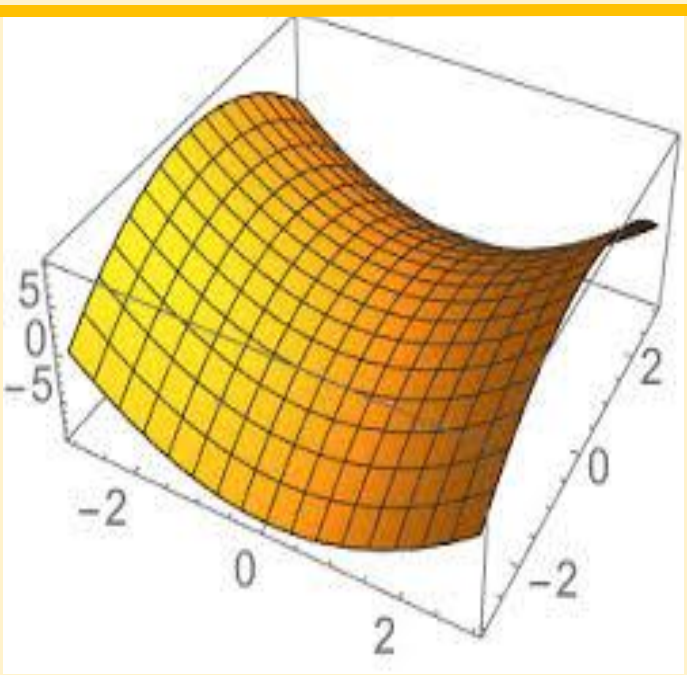
People's Democratic Republic of Algeria Ministry
of Higher Education and Scientific Research
University Abderrahmane Mira - Bejaia
Faculty of Natural and Life Sciences
Department of Microbiology

Course

INTRODUCTION TO DESIGN OF EXPERIMENTS



DR DJOUDI WARDA

2024

Preface

The experimentation plays an important role in Science, Engineering, and Industry. The experimentation is an application of treatments to experimental units, and then measurement of one or more responses. It is a part of scientific method. It requires observing and gathering information about how process and system works. In an experiment, some input x's transform into an output that has one or more observable response variables y. Therefore, useful results and conclusions can be drawn by experiment. In order to obtain an objective conclusion an experimenter needs to plan and design the experiment, and analyze the results.

Design of experiments (DOE) is a formal structured technique for studying any situation that involves a response that varies as a function of one or more independent variables. DOE is specifically designed to address complex problems where more than one variable may affect a response and two or more variables may interact with each other.

There are many types of experiments used in real-world situations and problems. When treatments are from a continuous range of values then the true relationship between y and x's might not be known. The approximation of the response function $y = f(x_1, x_2,\ldots,x_i) + \varepsilon$ is called Response Surface Methodology.

The present course is intended for master's students from different specialties of the Faculty of Natural and Life Sciences either for other faculties as well as for our doctoral students who want to model and optimize the results of their research works to have the advantage of organizing their experiments and to reduce the time to carry out their trials, in order to obtain the maximum amount of information on the process studied.

This course offers the students to introduce and understands the principle and steps of DOE modeling. The three types of Response Surface Methodology, the first-order and the second-order models, will be described and explained in this course. Some examples will be provided mainly for two levels factorial design and Centrale Composite Design (CCD) using excel software for statistical analysis of the regression. Demonstration how to manipulate with Minitab computer software and using excel software will be also given. Finally, a series of exercises are presented too.

List of Figures

Figure 1 : Black box model.

Figure 2: Progressive use of DOE as scientific method over past two decades.

Figure 3 : Factor’s domain.

Figure 4 : Experimental space.

Figure 5 : Study domain.

Figure 6 : Response surface.

Figure 7: Illustration of global effect and main effect.

Figure 8: Interactions plots.

Figure 9: 2^k factorial designs (a) k=2, 2^2 factorial design requires four experiments (b) k=3, 2^3 factorial design requires eight experiments

Figure 10: Effect of the factors on the response.

Figure 11: Interaction profile illustrating the importance of the interactions.

Figure 12: Interaction profiles regrouped into a single table.

Figure 13: Standardized Pareto plot of effects for the above experiment

Figure 14: Main Effects plot for the experiment

Figure 15: Mixture designs types

Figure 16: Exemple of Contour plot (a) and Response surface plot (b)

Figure 17: Illustration of central composite designs for **(a)** two factors and **(b)** three factors optimization. Every design consists of factorial points (●), star points (*), and central points (□).

Figure 18: Illustration of the two graphical forms for the three factors BBD: (a) the cube for BBD and (b) three interlocking 2^2 factorial design.

Figure 19: Illustration of the model of the Doehlert design for the optimization of (a) two variables and (b) three variables.

Figure 20: Minitab icon.

Figure 21: Minitab home window (worksheet).

Figure 22: Menus giving access to design of experiments.

Figure 23: Window offering several types of experimental designs.

Figure 24: Window allowing you to choose a full or fractional factorial design

Figure 25: Window allowing to define the characteristics of the factors

Figure 26: Window allowing to specify the characteristics of the experimental design.

Figure 27: Worksheet of the design created.

Figure 28: Worksheet of the design and responses.

Figure 29: Window allowing the analysis of the design

Figure 30: Window allowing response column selection to analysis

Figure 31: Window allowing factors selection to analysis.

Figure 32: Window offering several types of experimental designs.

Figure 33: Window allowing the choice a full factorial design.

Figure 34: Window allowing to specify the factors.

Figure 35: Window allowing to specify the experimental design.

Figure 36: Window allowing to indicate the results of the mathematical analysis.

Figure 37: Minitab Worksheet for 2 factors.

Figure 38: Responses are entered into the Minitab worksheet.

Figure 39: Drop-down menus allowing you to begin the analysis of the experimental design.

Figure 40: Window allowing you to specify the analysis methods.

Figure 41: Window allowing you to choose the mathematical model.

Figure 42: Window allowing you to select useful results.

Figure 43: Outputs choice.

Figure 44: The results analysis is in the session window.

Figure 45: The analysis results are also displayed in the worksheet.

Figure 46: Worksheet with the example data.

Figure 47: Session sheet with results: Effects and coefficients.

Figure 48: Window allowing us to choose the effects graph, the interactions graph and the cube graph.

Figure 49: Definition of effects graphs.

Figure 50: Effects graph.

Figure 51: Interaction graph.

Figure 52: Menu allowing you to create a response surface design

Figure 53: Menu allowing you to specify the response surface design

Figure 54: Menu allowing you to specify the characteristics of the plan
Figure 55: Minitab worksheet the CCD response surface design
Figure 56: Design matrix and responses of CCD
Figure 57: Analyze a response surface design window allowing you to define the desired analyses.
Figure 58: Analyze Response Surface design window - Terms for defining the mathematical model.
Figure 59: Minitab Worksheet with responses and coefficients
Figure 60: Statistical analysis results in session window
Figure 61: Contour Plot/Surface Plot window
Figure 62: Contour Plot/Surface contour Plot window
Figure 63: Window allowing you to specify the nature of the isoresponse graph
Figure 64: Contour plots of the surface roughness
Figure 65: Menu allowing you to create a Box-Behnken design
Figure 66: Minitab worksheet with the Box-Behnken design
Figure 67: Minitab worksheet with the Box-Behnken design and responses
Figure 68: Session sheet giving the coefficients of the quadratic model
Figure 69: Contour Plot/Surface Plot window of BBD
Figure 70: Response and factors choice to illustrate in the graph
Figure 71: We impose a level to the factors which are not illustrated on the graph
Figure 72: Acid loss as a function of the concentration and the pH for a dilution of 0.5

List of Tables

Table 1: Experimental matrix of factorial designs (2^k)

Table 2: Calculation matrix of coefficient for k=3 factors

Table 3: Calculation coefficient matrix for k= 2

Table 4: Factors and study domain

Table 5: Experimental matrix and results

Table 6: Effects and interactions of the factors (coded units) Response: Deposition speed

Table 7: An 8 run geometric P–B design

Table 8: List of process parameters and their levels for the experiment

Table 9: Experimental Layout for 12 run P–B design with response values

Table 10: Values of α and N_0 according to the properties of the composite design

Table 11: Calculation matrix of CCD for 2 factors and N_0= 4

Table 12: Calculation matrix of CCD for 3 factors and N_0= 4 replicates

Table 13: Levels and factors values

Table 14: Design matrix and response

Table 15: Effect calculation matrix and the calculated effect

Table 16. Residual calculation results

Table 17: Analysis of variance (ANOVA) table

Table 18: Factors and study domain

Table 19: Factors values at different levels

Table 20: Study domain of CCD

Table 21: Design domain of BBD

Table of contents

Preface

List of Figures

List of Tables

1. Introduction ……………………………………………………………………1
2. Advantages of DOE……………………………………………………………3
3. Historical perspective…………………………………………………………..4
4. Terminology……………………………………………………………………5
   4.1.Factors types…………………………………………………………………..5
   4.2.Factor's domain………………………………………………………………5
   4.3. Study domain…………………………………………………………………6
   4.4. The Response Surface………………………………………………………..6
   4.5.Centered and Scaled Variables…………………. …………………………………7
   4.6. Experimental matrix…………………………………………………………..8
   4.7. Calculation matrix of effects………………………………………………9
   4.8.Factor main effect……………………………………………………………10
   4.9.Interaction effect……………………………………………………………..11
5. Mathematical modeling of the response ……………………………………11
6. Types of designs of experiments…………………………………………………14
   6.1.First order factorial designs………………………………………………14
   6.2.Second order experimental designs (Response Surface Methodology) ………..….27
7. Statistical analysis of results……………………………………………………..33
   7.1.Test of coefficients significance……………………………………………..33
   7.2. Analysis of Variance (ANOVA)……………………………………………36
   7.3.The coefficient of determination R^2……………………………………………39
8. Optimization…………………………………………………………………..40
9. DOE Softwares…………………………………………………………………..42
10. Demonstration using Minitab software manipulation………………………....43
    10.1. Starting Minitab……………………………………………………………43
    10.2. Access to design of experiments……………………………………………44

10.3. Design of experiments definition……….…………………… ……………45
10.4. Factors definition ……………………..…………………………………47
10.5. Effects plots in Minitab……………………………………………………51

11. Applications using Minitab 16 software package ……………………….……..53
12. Application using Excel software………………………………………………81
13. Exercises……………………………………………………………………...88
13. Definitions of some matrix………………………………………………...94
14. Student table………………………………………………….…………..97
15.Fisher table ………………………………………………………………98
References……………………………………...………………...……………..99

1. Introduction

Investigators perform experiments in virtually all fields of inquiry, usually to discover something about a particular process or system. Each experimental **run** is a **test**. More formally. we can define an **experiment** as a test or series of runs in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes that may be observed in the output response. We may want to determine which input variables are responsible for the observed changes in the response, develop a model relating the response to the important input variables and to use this model for process or system improvement or other decision-making [1].

In general, experiments are used to study the performance of processes and systems. The process or system can be represented by a **black box** model shown in figure 1. We can usually visualize the process as a combination of operations. Machines, methods, people, and other resources that transforms some input (often a material) into an output that has one or more observable response variables. Controllable variables $(x_1.\ x_2.\ \ldots\ldots x_p)$ can be varied easily during an experiment and such variables have a key role to play in the process characterization. Uncontrollable variables ($z_1.\ z_2.\ .\ .\ .\ .\ zq$) are difficult to control during an experiment. These variables or factors are responsible for variability in process performance. It is important to determine the optimal settings of *x*'s in order to minimize the effects of *z*'s. This is the fundamental strategy of robust design [1, 2].

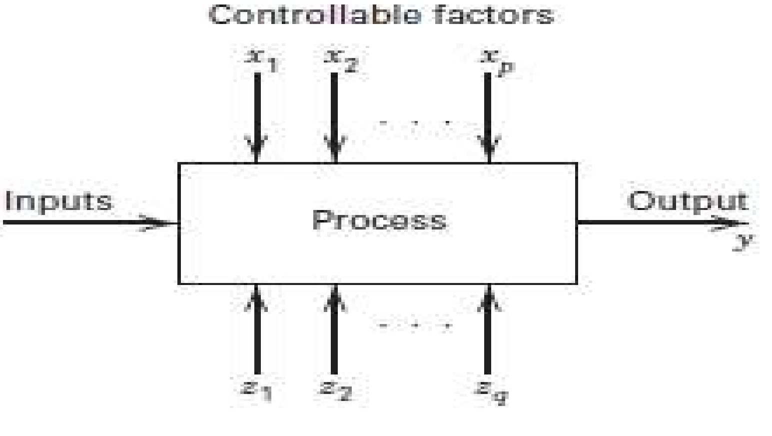


Figure 1 : Black box model [1]

Experimentation plays an important role in technology commercialization and product realization activities, which consist of new product design and formulation. manufacturing process development. and process improvement. The objective in many cases may be to develop a robust process. that is. a process affected minimally by external sources of variability. There are also many applications of designed experiments in a nonmanufacturing or non-product-development setting. such as marketing, service operations. and general business operations. A well-designed experiment is important because the results and conclusions that can be drawn from the experiment depend to a large extent on the manner in which the data were collected [1-4].

Statistical experimental design; also known as design of experiments (DOE); is the methodology of how to conduct and plan experiments in order to extract the maximum amount of information with the lowest number of trials. A designed experiment is a tool or set of tools used for gathering test data. Typical characteristics of an experimental design are planned testing. data analysis approach. simultaneous factor variability and scientific approach [5].

DOE is a branch of applied statistics that is used for conducting scientific studies of a system. process or product in which input variables **(x_i)** were manipulated to investigate its effects on measured response variable **(y).** 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. food industry. energy and architecture, and chromatography [3].

DOE is applied in experimental situations where several independent variables potentially impact one or more response variable. The experimenter controls the independent variable in a designed experiment, while the response variable is an observed output of the experiment. Changing more than one variable simultaneously, rather than changing one variable at a time, leads to effective results. Interactions between variables can cause problems that none can see until change has been made. DOE has been applied in many functional areas, one being research to quantify the inter-relationship between variables and to screen a large number of variables to identify important ones [2, 6].

The popularity of DOE is due to its tremendous power and efficiency. When used correctly, DOE can provide the answers to specific questions about the behavior of a system.

using an optimum number of experimental observations. Since designed experiments are structured to answer specific questions with statistical rigor. experiments with too few observations won't deliver the desired confidence in the results and experiments with too many observations will waste resources. DOE gives the answers that we seek with a minimum expenditure of time and resources [3, 7].

General practical steps and guidelines for planning and conducting DOE are listed below [3]:

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: S**election of independent variable (factors) that cause change in the response variable.

4. Determine 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.

6. Data analysis: Using statistical methods such as regression analysis and ANOVA.

7. Practical conclusions and recommendations: Graphical representation of the results and validation.

2. Advantages of DOE

Advantages of the design of experiments [3-5]:

• Helps to handle experimental error.

• Helps to determine the important variables that need to be controlled and find the unimportant variables that need not be controlled.

• Helps to measure interactions. which is very important.

• Allows extrapolation of data and search for the best possible product within the test variable ranges.

• Allows plotting graphs to depict how variables are related and what level of variables give the optimum product. Use of statistical models shows us the interrelationship between variables.

3. 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. The traditional approach demands considerable material expense and is more time consuming. The major disadvantage of the OFAT strategy is that it fails to consider any possible **interaction** between the factors [3].

Testing multiple variables at a time is better especially in cases where data must be analyzed carefully. In the 1920s and 1930s **Ronald A. Fisher** conducted a research in agriculture. he was the first one who started using DOE. In 1935, he wrote a book on DOE. 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 (figure 2). Thanks to software development in 1990s and later, the use of DOE in research over various scientific areas has risen sharply [3, 8].

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.

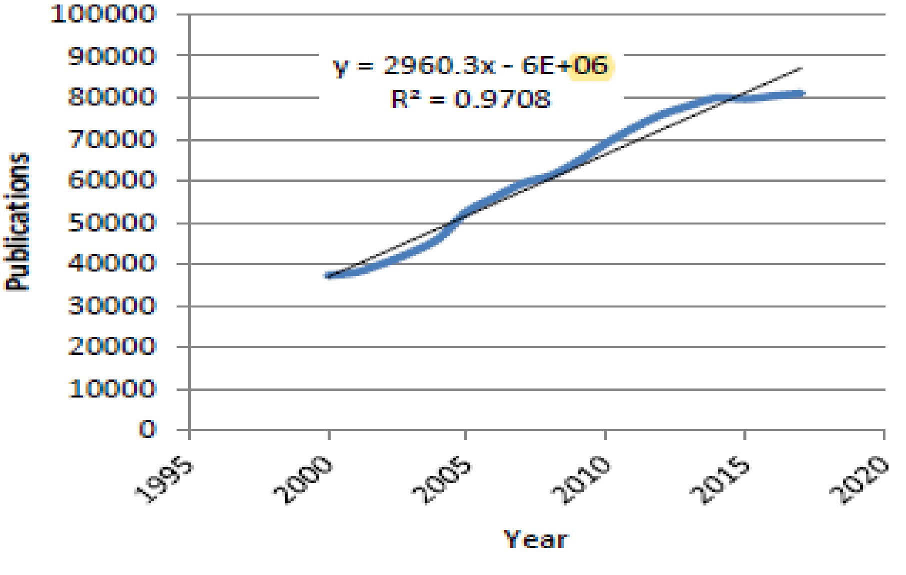


Figure 2: Progressive use of DOE as scientific method over past two decades [3]

4. Terminology

To simplify the communication a few different terms are introduced and defined

4.1. Factors types

Experimental variables that can be changed independently of each other, two types of factors exist [4, 6, 8]:

- ✓ **Continuous Variables:** Independent variables that can be changed continuously like pression. temperature. concentration….
- ✓ **Discrete Variables**: Discrete factors can take only particular values. These values are not necessarily numeric. The color of a product (blue. red. or yellow) is an example of a discrete factor, for example. size may be represented as large, medium. or small……

4.2. Factor's domain

The value given to a factor while running an experimental trial is called a **level**. The lower limit is the **low level (-1)** and the upper limit is the **high level (+1).** The set containing all the values between the low and the high level that the factor can take is called the **factor's domain of variation** or. more simply, the **factor's domain** (figure 3) [5, 6].

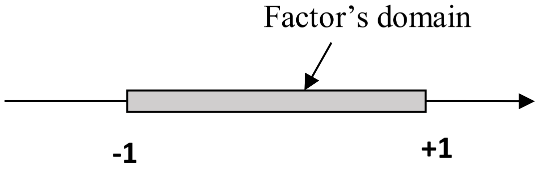


Figure 3: Factor's domain [4]

One continuous factor can be represented by a directed and graduated axis. If there is a second continuous factor, it is represented by a similar axis drawn orthogonally to the first. This area is called **the experimental space** (figure 4). The experimental space is composed of all the points of the plane factor 1 × factor 2 where each point represents an experimental trial.

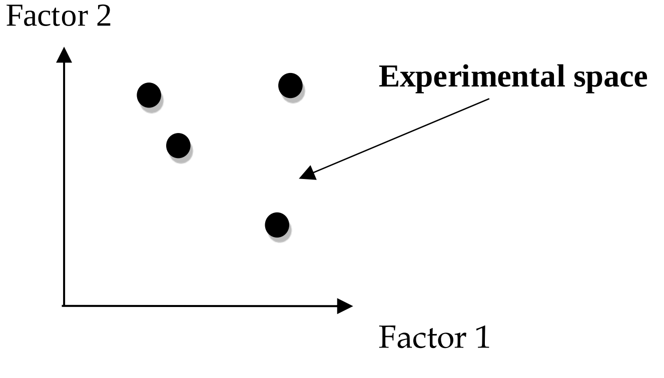


Figure 4 : Experimental space [4]

4.3. Study domain

The study domain is defined as a portion of experimental space to carry out the study, this domain is defined by the high and low levels of all the factors (figure 5) [4, 9].

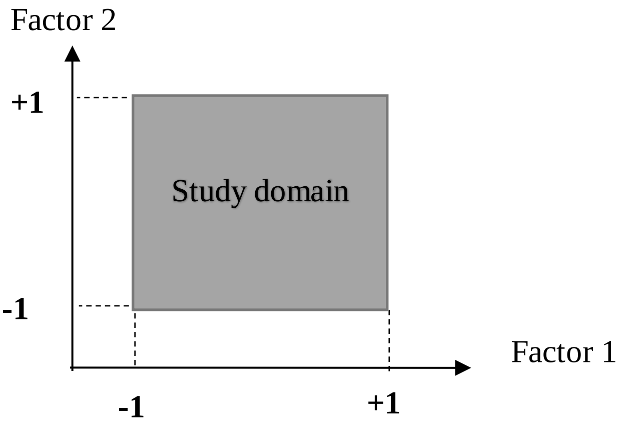


Figure 5 : Study domain [4]

4.4. The Response Surface

The collection of responses that correspond to all the points in the study domain forms the response surface. To obtain the response surface. it is necessary to interpolate using a mathematical model (figure 6) [4].

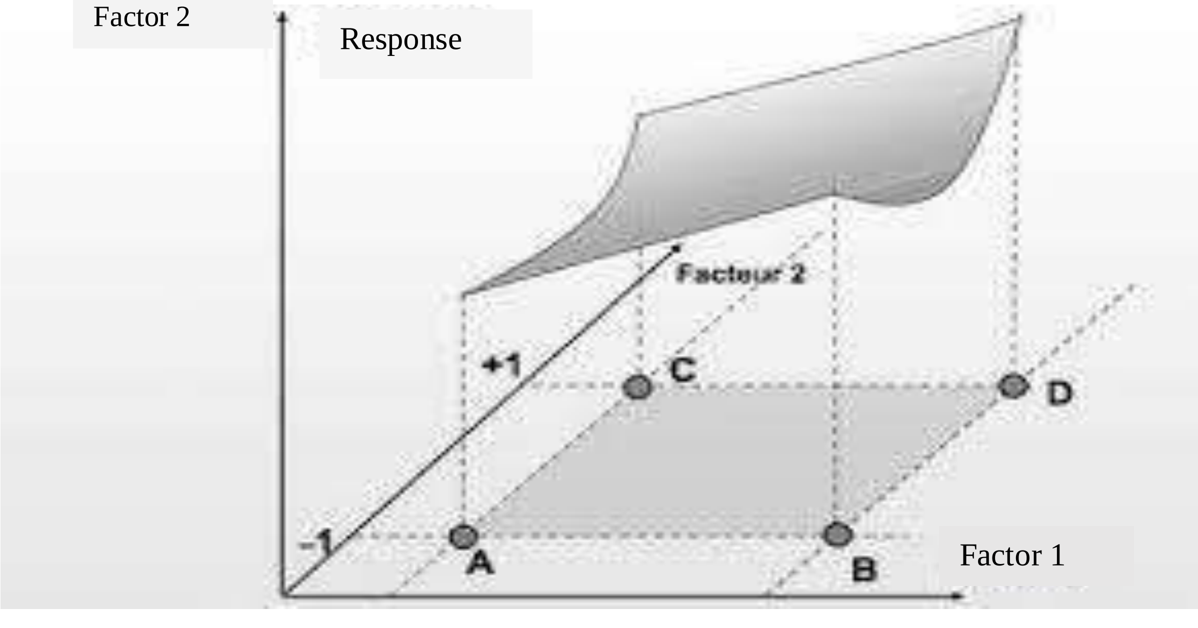


Figure 6 : Response surface [4]

4.5. Centered and Scaled Variables

When the lower level of a factor is represented by (-1) and the upper level is represented by (+1), two important changes occur. these two changes involve the introduction of new variables called **centered and scaled variables**. The conversion of the original variables $\boldsymbol{A_j}$ to the coded variables x_j (and vice versa) is given by the following formula. where $\boldsymbol{A_0}$ is thex central value [1, 4, 5].

Where: $$\boldsymbol{x_j = \frac{A_j - A_0}{Step}} \quad \textbf{(1)}$$

With:

$$A_0 = \frac{A_{+1} + A_{-1}}{2} \qquad Step = \frac{A_{+1} - A_{-1}}{2}$$

Example 1:

An experimenter chooses for the temperature factor to be 20°C at the low level (-1) and 60°C at the high level (+1). In coded units. what is the corresponding temperature for 30°C? Let's calculate the step for the speed factor. It's equal to half the difference between the high and low levels, so:

$$Step = \frac{60 - 20}{2} = 20°C$$

A_0 is the center value between the high and low levels; that is, it is half of the sum of the high and low levels:

$$A_0 = \frac{60 + 20}{2} = 40°C$$

$$X = \frac{A - A_0}{Step} = \frac{30 - 40}{20} = -0.5$$

A temperature of 30 °C is therefore, for this example. equal to **-0.5** in coded values.

Example 2:

We may also want the value in original units. knowing the coded value. What is the value of the temperature factor corresponding to +0.5 in coded units?
Write equation (1):

$$+0.5 = \frac{A - 40}{20}$$

So: A= 40+ (20×0.5) =50

The coded temperature 0.5 corresponds to a temperature of **50 °C**.

The advantage to using coded units lies in their power to present designed experiments in the same way, regardless of the chosen study domains and regardless of the factors. Seen this way. DOE theory is quite generalizable**.**

4.6.Experimental matrix

The experimental matrix (or design matrix) is the table that indicates the number of experiments to be carried out with how to vary the factors and the order in which the experiments must be carried out, this table can be arranged using either the original variables or the coded variables (-1 and +1) [4, 9].

For 2^k design. build a table with 2^k rows and k columns, the rows are labeled with factor-level combinations in standard order, and the columns are labeled with the k factors. In principle, the body of the table contains +1's and −1's. with +1 indicating a factor at a high level, and −1 indicating a factor at a low level [4, 9].

In Table 1, the factorial designs for 2, 3 and 4 experimental variables are shown. for example, for 2 factors factorial design (2^2), the first column of the matrix is used to designate the trials numbers. The second column holds the first factor (x_1), with its designated levels listed in order. The third column holds the second factor (x_2) and also lists the experimental runs in order. The results are written in the fourth column of the experimental matrix.

Table 1: Experimental matrix of factorial designs (2^k)

Two variables 2^2			**Three variables 2^3**				**Four variables 2^4**				
N°	x_1	x_2	**N°**	x_1	x_2	x_3	**N°**	x_1	x_2	x_3	x_4
1	**- 1**	**-1**	**1**	**-1**	**-1**	**-1**	**1**	**-1**	**-1**	**-1**	**-1**
2	**+1**	**-1**	**2**	**+1**	**- 1**	**-1**	**2**	**+1**	**- 1**	**-1**	**-1**
3	**- 1**	**+1**	**3**	**-1**	**+1**	**-1**	**3**	**-1**	**+1**	**-1**	**-1**
4	**+1**	**+1**	**4**	**+1**	**+1**	**-1**	**4**	**+1**	**+1**	**-1**	**-1**
			5	**-1**	**- 1**	**+1**	**5**	**-1**	**- 1**	**+1**	**-1**
			6	**+1**	**- 1**	**+1**	**6**	**+1**	**- 1**	**+1**	**-1**
			7	**-1**	**+1**	**+1**	**7**	**-1**	**+1**	**+1**	**-1**
			8	**+1**	**+1**	**+1**	**8**	**+1**	**+1**	**+1**	**-1**
							9	**- 1**	**-1**	**-1**	**+1**
							10	**+1**	**-1**	**-1**	**+1**
							11	**-1**	**+1**	**-1**	**+1**

		12 +1 +1 -1 +1 **13 -1 -1 +1 +1** **14 +1 -1 +1 +1** **15 -1 +1 +1 +1** **16 +1 +1 +1 +1**

4.7.Calculation matrix of effects

The calculation matrix of effects used to calculate the model coefficients; it is obtained by adding to the left of the experiment matrix. a column containing only 1 s, corresponding to the fictive variable x_0. The other columns correspond to the interactions of the different factors, they are obtained by performing the line by line product of the columns of the corresponding factors (table 2).

Once the signs for the main effects have been established. the signs for the remaining columns can be obtained by multiplying the appropriate preceding columns row by row. For example, the signs in ($x_1.x_2$) column are the product of the x_1 and x_2 column signs in each row.

Table 2: Calculation matrix of coefficient for k=3 factors [4, 9]

x_0	x_1	x_2	x_3	$x_1 x_2$	$x_1 x_3$	$x_2 x_3$	$x_1 x_2 x_3$	Y
1	-1	-1	-1	1	1	1	-1	y_1
1	1	-1	-1	-1	-1	1	1	y_2
1	-1	1	-1	-1	1	-1	1	y_3
1	1	1	-1	1	-1	-1	-1	y_4
1	-1	-1	1	1	-1	-1	1	y_5
1	1	-1	1	-1	1	-1	-1	y_6
1	-1	1	1	-1	-1	1	-1	y_7
1	1	1	1	1	1	1	1	y_8

Table 2 has several interesting properties: (1) Except for column x_0. every column has an equal number of plus and minus signs, (2) The sum of the products of the signs in any two columns is zero, (3) Column x_0 multiplied times any column leaves that column unchanged, (4) The product of any two columns yields a column in the table. For example, $x_1 \times x_2 = x_1 x_2$.

4.8. Factor main effect

The effect of a factor "A" on the response "*y*" is obtained by comparing the values taken by "*y*" when A increases from level (-1) to level (+1). Let y_1 and y_2 are these values (figure 7) [2, 4- 6].

We distinguish:

- Global effect of factor A by $(y_2 - y_1)$.
- Main effect of factor A by $(y_2 - y_1)/2$.

A main effect plot (figure 7) is a plot of the mean response values at each level of a design process variable. One can use this plot to compare the relative strength of the effects of various factors. The sign and magnitude of a main effect would tell us the following:

- ➢ The sign of a main effect tells us of the direction of the effect. i.e. if the average response value increases or decreases.

- ➢ The magnitude tells us of the strength of the effect.

If the effect of a design variable is positive, it implies that the average response is higher at high level than at low level of the parameter setting. In contrast, if the effect is negative, it means that the average response at the low level setting of the parameter is more than at the high level [1, 4, 5].

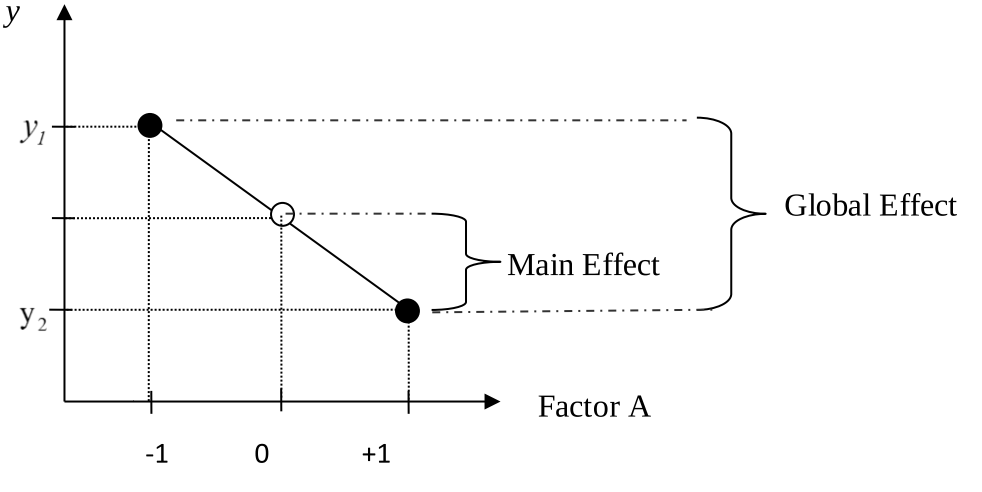


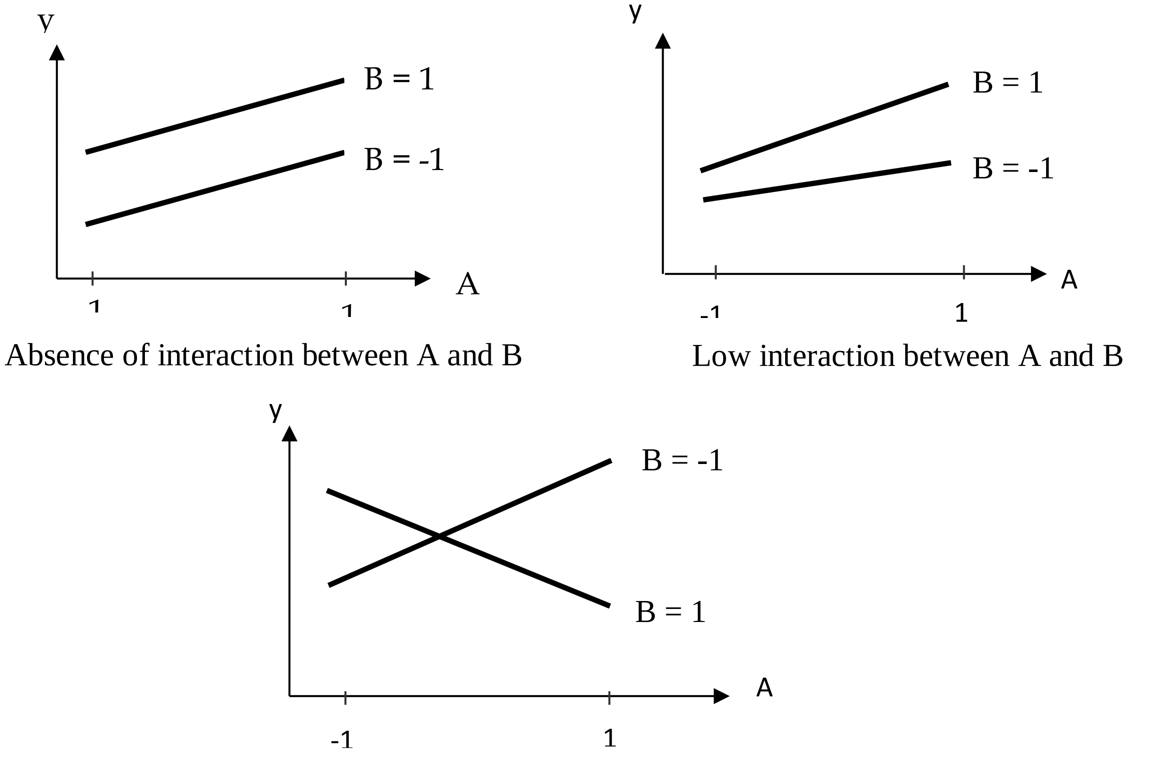
Figure 7: Illustration of global effect and main effect [5, 6].

4.9. Interaction effect

Interaction occur between two factors A and B if the effect of A on the response depends on the level of B or vice versa (figure 8). In other words. the effect of A on the response is different at different levels of B. The interaction between A and B can be computed using the following equation [1, 2, 4, 8]:

$$I_{A.B} = \frac{1}{2}\left(E_{A.B(+1)} - E_{A.B(-1)}\right) \qquad \textbf{(2)}$$

Where $E_{A.B(+1)}$ is the effect of factor ‘A’ at high level of factor ‘B’ and where $E_{A.B(-1)}$ is the effect of factor ‘A’ at low level of factor ‘B’.



Absence of interaction between A and B

Low interaction between A and B

High interaction between A and B

Figure 8 : Interactions plots [1, 2, 4, 8]

5. Mathematical modeling of the response

It is reasonable to assume that the outcome of an experiment is dependent on the experimental conditions. This means that the result can be described as a function based on the experimental variable [2, 5-8]:

$$y = f(x_i) + \varepsilon \qquad \textbf{(3)}$$

The function $f(x_i)$ is approximated by a polynomial function and represents a good description of the relationship between the experimental variables and the responses within a limited experimental domain. This function is expressed as [2, 5-8]:

$$\hat{y} = b_0 + \sum_{j=1}^{k} b_j x_j + \sum_{\substack{u,\ j=1 \\ u \neq j}}^{k} b_{uj} x_u x_j + \sum_{f=1}^{k} b_{ujf} x_u x_j x_f + \sum_{j=1}^{k} b_{jj} x_j^2 \quad \textbf{(4)}$$

Where: $\hat{y}$ is the response also called dependent variable;

ε is the pure error which comes from the response measurement;

x_u represents a level of factor u;

x_j represents a level of factor j;

b_0, b_j, b_{uj}, b_{ujf} and b_{jj} are the coefficients of the polynomial model.

This model is called the ***a priori*** model, or the ***postulated*** model.

Then: $y = \hat{y} + \varepsilon$

Three types of polynomial models will be discussed and exemplified with two variables, x_1 and x_2.

The simplest polynomial model contains only linear terms and describes only the linear relationship between the experimental variables and the responses. In a ***linear model***, the two variables x_1 and x_2 are expressed as:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 \quad \textbf{(5)}$$

The next level of polynomial models contains additional terms that describe the interaction between different experimental variables. Thus. a ***second order interaction model*** contains the following terms:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 \quad \textbf{(6)}$$

The two models above are mainly used to investigate the experimental system, i.e.. with screening studies, robustness tests or similar.

To be able to determine an optimum maximum or minimum. quadratic terms have to be introduced in the model. By introducing these terms in the model, it is possible to determine non-linear relationships between the experimental variables and responses.

The polynomial function below describes a quadratic model with two variables:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2 \qquad (7)$$

- **Linear regression**

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. …y_n. and that there is observation at each input variable x_{ij}. ($i = 1. 2. .... n$) than it can be represented as matrix notation [3, 4, 8]:

$$\boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (8)$$

where:

y is the **response vector**;

$\boldsymbol{X}$ is the **model matrix** or the **design matrix** which depends on the experimental points used in the design and on the postulated model;

$\boldsymbol{\beta}$ is the **coefficients matrix**;

ε is the **error matrix**;

The general matrix form of the model becomes as follows:

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

This system of equations cannot be. in general, solved simply because there are fewer equations than there are unknowns. To find the solution., we must use special matrix methods generally based on the criterion of least squares. The results are estimations of the coefficients. denoted as β.

The algebraic result of the least-squares calculations is [3, 4, 8]:

$$B = \left[X^T.X\right]^{-1}.[X]^T.Y \qquad (9)$$

Where X^T is the transpose matrix of X (appendix 1).

Two matrices appear frequently in the theory of experimental design:

- The **information matrix** $X^T X$
- The **dispersion matrix** $[X^T X]^{-1}$

6. Types of designs of experiments

Different types of designs are available; their choice is determined by the objectives of the experiment and the current state of knowledge about the experimental environment. They can be categorized as follows:

- Screening;
- Factorial design;
- Mixture design;
- Response surface design

In this course. we will essentially detail the full factorial design as an example of a first order model development, and the centered composite design (CCD) as an example for RSM modeling, for the other designs we will give some principles.

6.1.First order factorial designs

6.1.1. Full (2^k) or fractional ($2^{k–r}$) factorial experimental designs

Factorial design and fractional factorial design which both of them with two levels for each factor (k) were commonly used in process of screening design due to their efficiency and economical consideration.

In a full factorial experiment, responses are measured at all combinations of the experimental factor levels. The combinations of factor levels represent the conditions at which responses will be measured. Each experimental condition is a called a "run" and the response measurement an observation. The entire set of runs is the "design". Fractional factorial design enables the evaluation of a relatively large number of factors in small number of runs or experiments. This method was designed by fractioning a full factorial design of 2^k combinations into 2^{k-r} combinations; It should be noted that the fractional factorial design can reduce the number of runs or experiments but it does not possible to estimate all major and interaction effect separately [2, 3, 5-10].

The expression of full factorial design regression equation with interactions is as follow [1, 2, 4-11].

$$\hat{y} = b_0 + \sum_{j=1}^{k} b_j x_j + \sum_{\substack{u,\ j=1 \\ u \neq j}}^{k} b_{uj} x_u x_j + \sum_{f=1}^{k} b_{ujf} x_u x_j x_f \qquad \textbf{(10)}$$

Where ' $\hat{y}$ ' stands for the predicted response, $x_{j.}$ stands for the settings (factors), $b_j.b_{uj}$ and b_{ujf} are the respective coefficients and b_0 stands for the intercept of mean.

Majority of factorial experiments are composed of only two-level factors with four treatment combinations in total (2^k. where k is the number of factors) and are generally called as 2x2 factorial designs (figure 9 (a)). If there are three factors. the full factorial design points are at the vertices of a cube (figure 9 (b)) and for more factors. the design points are the vertices of a hypercube [3, 5, 6, 9, 11].

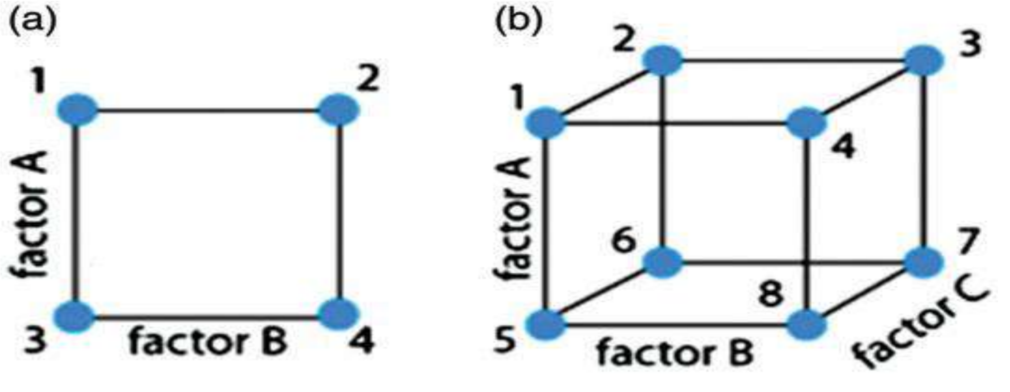


Figure 9: 2^k factorial designs (a) k=2, 2^2 factorial design requires four experiments (b) k=3, 2^3 factorial design requires eight experiments

- **Mean effects coefficients estimation**

  The simple calculation of the model coefficients comes from the algebraic properties of the effect's matrix of the factorial designs. **Fisher** and **Yate** showed that an orthogonal matrix leads to the independence of the model coefficient estimates. The scientist **Jacques Hadamard** (appendix 1) demonstrated the following expression [3, 5, 9-11]:

$$[X]^T.[X] = N[I]$$

Where. [I] is the identity matrix. N the number of experiments and $[X]^T$ the transpose matrix of [X].

The dispersion matrix $[X^T.X]^{-1}$ is written as follows:

$$[X^T.X]^{-1} = \begin{bmatrix} 1/N & .... & 0 \\ . & 1/N & . \\ 0 & .... & 1/N \end{bmatrix}$$

The calculation of the coefficients is then done by the scalar product of column (***y***) by the corresponding column **(*x_j*)**, divided by the number of trials **N**. Thus, for linear effects. the values of the coefficients are determined by:

$$b_j = \frac{1}{N}\sum_{i=1}^{N} x_{ji}y_i \ , \quad j = 0,1,..k \tag{11}$$

$$b_{uj} = \frac{1}{N}\sum_{i=1}^{N} (x_u x_j)_i \, y_i \quad . \quad j = 1.....k \quad . \quad u = 1..... k \quad . \quad j \neq u \tag{12}$$

$$b_{ujf} = \frac{1}{N}\sum_{i=1}^{N} (x_u x_j x_f)_i \, y_i \ , \ f = 1.......k \tag{13}$$

Example 1:

We want to test the influence of **P**ressure (x_1) and **T**emperature (x_2) (two factors) on the yield **(Y)** of a chemical reaction (response), for this purpose, a 2^2 factorial design is used. The design matrix is as follow (table 3) [4, 6, 14-16].

As the number of trials N= 4. we deduce that:

$$\boldsymbol{b_0} = \frac{1}{4}(\mathbf{60+78+63+89}) = \mathbf{72.5}$$

$$\boldsymbol{b_1} = \frac{1}{4}(\mathbf{-60+78-63+89}) = \mathbf{11}$$

$$\boldsymbol{b_2} = \frac{1}{4}(\mathbf{-60-78+63+89}) = \mathbf{3.5}$$

$$\boldsymbol{b_{12}} = \frac{1}{4}(\mathbf{60-78-63+89}) = \mathbf{2}$$

Table 3: Calculation coefficient matrix for k= 2 factors

Trial	x_0	x_1	x_2	$x_1.x_2$	Y (%)
1	**1**	**-1**	**-1**	**+1**	**60**
2	**1**	**+1**	**-1**	**-1**	**78**
3	**1**	**-1**	**+1**	**-1**	**63**
4	**1**	**+1**	**+1**	**+1**	**89**
Level (–1)	2 bars		50 °C		
Level (+1)	4 bars		70 °C		

These results allow us to write the model giving chemical reaction yield according to the levels of the two factors (in coded units):

$$\hat{y}= 72.5 + 11.x_1 + 3.5.x_2 + 2.x_1.x_2$$

The function above is now describing how the experimental variables and their interactions influence the response $\hat{Y}$. The model shows that variable x_1 (the pression) has the largest influence on the yield, because its coefficient ($b_1 = \mathbf{11} > b_2 = \mathbf{3.5}$) is the most important. Besides this, the pression coefficient has positive sign, this means that an increase of the pression from (-1) to (+1) results in an increase of the reaction yield by 2×11=22% (2 x Main effect).

Example 2:

Gold-plated jewelry is covered with a thin layer of gold that must look identical to solid gold. and also have mechanical resistance to ensure a long life. Three variables are chosen to study their influence on gold deposition speed by electrolysis (electrochemical process). The low and high levels of factors are summarized in table 4 and the experimental matrix and results is given in table 5.

The study objective is to carry out the deposition as quickly as possible while maintaining quality (the response is the speed of gold deposition) [4].

Table 4: Factors and study domain

Factor	Low Level (-)	High Level (+)
Gold concentration (1)	2 g/L	15 g/L
Density of the current (2)	5 A/dm^2	25 A/dm^2
Cobalt concentration (3)	0.5 g/L	1.5 g/L

Table 5: Experimental matrix and results

Trial	Gold Concentration (g/L)	Density of the Current (A/dm^2)	Cobalt Concentration (g/L)	Speed (mg/min)
	Factor 1	*Factor 2*	*Factor 3*	
1	−1	−1	−1	53
2	+1	−1	−1	122
3	−1	+1	−1	20
4	+1	+1	−1	125
5	−1	−1	+1	48
6	+1	−1	+1	70
7	−1	+1	+1	68
8	+1	+1	+1	134
−1 Level	2 g/L	5 A/dm^2	0.5 g/L	
+1 Level	15 g/L	25 A/dm^2	1.5 g/L	

Since we have three factors each taking two levels. and since we think that a linear model is sufficient to explain the phenomena, a 2^3 factorial design is used.

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

The different coefficients are calculated and summarized in table 6.

Table 6: Effects and interactions of the factors (coded units)
Response: Deposition speed

Effect	Value
Intercept	80
Gold concentration (1)	32.75
Current density (2)	6.75
Cobalt concentration (3)	0
1×2 Interaction	10
1×3 Interaction	−10.75
2×3 Interaction	14.25
1×2×3 Interaction	1

These results allow us to write the model giving deposition speed according to the levels of the three factors (in coded units):

$$y_{speed} = 80 + 32.75x_1 + 6.75x_2 + 10x_1x_2 - 10.75x_1x_3 + 14.25x_2x_3 + x_1x_2x_3$$

Factor 1 (x_1: gold concentration) is the most influential followed by Factors 2 (x_2). Factor 3 (x_3) is not directly influential, but it is influential via the interaction with other factors. The three second order interaction terms are close to one another except for the difference in sign. The third-order interaction is small by comparison.

We use this equation to make calculations and to draw graphs using Minitab package software as applications in this course.

Two diagrams are useful for a deeper understanding of the influence of the factors: the **effect diagram**, which indicates the principal effects of the factors. and the **interaction diagram**, which shows the second-order interactions among the factors. The different steps to draw these diagrams of this example using Minitab software package are given in applications part of this course (example 2).

- **Effect Diagram**

The prediction profiler shows the principal effects of the factors, i.e., the coefficients of the first-degree terms of the mathematical model. The diagram can be constructed with coded units or with original or natural units (figure 10), since the appearance is the same. When presenting results, it is much easier to use the natural units. which give immediate values for comparison [1, 4].

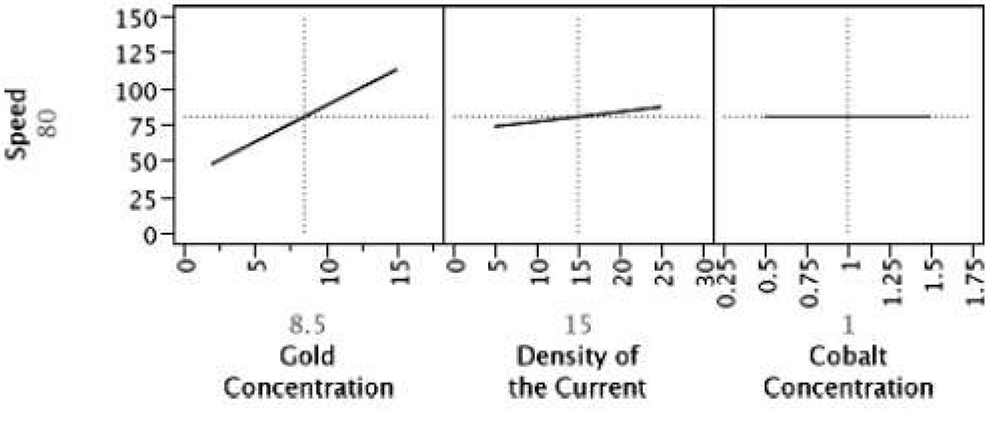


Figure 10: Effect of the factors on the response [1, 4]

The deposition rate grows larger as the solution contains more gold and as the current density is slightly raised. The cobalt concentration of the electrolytic solution does not seem to play any part in the reaction. But, to have a complete interpretation of the results, we have to take the interactions into account, and we have seen that they are not negligible.

An interaction diagram shows the effects of a factor at low and high levels on another factor. The diagram in figure 11 can be interpreted as follows.

The response is shown on the y-axis and the scales of the factors are on the x-axis. In the upper right square, the effect of the current density factor is shown for low (2) and high (15) levels of gold concentration. In the lower left square. the gold concentration factor is shown for low (5) and high (25) levels of the current density factor.

If the lines are not parallel. there is a significant interaction. This is the case for two factors where the slopes of the effects are different. and the lines cross each other.

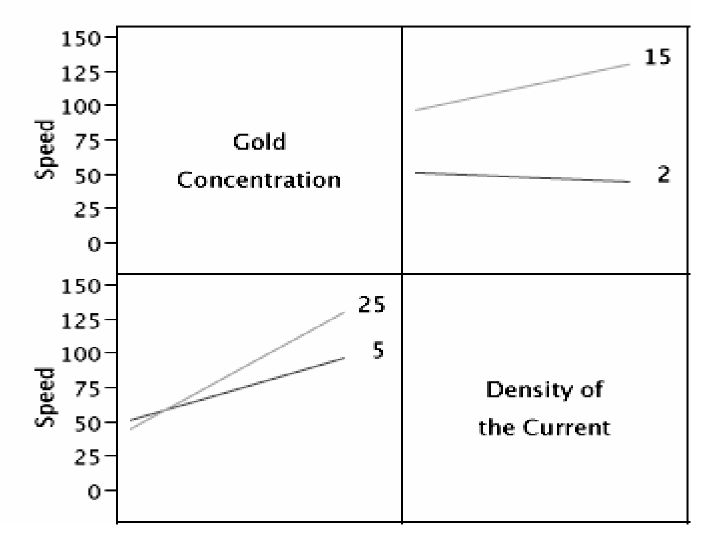


Figure 11: Interaction profile illustrating the importance of the interactions [1, 4]

We could also present the interaction plot by taking the interactions two at a time and drawing them in a single table as in figure 12.

The interaction plot clearly shows that the interactions are not negligible and that they must be taken into account during the interpretation of the results.

The maximum deposition speed is attained when the three factors are at their high levels: 15 g/L of gold. 1.5 g/L of cobalt in the electrolytic solution. and a current density of 25 A/dm^2. which gives a deposition speed of 134 mg/min.

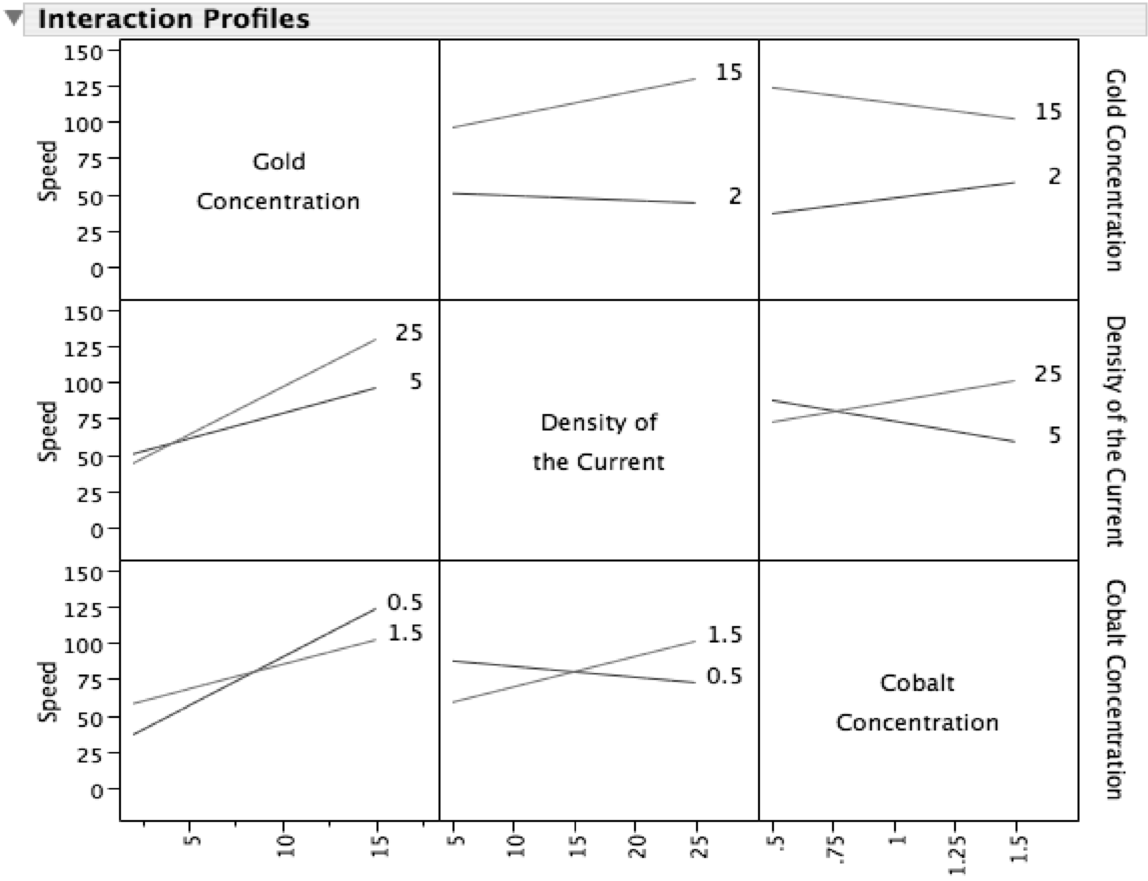


Figure 12: Interaction profiles regrouped into a single table

6.1.2. Plackett Burman design

A popular class of screening designs is the Plackett-Burman design (PBD), developed by R.L. Plackett and J.P. Burman in 1946. It was designed to improve the quality control process that could be used to study the effects of design parameters on the system state so that intelligent decisions can be made. Plackett and Burman (PB) devised orthogonal arrays are useful for screening, which yield unbiased estimates of all main effects in the smallest design possible. Various number or 'n' factors can be screened in an 'n + 1' run PB design (table 7) [1, 2, 5, 6, 10].

Plackett-Burman designs were applied as a screening method to evaluate the most significant factors with the fewest experiments, they are based on Hadamard matrices in which the number of experimental runs or trials is a multiple of four, i.e, N=4. 8. 12. 16. . . . and so on, where N is the number of trials/runs, and the results are interpreted using the first-degree polynomial model [1-6, 9,10].

$$\hat{y} = b_0 + \sum_{j=1}^{k} b_j x_j \qquad (10)$$

Where ' $\hat{y}$ ' stands for the predicted response, x_j stands for the settings (factors), *bj* are the respective coefficients and b_0 stands for the intercept of mean.

For screening designs. experimenters are generally not interested to investigate the nature of interactions among the factors. The aim is to study as many factors as possible in a minimum number of trials and identifying those that need to be studied in further rounds of experimentation [3, 9].

Table 7 illustrates the competed design matrix for 8 run P–B design. this allows one to study up to 7 factors at 2-levels.

Table 7: An 8 run geometric P–B design

A	B	C	D	E	F	G
+1	−1	−1	+1	−1	+1	+1
+1	+1	−1	−1	+1	−1	+1
+1	+1	+1	−1	−1	+1	−1
−1	+1	+1	+1	−1	−1	+1
+1	−1	+1	+1	+1	−1	−1
−1	+1	−1	+1	+1	+1	−1
−1	−1	+1	−1	+1	+1	+1
−1	−1	−1	−1	−1	−1	−1

The pattern for the first row (or column) determines the entire design. Each subsequent row (or column) is simply the previous row, say, shifted one step to the right, with the final symbol from the previous row being placed at the start of the next row. As such it is simply a cyclical arrangement of the first row (or column). The final row (in the example below) is set to all minus (-) [9].

Example:

We consider a plastic foam extrusion process. A process improvement team was formed to investigate what effects porosity of plastic parts. After a thorough brainstorming session with quality engineers, process manager and operators, it was identified that eight process parameters might have some impact on porosity. Table 8 presents the list of parameters and their levels for the experiment. Each factor was studied at 2-levels. As the total degrees of freedom for studying 8 factors at 2-levels is equal to 8, it was decided to choose a 12 run P–B design with 11 degrees of freedom. The extra 3 degrees of freedom can be used to estimate experimental error. Table 9 shows the experimental design with response [1, 4].

Table 8: List of process parameters and their levels for the experiment

Process parameters	Labels	Low level (−1)	High level (+1)
Temperature profile	A	1	2
Temperature after heating	B	210 °C	170 °C
Temperature after expansion	C	170 °C	150 °C
Temperature before coating die	D	130 °C	115 °C
Extrusion speed	E	6 m/min	4.5 m/min
Adhesive coating thickness	F	0.7 mm	0.4 mm
Adhesive coating temperature	G	115 °C	100 °C
Expansion angle	H	Max	Min

Table 9: Experimental Layout for 12 run P–B design with response values

Run	A	B	C	D	E	F	G	H	Porosity (%)
1 (6)	+1	+1	−1	+1	+1	+1	−1	−1	44.8
2 (11)	+1	−1	+1	+1	+1	−1	−1	−1	37.2
3 (9)	−1	+1	+1	+1	−1	−1	−1	+1	36.0
4 (7)	+1	+1	+1	−1	−1	−1	+1	−1	34.8
5 (2)	+1	+1	−1	−1	−1	+1	−1	+1	46.4
6 (1)	+1	−1	−1	−1	+1	−1	+1	+1	24.8
7 (5)	−1	−1	−1	+1	−1	+1	+1	−1	43.6
8 (12)	−1	−1	+1	−1	+1	+1	−1	+1	44.8
9 (3)	−1	+1	−1	+1	+1	−1	+1	+1	24.0
10 (8)	+1	−1	+1	+1	−1	+1	+1	+1	34.4
11 (4)	−1	+1	+1	−1	+1	+1	+1	−1	27.2
12 (10)	−1	−1	−1	−1	−1	−1	−1	−1	49.6

Note: Numbers in parentheses represent the random order of experimental runs or trials.

The objective of the experiment was to determine the key parameters which affect percentage porosity. Minitab software system is used for analysis purposes. Figure 13 illustrates a standardized Pareto plot of effects for the experiment drawn with Minitab software [1, 4].

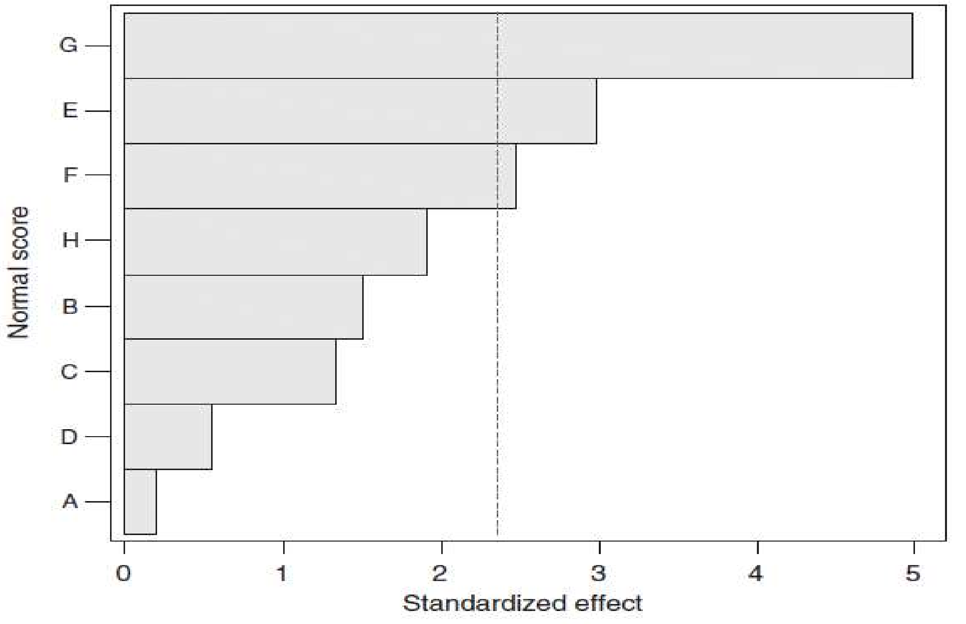


Figure 13: Standardized Pareto plot of effects for the above experiment [4]

Figure 13 shows that process parameters such as G (adhesive coating temperature). E (extrusion speed) and F (adhesive coating thickness) have significant impact on porosity.

These parameters should be further explored using full fractional designs and more advanced methods such as response surface methods. if necessary. In the next stage of experimentation, one should analyses the interactions among the parameters E. F and G. In order to identify what levels of these parameters yields minimum porosity. we may consider an effects plot (figure 14). Figure 14 shows that E at high level. F at low level and G at high level yields minimum porosity.

The figure shows that porosity decreases as temperature is kept at high level (100 °C). Similarly, porosity decreases as extrusion speed is kept at high level (4.5 m/min) and coating thickness at low level (0.7 mm).

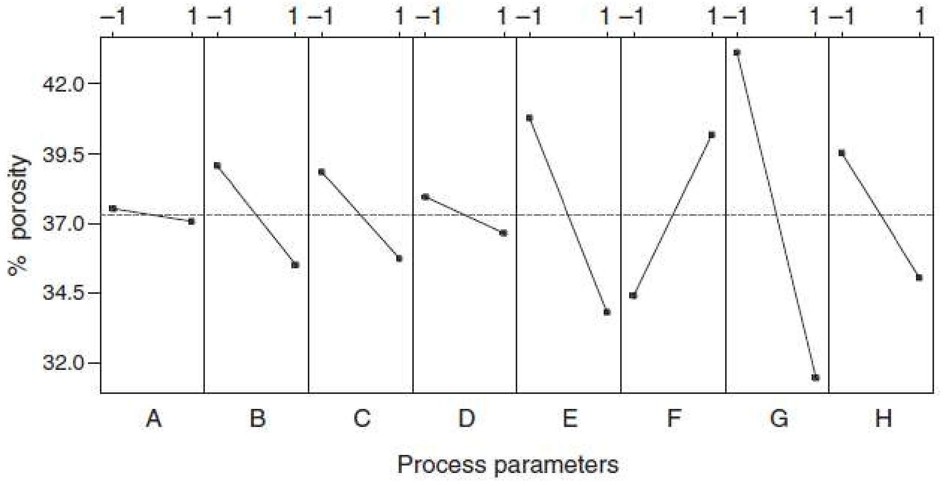


Figure 14: Main Effects plot for the experiment [4]

6.1.3. Mixture designs

A mixture design is a particular kind of DOE in which the factors are ingredients or components of a mixture. Mixtures are different from other types of experimental design because the proportions of the constituents must add up to 100%. Increasing the level of one constituent necessarily reduces the level of the others. The constraint that the proportions all of the ingredients must add up to 100% creates a unique design region that differs from classical design settings. The additional mixture contraint is as follow [1-4, 11] :

$$0 \leq x_i \leq 1 \; i = 1,2,\ldots,q \qquad \sum_{i=1}^{q} x_i = 1 \tag{14}$$

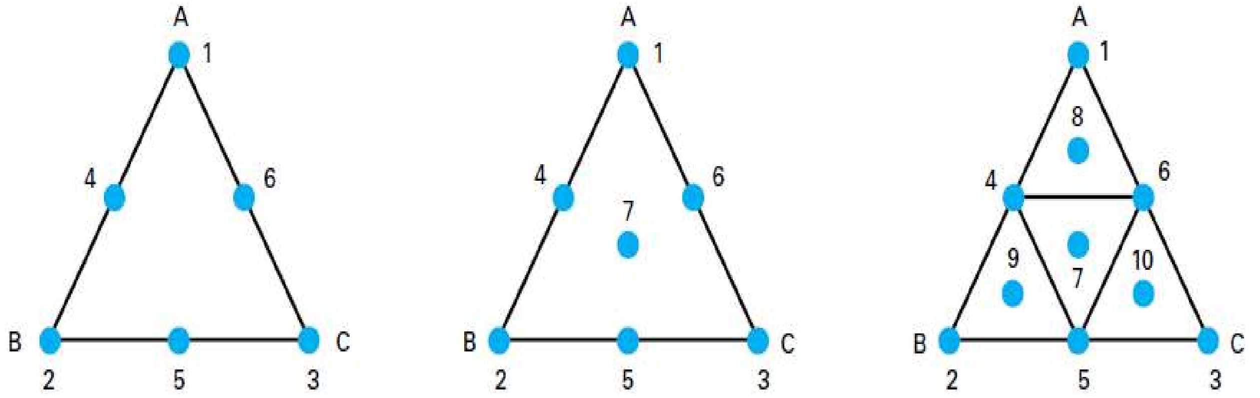
where q represents the number of ingredients in the system under study and the proportion of the i th ingredient in the mixture is denoted by xi.

where x_i represents the proportion of the ith component in the mixture.

- **Experimental points location**

A mixture with three constituents and when there are no constraints, the experiment points are distributed throughout the study domain (figure 15). Depending on the arrangement of these points, we distinguish several types of mixing designs:

1. Lattice designs.
2. Simplex centroid designs.
3. Augmented simplex centroid designs.



Lattice designs | Simplex centroid designs | Augmented simplex centroid designs

Figure 15 : Mixture designs types

- **Mixture design mathematical models**

For first order models with three component mixture. in a given point. the model can be written [1- 4, 9, 10]:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3$$

But the fundamental mixture constraint must be taken into account. The proportions x_i are not independent. Using the constraint. we know:

$$x_1 + x_2 + x_3 = 1$$

The equation above can then write:

$$\hat{y} = b_0(x_1 + x_2 + x_3) + b_1x_1 + b_2x_2 + b_3x_3$$

After regrouping the parameters. we obtain:

$$\hat{y} = (b_0 + b_1)x_1 + (b_0 + b_2)x_2 + (b_0 + b_3)x_3$$

This model has no constant and if we write:

$a_1 = (b_0 + b_1)$. $a_2 = (b_0 + b_2)$. $a_3 = (b_0 + b_3)$

The model takes the following form:

$$\hat{y} = a_1x_1 + a_2x_2 + a_3x_3$$

- **Second order Models**

The second order mathematical model for three component mixture has the following expression:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2$$

Taking into account the fundamental mixture constraint: $x_1 + x_2 + x_3 = 1$

which can be written in terms of x_1: $x_1 = 1 - x_2 - x_3$

multiplying each side by x_1 gives: $x_1^2 = x_1\,(1 - x_2 - x_3)$

$$x_1^2 = x_1 - x_1x_2 - x_1x_3$$

This shows that the squared term is in fact equal to a first-degree term and interaction term. So, the second-degree model therefore contains only first order and interaction terms, and can be written: $\hat{y} = a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$

6.2. Second order experimental designs (Response Surface Methodology)

Response Surface Methodology (RSM) is a compilation of statistical and mathematical techniques useful for modeling and analyzing problems. which predicts the response of interest influenced by several variables to optimize the product. RSM was first introduced by Box and Wilson in 1951 and now it is comprehensively used for different purposes in chemical and biological processes [2, 4, 8, 11].

In many instances. RSM uses the statistical experimental designs like Central Composite Design (CCD); which will be detailed in this course; and Box Behnken Design (BBD) to develop empirical models, which relate a response and mathematically depicts the relationships existing among the independent (inputs or causes. i.e. potential reasons for variation) and dependent variables (output or outcome. i.e. the values which result from the independent variables) of the process, the RSM provides contour plots (figure 16(a)) and three-dimensional graphs (3D) (figure 16 (b)) to visualize the shape of response surface.

In the analysis of data, it is desirable to provide both graphical and statistical analyses. Plots that illustrate the relative responses of the factor settings under study allow the experimenter to gain a feel for the practical implications of the statistical results and to communicate effectively the results of the experiment to others.

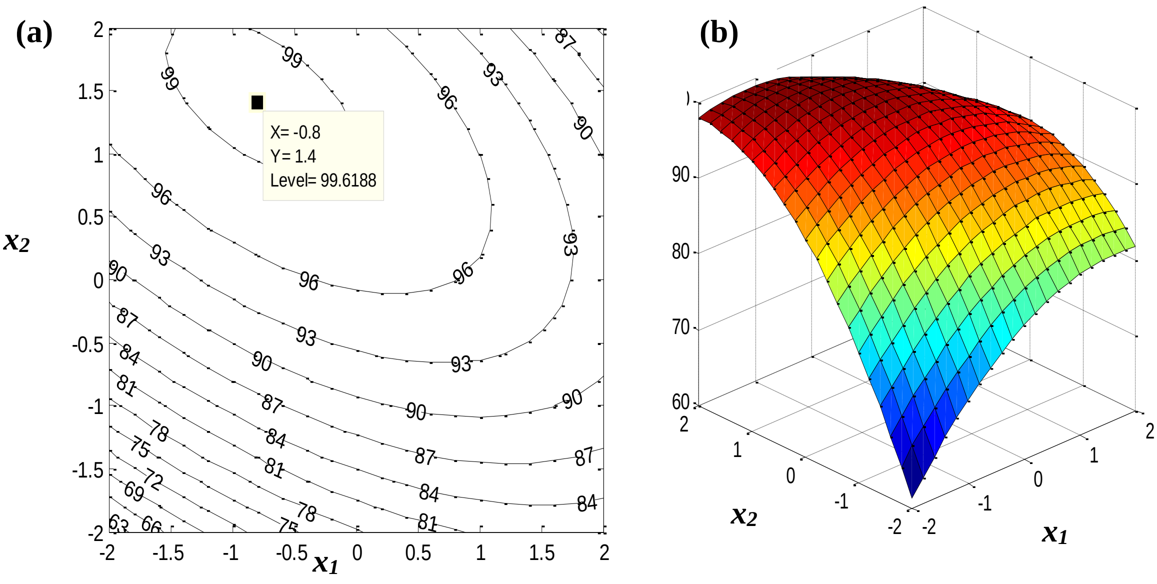


Figure 16: Example of Contour plot (a) and Response surface plot or 3D plot

6.2.1. Central Composite Design (CCD)

The two-level designs were useful to be applied in the screening study. However, the two-level designs lack information about maxima or any non-linear relationships since its application lead only on linear models. Performing a full factorial design with the level more than two will affect the effectiveness of the design itself due to the greater number of experiments that should be done. Hence, it is important to develop a design which allows greater level numbers without running every combination experiment. Presented by Box and Wilson (1951), the CCD becomes solution to overcome these problems [1, 4, 8, 10, 11].

Composite designs lend themselves well to a sequential study. The first part of the study is a full- or fractional-factorial design supplemented by center points to check the validity of the first-degree. factorial model. If the validation tests are positive (the response measures at the center of the field are statistically equal to the predicted value at the same point), the study is generally completed. If the tests are negative. supplementary trials are undertaken to establish a second-degree model. The additional trials are represented by the design points located on the axes of the coordinates and by new central points. The points located on the coordinate axes are called star points. Composite plans therefore have three parts (figure 17), in the present [5, 8-11]:

- **Factorial design**

This is a full- or fractional-factorial design with two levels per factor. The experimental points are at the corners of the study domain.

- **Star design**

The points of a star design are on the axes and are in general, all located at the same distance from the center of the study domain.

- **Center points**

There are usually center points, at the center of the study domain for both the factorial designs and the star designs.

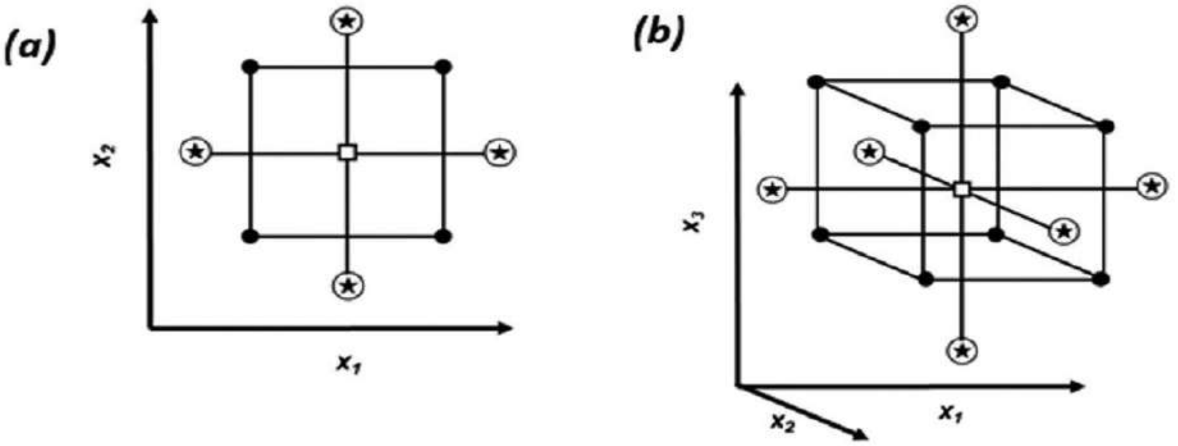


Figure 17: Illustration of central composite designs for **(a)** two factors and **(b)** three factors optimization. Every design consists of factorial points (●), star points (*), and central points (□).

To calculate the total number of trials. N, to carry out. sum the following:

- The trials from the star design (N_{α}= 2.k)
- The trials at the center (N_0)

So, the number of trials for a composite design is given by the equation.

$$N = N_f + N_{\alpha} + N_0$$

Star points are at some distance α from the center, based on the properties desired for the design and the number of factors in the design, the precise value of α are summarized in table 10 [4, 5, 9].

Table 10: Values of α and N_0 according to the properties of the composite design [1, 4, 9]

K	2	3	4	5 (2^5)	5 (2^{5-1})	6 (2^6)	6 (2^{6-1})
N_0 • Orthogonality	8	12	12	17	10	24	15
N_0 • Uniforme precison	5	6	7	10	6	15	9
N_0 • Rotatability	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1
α	1.41	1.68	2.00	2.00	2.38	2.83	2.38
N_f • Orthogonality	16	26	36	59	36	100	59
N_f • Uniforme precision	13	20	31	52	32	91	53

A CCD supports the building of a polynomial equation which takes into account the individual, interactive and quadratic terms and basically reads as follows [1, 4, 8, 11].

$$\hat{y} = b_0 + \sum_{j=1}^{k} b_j x_j + \sum_{\substack{u,\, j=1 \\ u \neq j}}^{k} b_{uj} x_u x_j + \sum_{j=1}^{k} b_{jj} x_j^2$$

Where $\sum_{j=1}^{k} b_j x_j$ is the individual effect of each factor; $\sum_{\substack{u,j=1 \\ u \notin j}}^{k} b_{uj} x_u x_j$ indicates the interactions amongst the variables; finally, the term $\sum_{j=1}^{k} b_{jj} x_j^2$ takes into account a possible non-linear/quadratic effect of some factors.

Table 11 and 12 and shows a typical scheme for a 2-factor and 3-factor CCD respectively.

Table 11: Calculation matrix of CCD for 2 factors and N_0= 4 replicates [1, 4]

Trial	x_0	x_1	x_2	$x_1 x_2$	x_1^2	x_2^2
1	1	-1	-1	1	1	1
2	1	1	-1	-1	1	1
3	1	-1	1	-1	1	1
4	1	1	1	1	1	11
5	1	-1.41	0	0	$(-1.41)^2$	0
6	1	1.41	0	0	$(1.41)^2$	0
7	1	0	-1.41	0	0	$(-1.41)^2$
8	1	0	+1.41	0	0	$(1.41)^2$
9	1	0	0	0	0	0
10	1	0	0	0	0	0
11	1	0	0	0	0	0
12	1	0	0	0	0	0

Table 12: Calculation matrix of CCD for 3 factors and N_0= 4 replicates [1, 4].

Trial	x_0	x_1	x_2	x_3	$x_1 x_2$	$x_1 x_3$	$x_2 . x_3$	x_1^2	x_2^2	x_3^2
1	1	-1	-1	-1	1	1	1	1	1	1
2	1	1	-1	-1	-1	-1	1	1	1	1
3	1	-1	1	-1	-1	1	-1	1	1	1
4	1	1	1	-1	1	-1	-1	1	1	1
5	1	-1	-1	1	1	-1	-1	1	1	1
6	1	1	-1	1	-1	1	-1	1	1	1
7	1	-1	1	1	-1	-1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1
9	1	**-1.68**	0	0	0	0	0	$\mathbf{(-1.68)^2}$	0	0
10	1	**+1.68**	0	0	0	0	0	$\mathbf{(+1.68)^2}$	0	0
11	1	0	**-1.68**	0	0	0	0	0	$\mathbf{(-1.68)^2}$	0
12	1	0	**+1.68**	0	0	0	0	0	$\mathbf{(+1.68)^2}$	0
13	1	0	0	**-1.68**	0	0	0	0	0	$\mathbf{(-1.68)^2}$

14	1	0	0	**+1.68**	0	0	0	0	0	**$(+1.68)^2$**
15	1	0	0	0	0	0	0	0	0	0
16	1	0	0	0	0	0	0	0	0	0
17	1	0	0	0	0	0	0	0	0	0
18	1	0	0	0	0	0	0	0	0	0

- **Coefficient estimation**

The following expression let us the coefficients estimation of CCD regression equation [1, 2, 4-6]. $B = \left[X^T.X\right]^{-1}.[X]^T.Y$

6.2.2. Box Behnken Design

Box-Behnken design (BBD) is a class of rotatable second-order response surface design based on three-level incomplete factorial design devised by Box and Behnken in 1960. This design was more efficient and economical than other three-level designs due to its ability to allow points selection from the three-level factorial arrangement BBD ensures that it does not contain combinations for which all factors are simultaneously at their highest or lowest levels. Besides, each factor requires only three levels instead of the five required for CCD, which may be experimentally more convenient and less expensive than CCD with the same number of factors but it is not suited for sequential experiments (figure 18) [5, 8].

The number of experiments (N) required for the development of BBD is defined as $N= 2k (k-1) + N_0$, where k is number of factors and N_0 is the number of central points (figure 18) [4, 5, 8, 11].

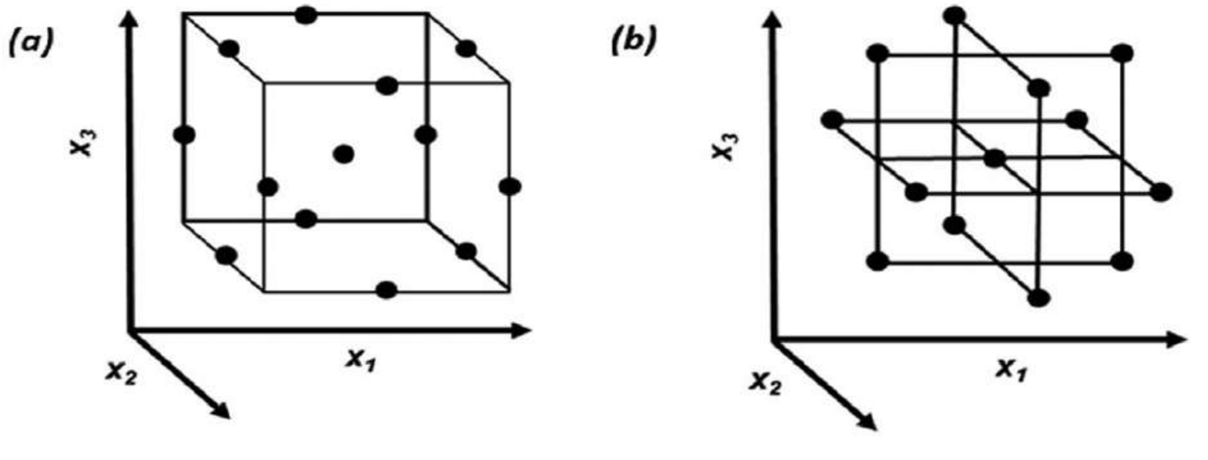


Figure 18. Illustration of the two graphical forms for the three factors BBD: (a) the cube for BBD and (b) three interlocking 2^2 factorial design [2, 11]

6.2.3. Doehlert design

Doehlert (1970) developed an alternative experimental design which has several advantages such as few experimental points, high efficiency and economically effective. Different from central composite and BBDs, these designs are not rotatable due to their number of estimations for varied factors. Nevertheless, Doehlert designs have different numbers of levels for different factors and allow to fill the provided factor space uniformly according to its possibility. Belonging to a second-order experimental design, Doehlert designs describe different characteristics for different levels: 1) a circular domain for two variables; 2) spherical domain for three variables; and 3) hyperspherical domain for four and more variables, which accents the uniformity of the studied variables in the experimental domain. Figure 19 illustrates the model of the Doehlert design [2, 5, 11].

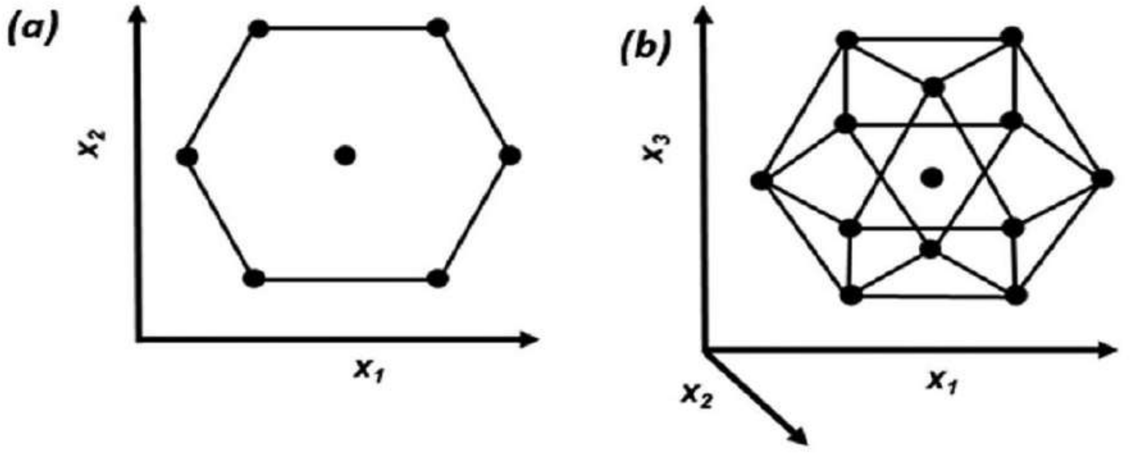


Figure 19: Illustration of the model of the Doehlert design for the optimization of (a) two variables and (b) three variables [2, 11]

7. Statistical analysis of the data

After the estimation of the factor regression coefficients and the first order or second order regression equation is developed, we have to test it validity and adequation.

7.1. Test of coefficients significance

To evaluate the importance of a coefficient. we apply statistical theory that compares the coefficient (b_j) with its standard deviation (S_{bj}) using the ratio $t_j = \frac{/b_j/}{S_{bj}}$. **(15)**

This ratio is called ***Student's*** *t* or the *t-ratio* [3-5, 9].

With:

$$S_{bj} = \frac{S_{rep}}{\sqrt{N}} \quad \textbf{(16)}$$

The variance of the measurements (or of reproducibility S_{rep}^2) is estimated generally by that calculated at the center of the experimental domain:

$$S_{\text{rep}}{}^2 = \frac{\sum_{i=1}^{n_0}(y_{i0} - \bar{y}_0)^2}{n_0 - 1} \quad \textbf{(17)}$$

$$\overline{y_0} = \frac{\sum_{i=1}^{n_0} y_{i0}}{n_0} \quad \textbf{(18)}$$

n_0: The trials number at the center;

y_{i0} : i th experimental response value;

$\overline{y_0}$: Mean of the trials replicated at the center domain;

Starting with the *t*-ratio. we can evaluate the probability that the coefficient is different from zero, or, said another way, if it is or is not significant using. either student table by reading the tabulated *t* (α. f=n_0-1) value to compare it to the calculated one. or by software of DOE, by the calculus of probability ***p-value***. If the *p*-value is close to zero (for biologists and chemists. *p-value* <0.05), the coefficient is influential and therefore is not equal to zero. If the ***p-value*** is close to one (*p-value* > 0.05), the coefficient cannot be distinguished from zero and is therefore not influential [1, 3, 4, 9].

Example:

We consider a chemical reaction yield which depends on two factors. temperature and pressure. The technician decides to carry out a first order experimental design without interactions with the following experimental domain (table 13) [1, 4]:

Table 13: Levels and factors values

	Low level : -1	High level :+1
Temperature : T	60°C	80°C
Pression : P	1 bar	2 bars

The response ***y*** studied; yield of the experiment; is given in table 14, for two factors we have to do 2^2= 4 trials.

Table 14: Design matrix and response

Trial	T	P	*y* (%)
1	-1	-1	60
2	+1	-1	65
3	-1	+1	75
4	+1	+1	85

The results of calculating the effects from the effect calculation matrix are given in table 15.

Table 15: Effect calculation matrix and the calculated effect

Trial	X_0	T	P	*y* (%)
1	+1	-1	-1	60
2	+1	+1	-1	65
3	+1	-1	+1	75
4	+1	+1	+ 1	85
Divider	4	4	4	
Affects	b_0 = 71.25	b_1 = 3.75	b_2 = 8.75	

The model equation is: $\hat{y} = 71.25 + 3.75x_1 + 8.75x_2$

To test the significance of the coefficients, we need to calculate the residual variance because of the absence of replicates ($S_{bj}^2 = S_{res}^2$) then the results of table 16.

Table 16. Residual calculation results

Trials	*y (%)*	*ŷ (%)*	e_i	e_i^2
1	60	58.75	1.25	1.5625
2	65	66.25	-1.25	1.5625
3	75	76.25	-1.25	1.5625
4	85	83.75	1.25	1.5625

$$e_i = y_i - \hat{y}_i$$

$$S_{bj}^2 = S_{res}^2 = \frac{1}{N-1} \sum e_i^2 = \frac{1}{4-3} \sum e_i^2 = 6.25$$

- Student test t consist to calculate:

$$t_j = \frac{/b_j/}{S_{bj}}$$

Bilateral Student test give for significance level α=0.05 and degree of freedom f= N-m =4-3 (N is the number of trials. m is the coefficient number in the model) (Appendix 2). The tabulated value t (0.05. 1) = 12.71

- For the effect b_1 = 3.75 of Temperature. we have t_1 = 3 < 12.71, we deduce that the effect of temperature T is not significant.
- For the effect b_2 = 8.75 of Pression. we have t_2 = 7 < 12.71, the effect of Pression P is not significant.

The conclusion of this study is that we must reject a linear model to explain the yield of this chemical reaction. It would be necessary to test study with a second-degree polynomial model.

7.2. Analysis of Variance (ANOVA)

The aim of applying the analysis of variance method is to answer the question: is the difference between the obtained response means for the tested factors a result of the influence of tested factors or has it occurred randomly.

Analysis of Variance (ANOVA) consists of finding the source of variation of the responses. Suppose that the responses have been calculated with a postulated model, by using the method of least squares [1, 4, 9]:

$$y_i = f(x_1, x_2, x_3, \dots\dots, x_n) + e_i \qquad \textbf{(19)}$$

In this case, the responses are written $\hat{y}_i$ and the errors as $\boldsymbol{e}$. These theoretical errors take particular values, written as r_i, and called residuals. The residuals are therefore particular values of the errors. We have:

$$\hat{y}_i = f(x_1, x_2, x_3, \dots\dots, x_n) \qquad \textbf{(20)}$$

With the new notation. the equation giving the response can be written as:

$$y_i = f(x_1, x_2, x_3, \dots \dots, x_n) + r_i \quad \textbf{(21)}$$

Classical analysis of variance uses not only the responses themselves but also the difference between the responses and their mean $(y_i - \bar{y})$ or $(\hat{y}_i - \bar{y})$. This difference is designated as "errors about the mean." In the case of calculated responses, we can also say "corrected for the mean" [1, 4].

In the case of the method of least squares. the mean of the observed responses is equal to the mean of the observed responses under the postulated model. Therefore. if $\bar{y}$ is the mean of the responses [1, 4].

$$y_i - \bar{y} = \hat{y}_i - \bar{y} + r_i \quad \textbf{(22)}$$

Squaring both side of the equation gives:

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum r_i^2 \quad \textbf{(23)}$$

This is the fundamental relation of analysis of variance. The left side is the sum of squares of the errors around the mean of the observed responses. This sum decomposes into two pieces: the sum of squares of the errors around the mean of the responses calculated with the model, and the sum of the squares of the residuals.

Suppose a polynomial regression model has been postulated for a given experiment, and the model assumptions appear to be satisfied, then it is appropriate to proceed with analysis of the data. The determination of significant factors affecting the dependent variables of interest (responses) is followed by **(ANOVA)** which uses tests based on variance ratios to determine whether or not significant differences exist among the means of several groups of observations. where each group follows a normal *the analysis of the variance* distribution [4, 9].

The analysis of variance is used very widely in the biological. social and physical sciences. The technique was first developed by R. A. Fisher and his colleagues in England in the 1920s [1, 4, 9].

7.2.1. Test of regression validation

In an F-test. the variance ratio between lack of fit (S^2_{res}) and pure experimental error (S^2_{rep}). is being compared to tabled values of F-distribution (Appendix 3). If the calculated F

exceeds the tabled *F*. then there is a significant **lack of fit** at the probability level that is chosen (usually *P*=0.05) and the model is incorrect [4-9].

$$S_{res}^2 = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N - l} \quad \textbf{(24)}$$

$$F = \frac{S_{res}^2}{S_{rep}^2} \quad \textbf{(25)}$$

l: the number of significant factors in the regression equation

7.2.2. Test of regression significance

In an analysis of variance. ANOVA. the total variation of the response is defined as a sum of two components; a regression variance (S_{reg}^2) and a component due to the residuals (S_{res}^2). The sum of squares of the total variation. corrected for the mean S_T^2. can thus be written as [1, 4, 9]:

$$S_T^2 = S_{reg}^2 + S_{res}^2 \quad \textbf{(26)}$$

$$S_{reg}^2 = \sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2 \,/\, l\text{-}1 \quad \textbf{(27)}$$

The regression component of the total variation is compared to the residual component. If the standard deviation of the response explained in the model regression is larger than the standard deviation of the residuals, then the model is significant at the chosen probability level (usually *P*=0.05) [4, 9].

$$F = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2 \,/\, \text{l-1}}{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \,/\, \text{N-l}} \quad \textbf{(28)}$$

$$\bar{y} = \frac{\sum_{i=1}^{N} y_i}{N} \quad \textbf{(29)}$$

The F distribution has df (degree of freedom) numerator = l– 1 degrees of freedom and df denominator = N– l denominator degrees of freedom. At a significance level of α = 0.05, since F calculated > $F_{0.05}$ tabulated. we must conclude that there are no significant differences between the two variances and the regression is valid [1, 4, 6- 9].

Software can construct ANOVA tables. The simplest of these tables has five columns (source of variation. sum of squares, degrees of freedom (df), mean square, and *F*-ratio) and four lines (column titles. model corrected for the mean. residuals and observed responses corrected for the mean) similar to Table 17 [1, 4, 6, 9].

- The first column shows the sources of variation.
- The second column shows the df (degree of freedom) of each sum of squares.
- The third column gives the sums of squares of the errors around the mean.
- The mean squares of the fourth column are the sums of squares divided by their df.
- The fifth column shows the *F*-ratio, which is the ratio of the mean square of the model to the mean square of the residuals.

F-ration allows the calculation of the probability that the two mean squares are not equal. In other words, if the *F*-ratio is high (small probability that the model is only due to the effect of the mean). the variations of the observed responses are likely due to variations in the factors. If the *F*-ratio is near 1 (strong probability that the model is not due to the effects). the variations of the observed responses are comparable to those of the residuals. The ***p*-value** corresponding to the *F*-ratio is also shown [1-4, 8- 10].

Table 17: Analysis of variance (ANOVA) table [1, 4]

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	10	965.30000	96.5300	46.3196
Error	5	10.42000	2.0840	**Prob > F**
C. Total	15	975.72000		0.0003*

7.2.3 The coefficient of determination R^2

The analysis of variance allows the calculation of a very useful statistic: R^2. This statistic is the ratio of the sum of squares of the predicted responses (corrected for the mean) to the sum of squares of the observed responses (also corrected for the mean) [1-4, 6, 9]:

$$R^2 = \frac{\text{Sum of squares (Model)}}{\text{Sum of Squares (Total)}}$$

$$R^2 = \frac{\sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2}{\sum^{N} (y_i - \bar{y})^2} \tag{30}$$

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 $\overline{R^2}$ is introduced,which will not always increase with adding a new variable [1-4, 8-10]:

$$\overline{R}^2 = R^2 - (1 - R^2)\frac{\ell - 1}{N - \ell} \tag{31}$$

8. Optimization

Response surfaces are used to determine an optimum. In addition, it is a good way to graphically illustrate the relation between different experimental variables and the responses. To be able to determine an optimum it is necessary that the polynomial function contains quadratic terms, this is done by deriving the predicted response model (y) with respect to all variables or by the contour and surface response plots.

- **Example**

The results of modeling of antimicrobial production by a strain using Central Composite Design (CCD) for four operating factors (x_1, x_2, x_3, x_4) allowed to obtain the following regression equation after its validation [17]:

$$\hat{y} = 12.53 + 1.22x_2 + 1.46x_1x_3 + 1.31x_2x_3 - 1.32x_2x_4 - 0.94x_3^2$$

x_1 : KCl concentration (g/l) ;

x_2: K_2HPO_4 concentration (g/l) ;

x_3 : $MgSO_4$ concentration (g/l) ;

x_4 : Incubation time (days).

From this model. we are now able to calculate the optimal values of the operating parameters leading to maximum antibacterial activity against a target germ (*Salmonella typhi*).

to do this. simply solve the system of equations below. obtained by deriving the predictive model with respect to each of the variables x_1, x_2, x_3, x_4:

$\frac{d\hat{y}}{dx_1} = 0.$ $1.46\, x_3 = 0 \rightarrow x_3 = 0$

$\frac{d\hat{y}}{dx_2} = 0.$ $1.22 + 1.31\, x_3 - 1.32\, x_4 = 0 \rightarrow x_4 = 0.92$

$\frac{d\hat{y}}{dx_3} = 0.$ $1.46\, x_1 + 1.31\, x_2 - 1.88\, x_3 = 0 \rightarrow x_1 = 0$

$\frac{d\hat{y}}{dx_4} = 0.$ $-1.32\, x_2 = 0 \rightarrow x_2 = 0$

The resolution of this equations system gives:

- $x_1 = 0$. corresponding to KCl concentration of 0.5 g/L.
- $x_2 = 0$. corresponding to K_2HPO_4 concentration of 1 g/L.
- $x_3 = 0$. corresponding to $MgSO_4 . 7H_2O$ concentration of 0.5 g/L.
- $x_4 = 0.92$ corresponding to incubation time of 9 days.

The optimal antibacterial activity obtained by replacing the optimal values in the postulated model is **12.53 mm**.

The corresponding 2D (figure a) and 3D (figure b) dimensional response surfaces of the quadratic models are shown in figures (a, b) bellow. The figures are drawn in KCl (x_1) $MgSO_4 \cdot 7H_2O$ concentration (x_3) plan (the most important interaction) using MATLAB 7.0 software. The analysis of these figures shows that in the presence of a moderate KCl concentration the antimicrobial activity increases with reduced $MgSO_4 \cdot 7H_2O$ concentration. The maximum predicted yield is indicated by the surface confined in the smallest curve of the contour diagram which is equal to 14.79 mm, corresponding to an economic condition to that obtained above [17]:

- [KCl] =0.1 $g.L^{-1}$ ($x_1 = -2$),
- [K_2HPO_4] =1 $g.L^{-1}$ ($x_2 = 0$),
- [$MgSO_4 \cdot 7H_2O$] =0.2 $g.L^{-1}$, ($x_3 = -1.5$),
- Incubation time equal to 9 days ($x_4 = 0.92$).

R^2 was found to be 90,7 % , indicating that the models can explain 90.7% of total variations.

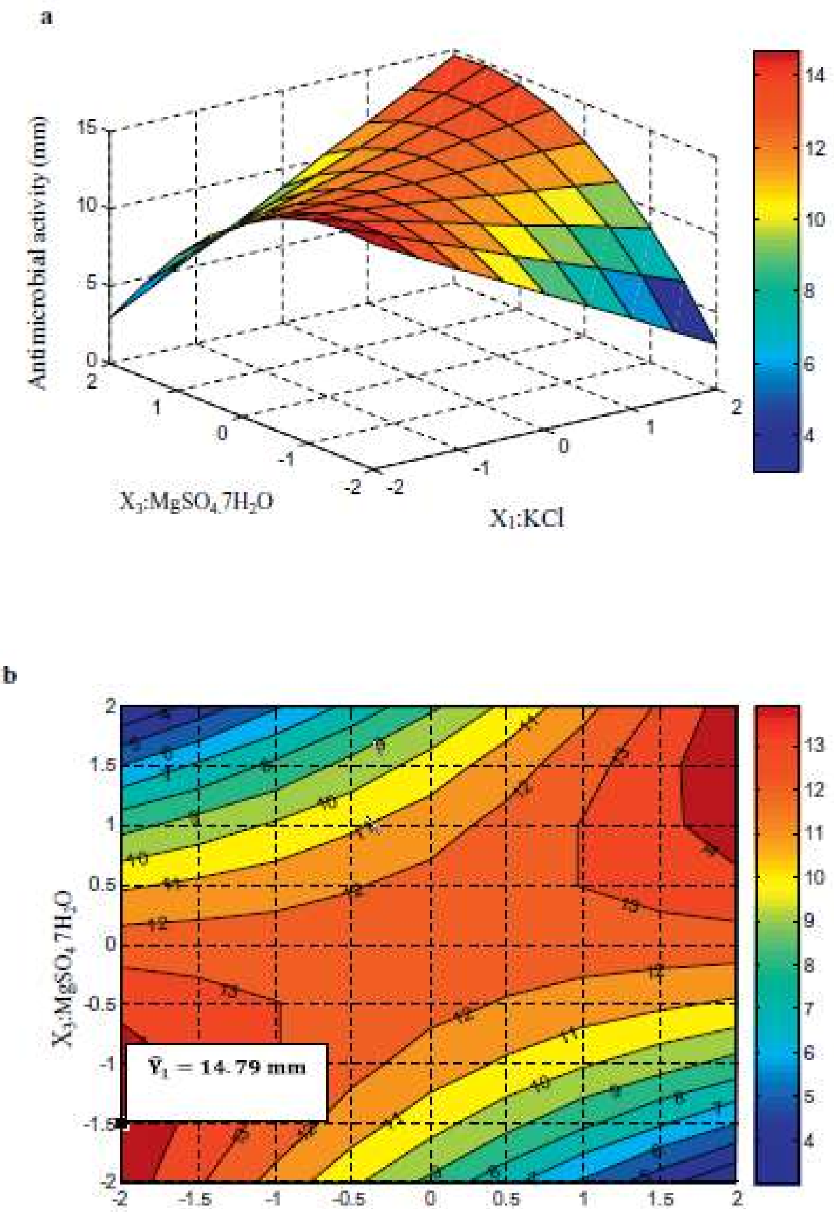


Figure (a, b): Response surface and isoresponse plots

9. 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. JMP. Design-Expert. Statgraphics. Prisma. etc. The most popular commercial packages Minitab and Statistica are equipped with user friendly interface and very good graphics output [1, 4, 7, 9].

Also, DOE design and analysis can be done easily in Microsoft Excel, using the procedure and formulas described above in the present course.

10. Demonstration using Minitab software

The MINITAB program interface is designed to be very simple and easy to use, in addition to the tools required to design and analyze experiments, this software supports most of the other statistical analyses and methods that most users need. MINITAB has a powerful graphics engine with an easy to use interface. Most graph attributes are easy to configure and can be edited after a graph is created [4, 12-16].

All the manipulation steps described below using Minitab software in this course will be applied in the form of practical work to students using a microcomputer.

10.1. Starting Minitab [12-16]

To open Minitab follow the following instructions:

- Double-click the Minitab icon (Figure 20).

Figure 20 : Minitab icon

We obtain a worksheet (figure 21) consisting of:

- Menu bar (There are 11 menu headings: ***File. Edit. Data. Calc. Stat. Graph. Editor. Tools. Window. Help and Wizard)***
- The icon bar
- "Session" window
- Worksheet similar to that of the Excel

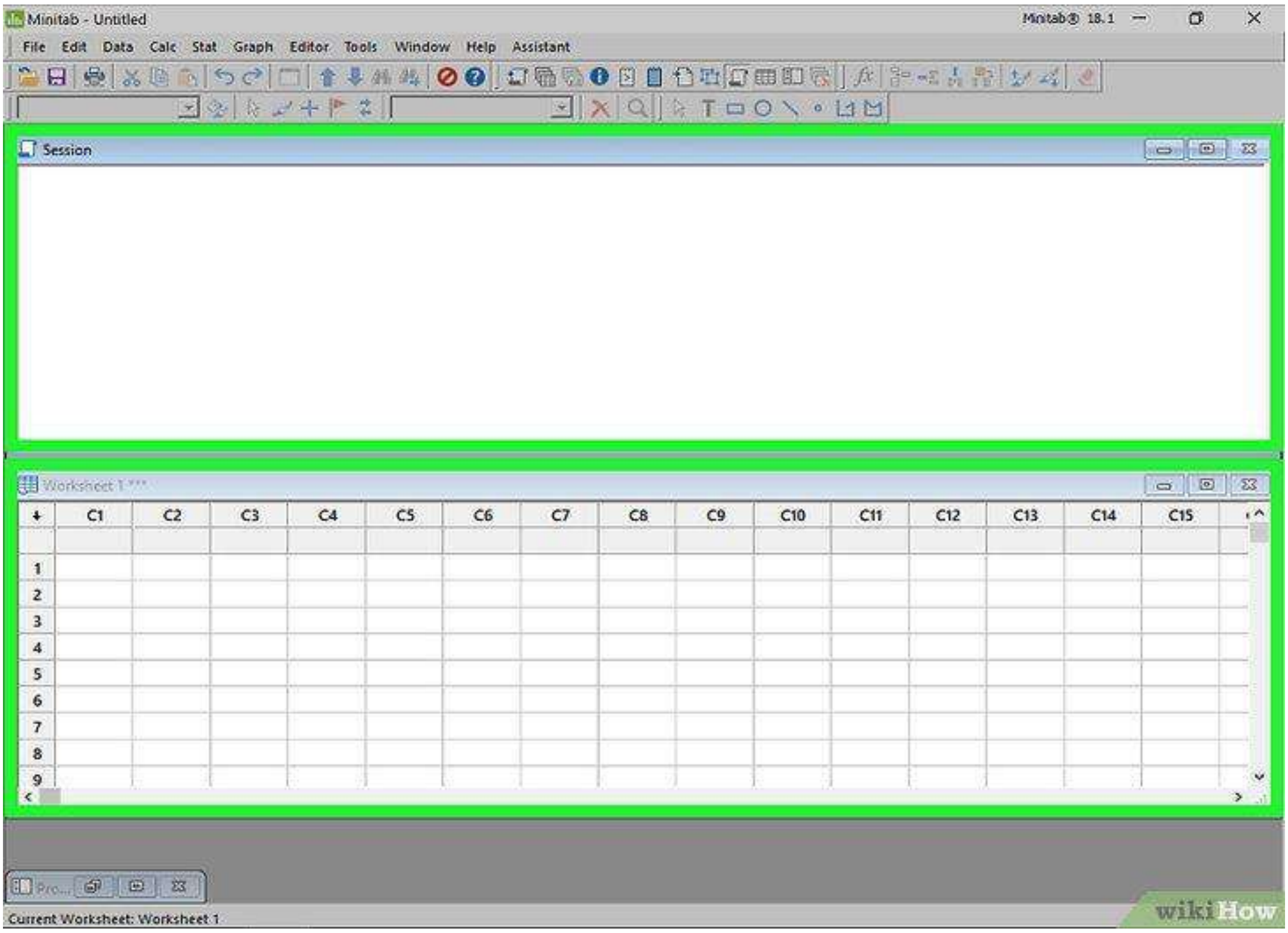


Figure 21: Minitab home window (worksheet)

10.2. Access to design of experiments

- Click on ***Stat*** from the main menu. A drop-down menu appears with ***DOE*** (Design of Experiments).
- Click on ***DOE*** (Design of Experiments). A new menu appears (Figure 22) in which there are four choices:
  - **Factorial design**
  - **Response surface**
  - **Mixture**
  - **Taguchi**

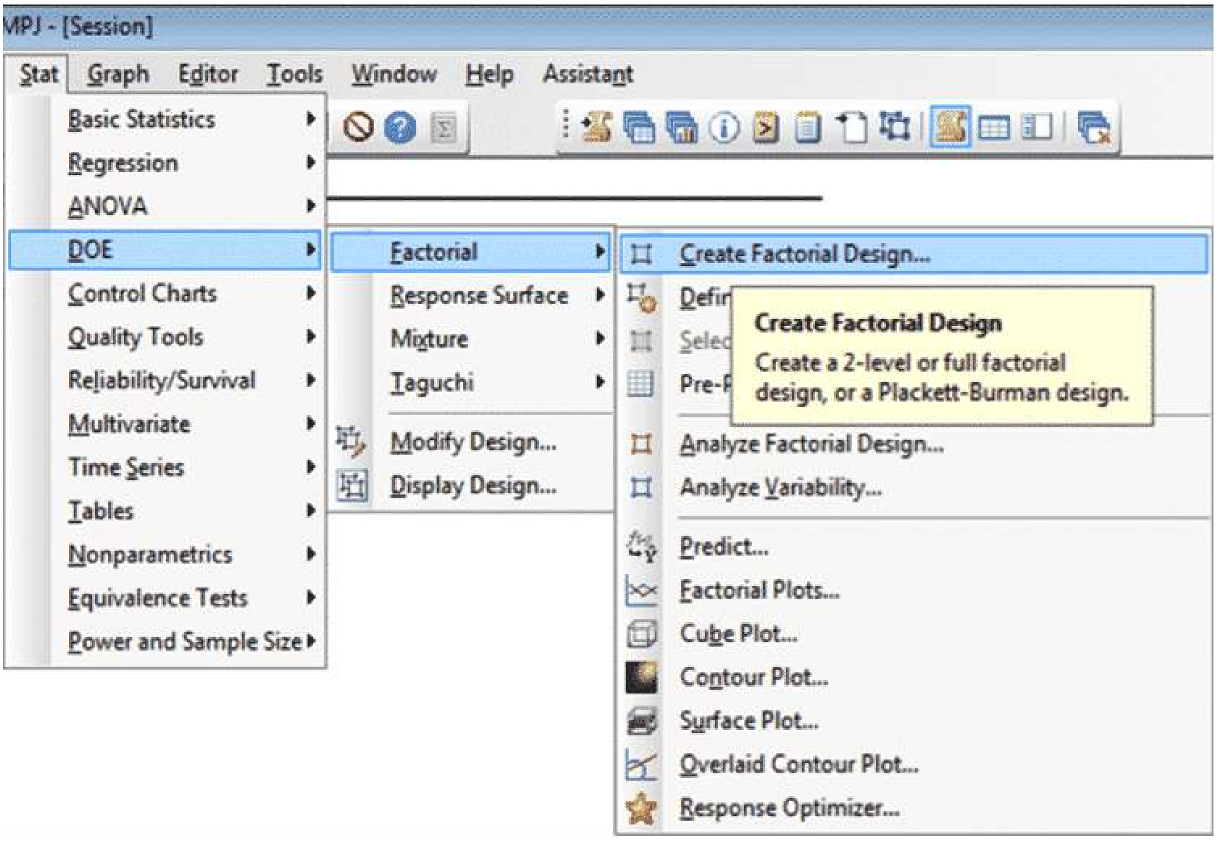


Figure 22: Menus giving access to design of experiments

10.3 Design of experiments definition

- **Global definition**

For all the designs. it is necessary to define the responses, the factors, the levels. The procedures for entering this data are practically the same for all designs (figure 22).

- Click Create ***factorial design*** for example. a window appears in which a choice of designs is offered (Figure 23).

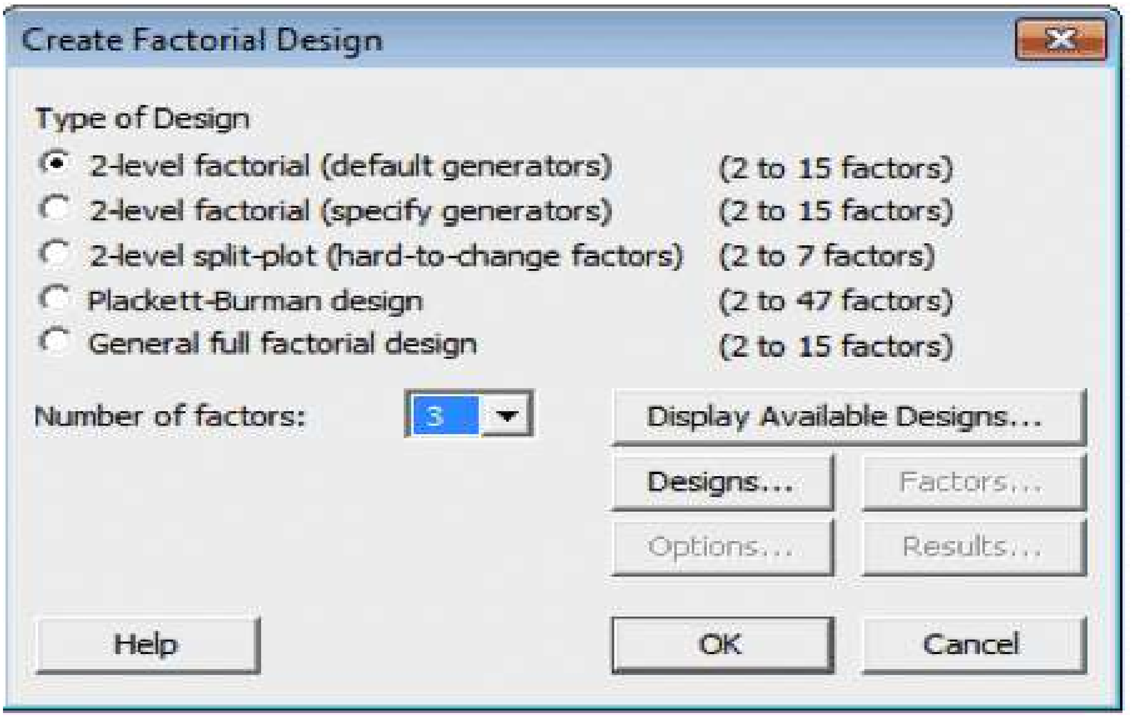


Figure 23: Window offering several types of experimental designs

- **Factors number**
  - ➢ Click on the ***Designs*** button of figure 23.
  - ➢ Click the ***OK*** button when you have finished your choice (figure 24).

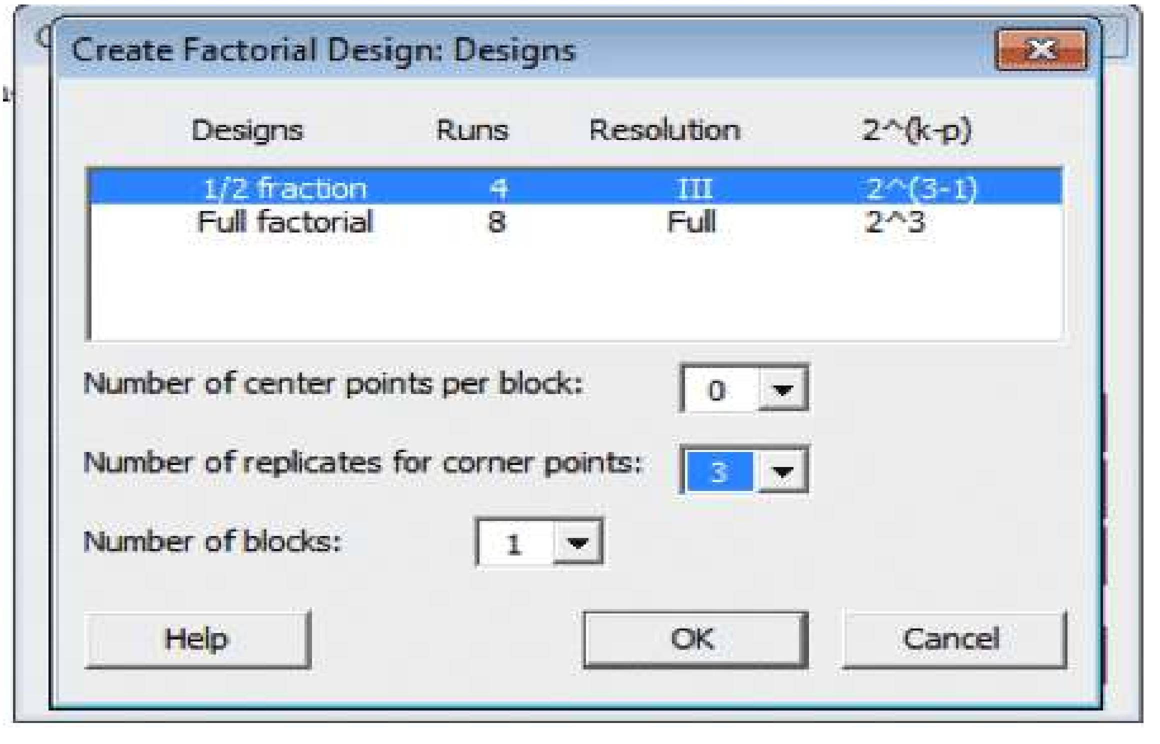


Figure 24: Window allowing you to choose a full or fractional factorial design

You then return to the ***Create a factorial design*** window (Figure 22) in which the three buttons previously grayed out are now accessible: Factors, Options, Results.

10.4 Factors definition

- Click on the ***Factors*** button of figure 23. you obtain figure 25 in which factors appear with their name (example: Pressure, Temperature, time), their type and their levels.

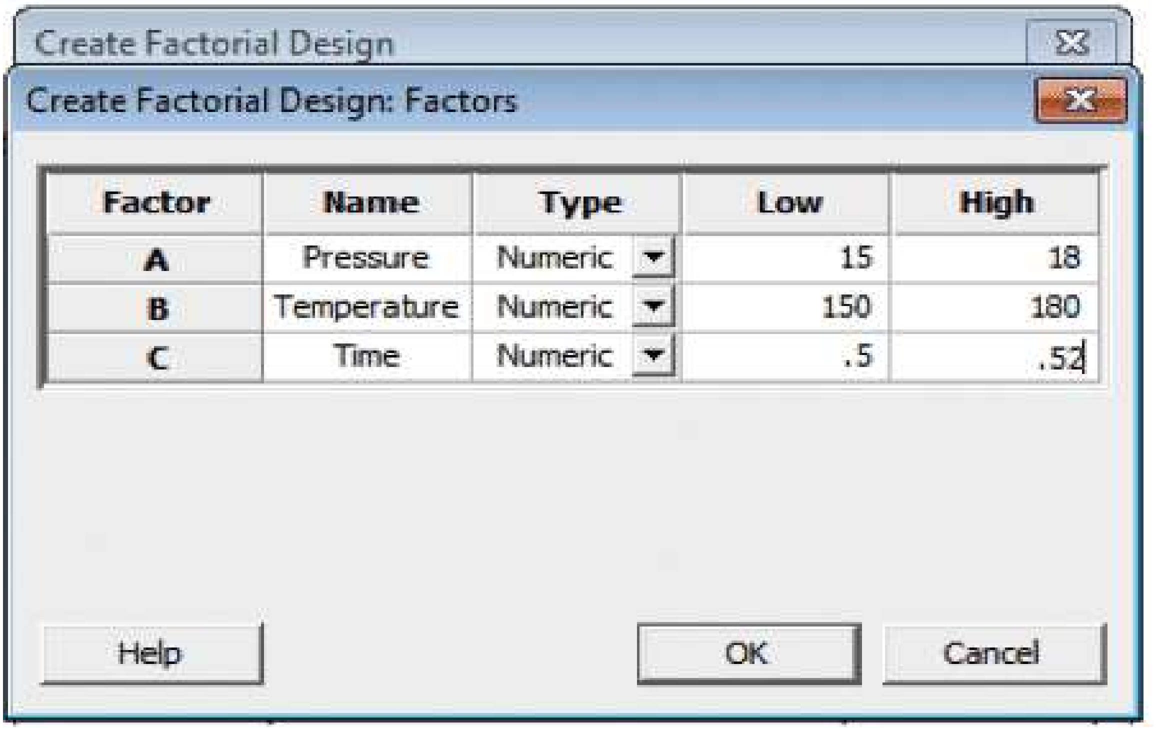


Figure 25: Window allowing to define the characteristics of the factors

10.5. Options definition

- Click on the ***Options*** button. you get the window in figure 26 in which you can further model your plan. replicate it or not. You can randomize the trials or not and request that the analysis results be stored in the worksheet.

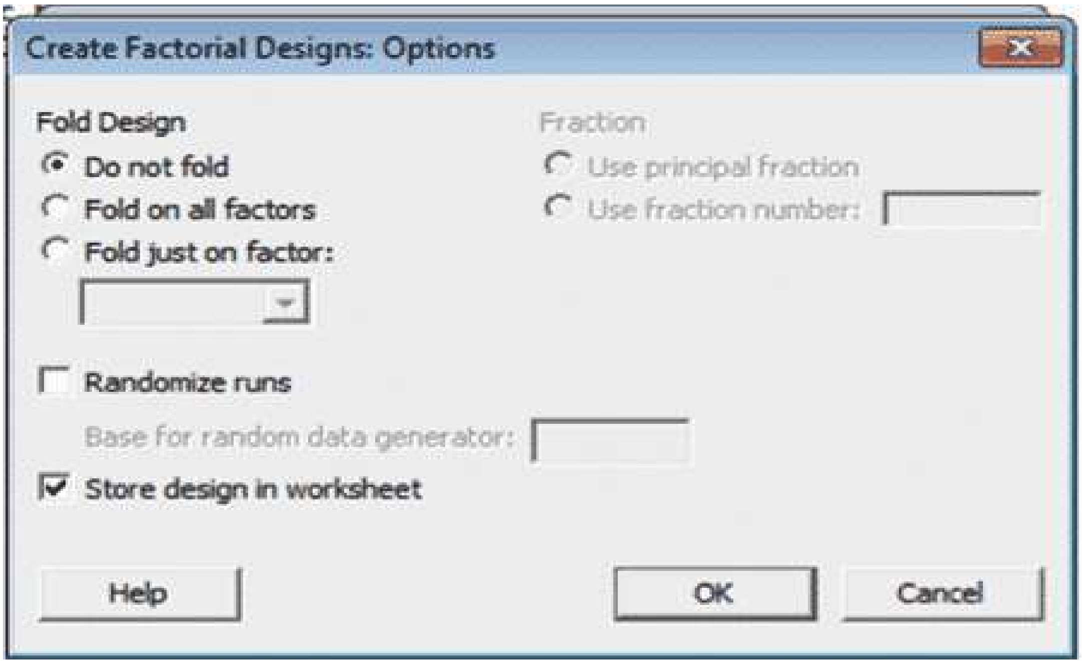


Figure 26: Window allowing to specify the characteristics of the experimental design

- Minitab will create a worksheet containing the DOE array (figure 27):

C1	C2	C3	C4	C5	C6	C7
StdOrder	RunOrder	CenterPt	Blocks	Pressure	Temperature	Time
1	1	1	1	15	150	0.52
2	2	1	1	18	150	0.50
3	3	1	1	15	180	0.50
4	4	1	1	18	180	0.52
5	5	1	1	15	150	0.52
6	6	1	1	18	150	0.50
7	7	1	1	15	180	0.50
8	8	1	1	18	180	0.52
9	9	1	1	15	150	0.52
10	10	1	1	18	150	0.50
11	11	1	1	15	180	0.50
12	12	1	1	18	180	0.52

Figure 27: Worksheet of the design created

- The first blank column in the worksheet (here C8) is reserved for the ***Response*** values

After running all of the experimental runs enter the results in to the worksheet (figure 28):

StdOrder	RunOrder	CenterPt	Blocks	Pressure	Temperature	Time	Average Fill
2	2	1	1	18	150	0.50	12.536
3	3	1	1	15	180	0.50	11.195
4	4	1	1	18	180	0.52	13.001
5	5	1	1	15	150	0.52	11.420
6	6	1	1	18	150	0.50	12.851
7	7	1	1	15	180	0.50	10.999
8	8	1	1	18	180	0.52	12.799
9	9	1	1	15	150	0.52	11.420
10	10	1	1	18	150	0.50	12.948
11	11	1	1	15	180	0.50	10.999
12	12	1	1	18	180	0.52	12.896

Figure 28: Worksheet of the design and responses

The second series of steps allow us to analyze the results as well as produce the charts and graphs that help us communicate our results (figure 29).

➢ Go to ***Stat*** > **DOE** > ***Factorial*** > ***Analyze Factorial Design***

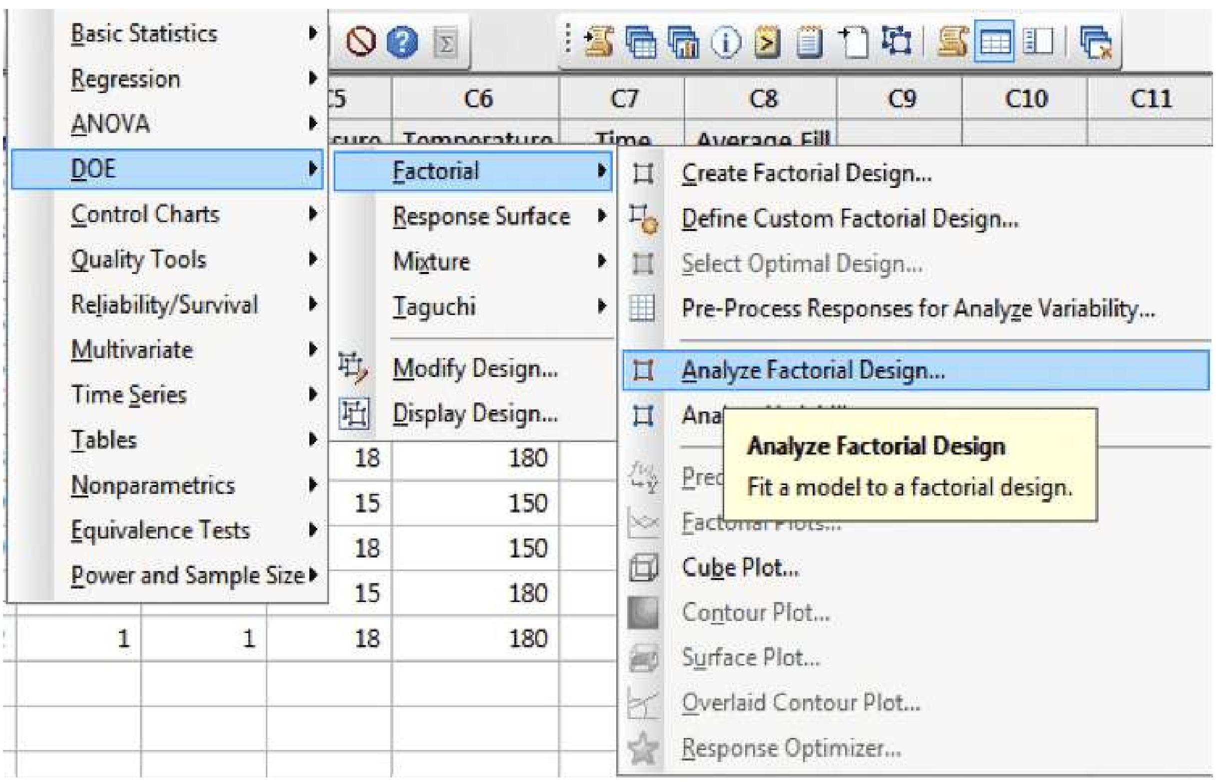


Figure 29: Window allowing the analysis of the design

- Enter the column (here C8) that contains the response in the open window called Responses (or just double-click on C8 in the left box) (figure 30).
- Then click on ***Terms.***

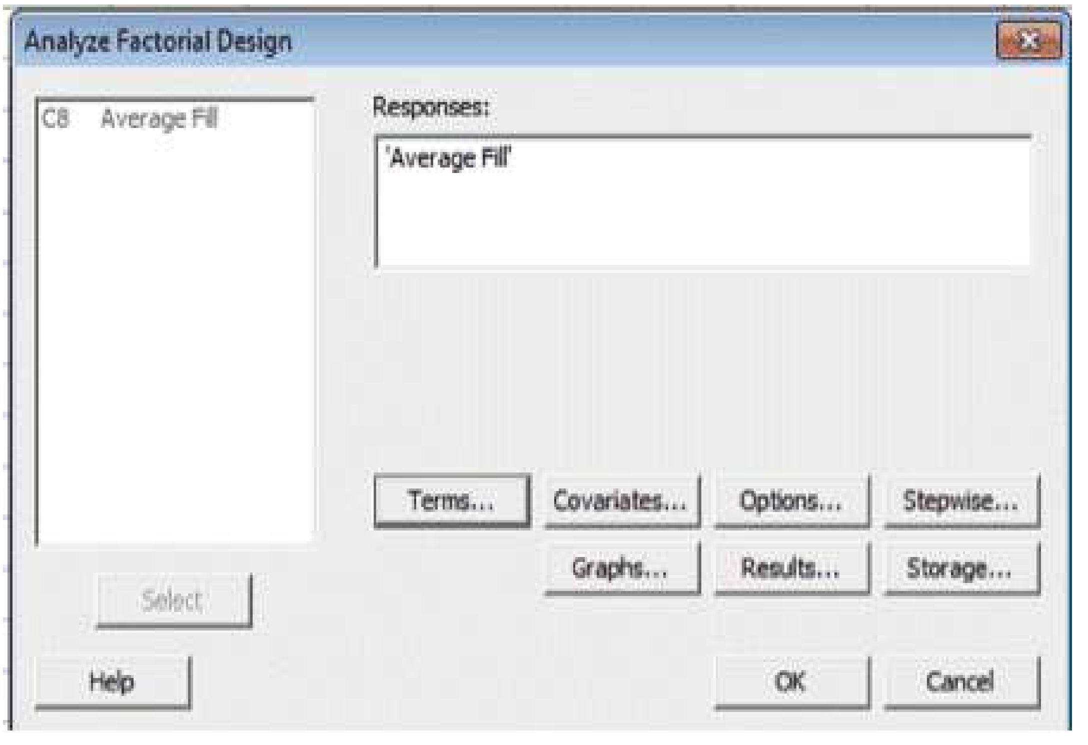


Figure 30: Window allowing response column selection to analysis

- Select the terms you want in the model (figure 31).
- Either double click on the ***term.***
- Then click OK.

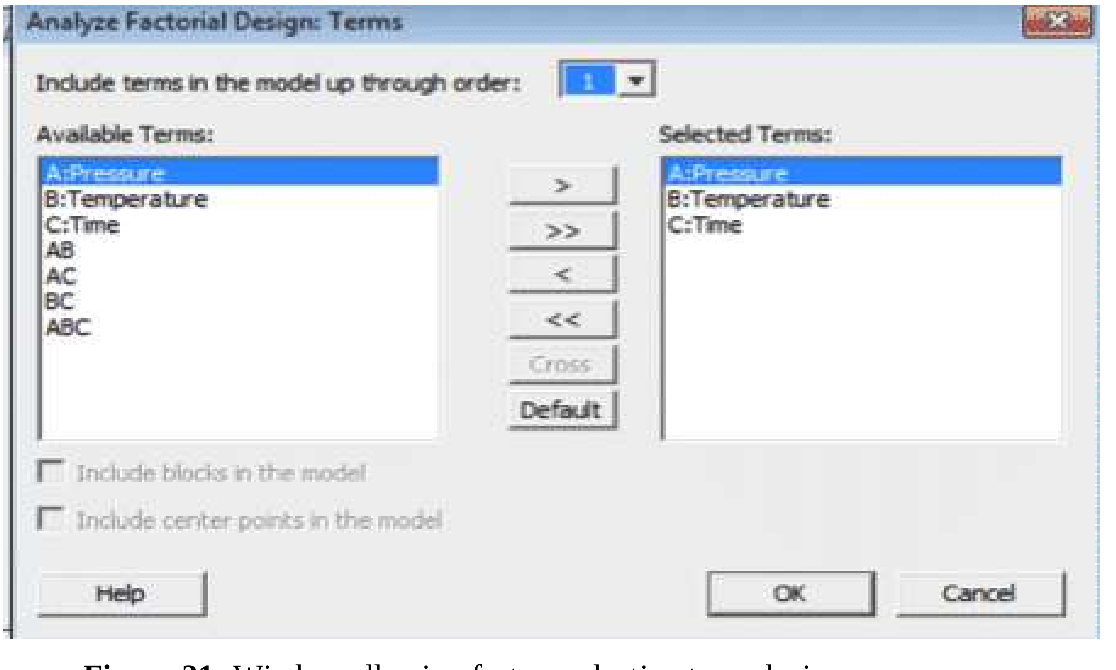


Figure 31: Window allowing factors selection to analysis

The details behind the analysis will be contained in the Minitab Worksheet

- This is the ANOVA table for the experiment:

Results for: DOE 3 factor.MTW

Factorial Regression: Average Fill versus Pressure, Temperature, Time

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	3	7.73886	2.57962	141.68	0.000
Linear	3	7.73886	2.57962	141.68	0.000
Pressure	1	7.49009	7.49009	411.37	0.000
Temperature	1	0.05410	0.05410	2.97	0.123
Time	1	0.19467	0.19467	10.69	0.011
Error	8	0.14566	0.01821		
Total	11	7.88452			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.134935	98.15%	97.46%	95.84%

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		12.0486	0.0390	309.32	0.000	
Pressure	1.5801	0.7900	0.0390	20.28	0.000	1.00
Temperature	-0.1343	-0.0671	0.0390	-1.72	0.123	1.00
Time	0.2547	0.1274	0.0390	3.27	0.011	1.00

Regression Equation in Uncoded Units

Average Fill = -2.40 + 0.5267 Pressure - 0.00448 Temperature + 12.74 Time

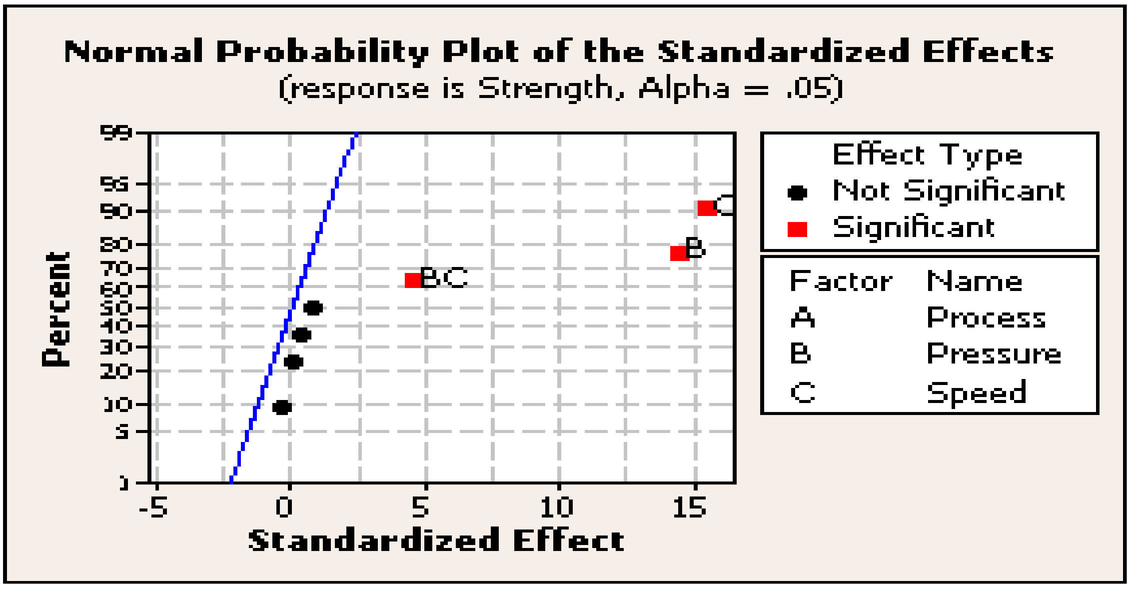
10.5. Effects plots in Minitab

The primary goal of screening designs is to identify the "vital" few factors or key variables that influence the response. Minitab provides two graphs that help you identify these influential factors: a normal plot and a Pareto chart. These graphs allow you to compare the relative magnitude of the effects and evaluate their statistical significance [12-16].

- **Normal Probability Plot of the Effects**

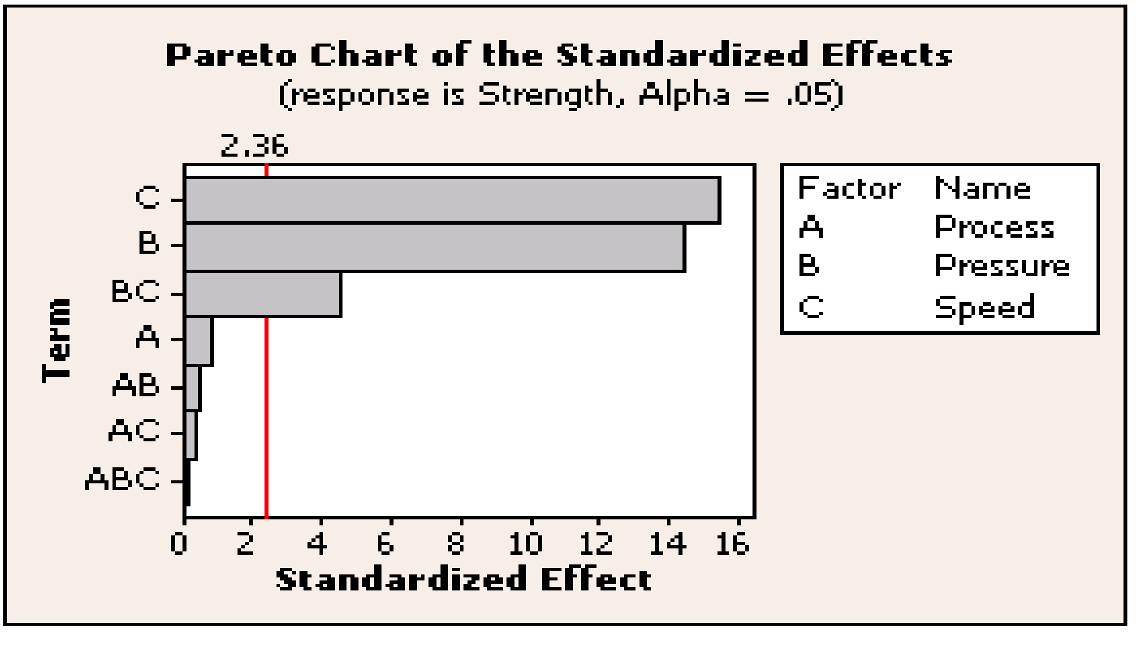
In the normal probability plot of the effects. points that do not fall near the line usually signal important effects. Important effects are larger and further from the fitted line than unimportant effects. Unimportant effects tend to be smaller and centered around zero. This plot shows that terms B, C, and BC are significant [12-16].

- Choose ***Stat*** > ***DOE*** > ***Factorial*** > ***Analyze Factorial Design***.
- Click ***Graphs***. Under **Effects Plots**. check Click ***OK*** in each dialog box.



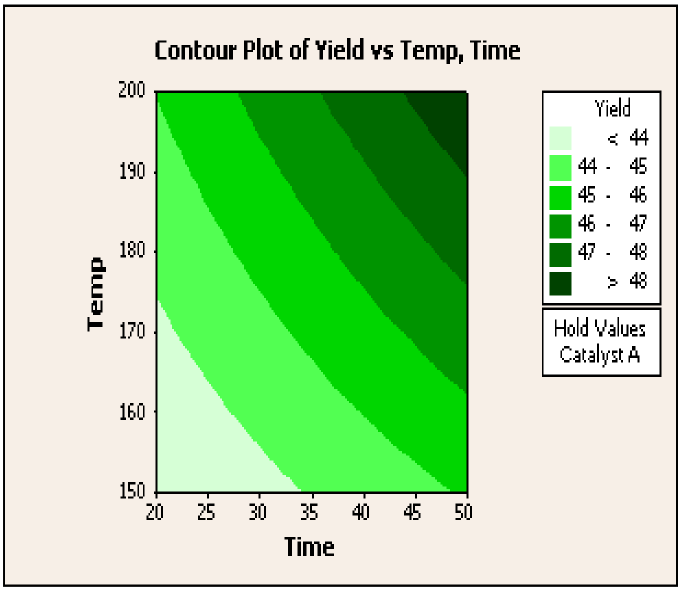
➢ **Pareto Chart of the Effects**

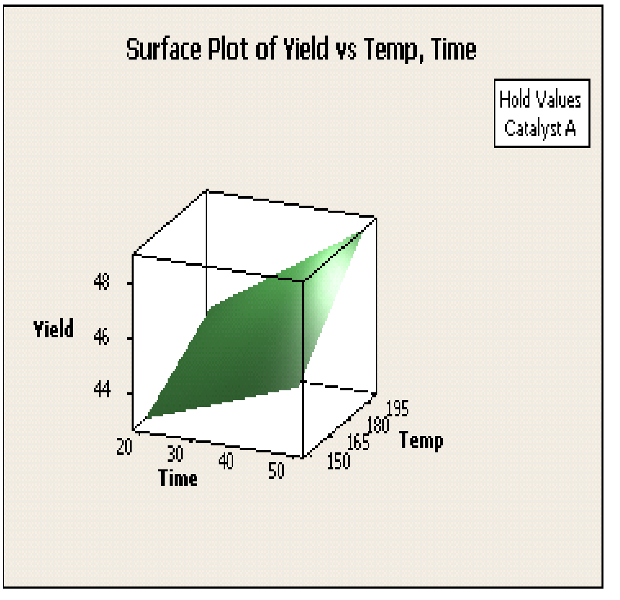
Use a Pareto chart of the effects to determine the magnitude and the importance of an effect. The chart displays the absolute value of the effects and draws a reference line on the chart. Any effect that extends past this reference line is potentially important [12-16].



➢ **Contour plot and a surface plot**

- Choose **Stat > DOE > Factorial > Contour/Surface Plots**.
- Check ***Contour plot*** and click **Setup**. Click ***OK***.
- Check ***Surface plot*** and click **Setup**. Click ***OK*** in each dialog box.





11. Applications using Minitab 16 software package

Example 1: 2^2 factorial design [4, 12-16]

A motorist wants to know the gas consumption of his car when he drives with or without extra weight. while driving fast or slowly. He decides to carry out a complete factorial design 2^2 to study the influence of two factors. their values and levels are given in table 18.

The response is the car's fuel consumption (mph).

Table 18: Factors and study domain

Factor	Low Level (–)	High Level (+)
Speed (1)	45 mph	70 mph
Additional weight (2)	0	550 lbs

- **DOE construction**

As this is a 2^2 design. the design creation windows are filled like Figure 22 to obtain figure 32 where you choose à number of factors (2 factors).

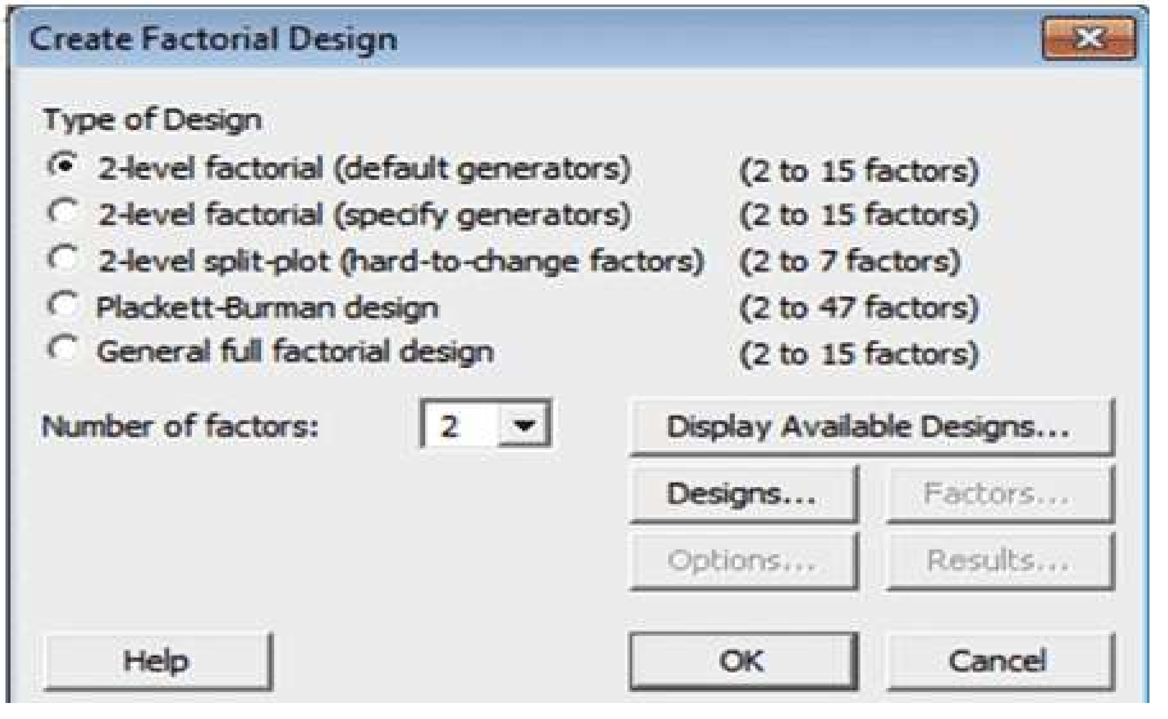


Figure 32: Window offering several types of experimental designs

➢ Click the ***OK*** button in Figure 32. you get the window in Figure 33.

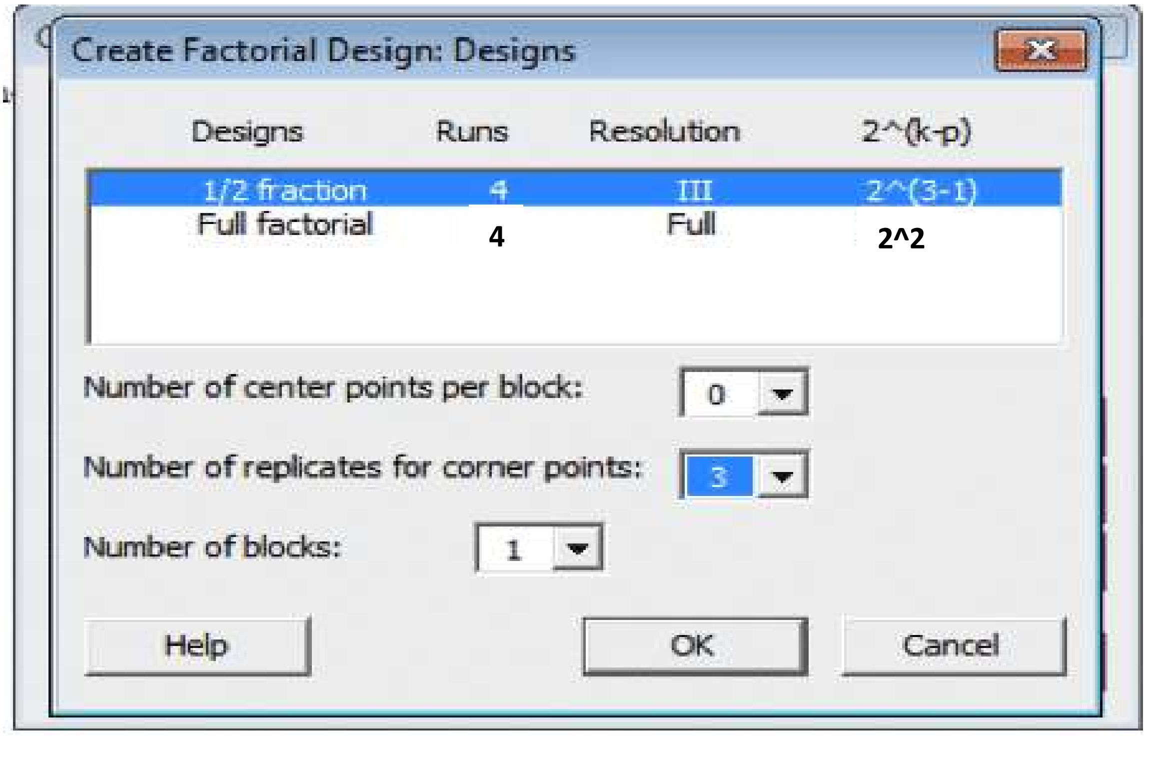


Figure 33: Window allowing the choice a full factorial design

- **Factors definition**

➢ Click the ***Factors*** button in Figure 23 and modify the factor names to obtain a window similar to figure 34.

➢ Click the ***OK*** button.

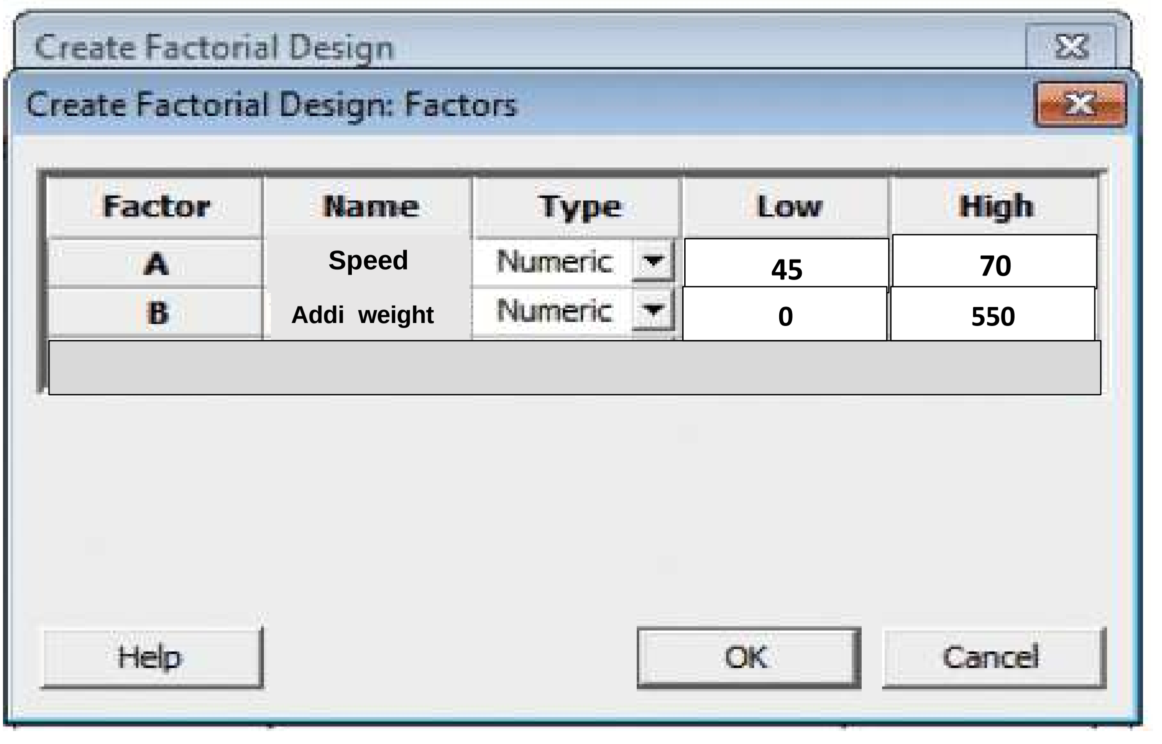


Figure 34: Window allowing to specify the factors

➢ Click on the ***Options*** button in Figure 23. you get the window in Figure 35.

➢ Click the OK button in the Create Factorial Design-Options window.

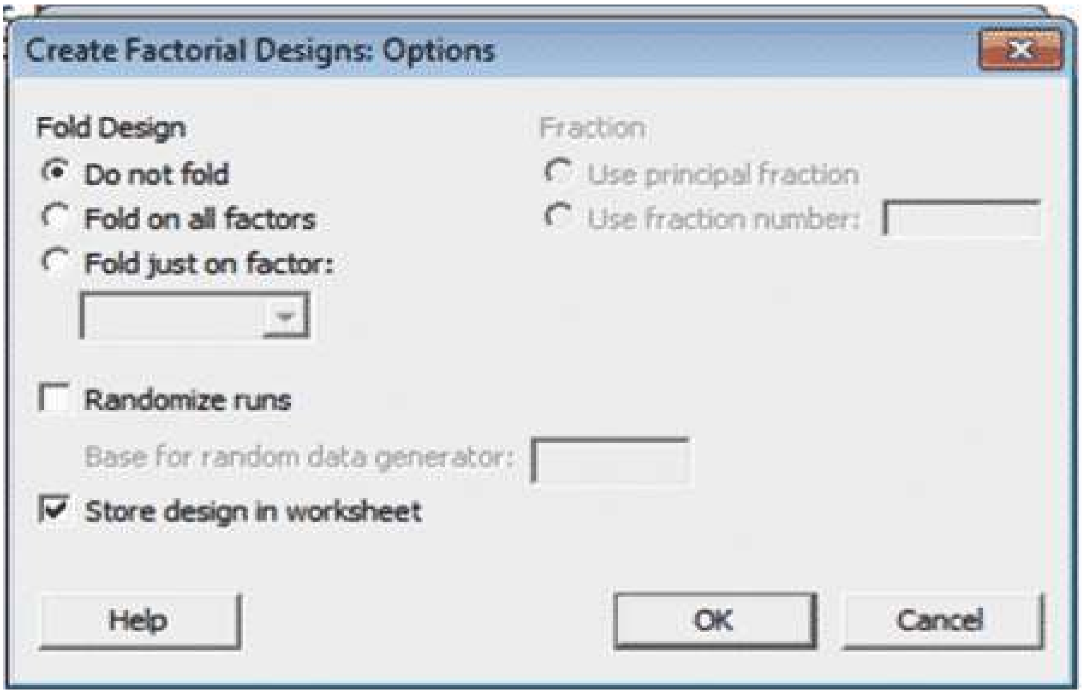


Figure 35: Window allowing to specify the experimental design

- **Outputs choice**

➢ Click the ***Results*** button in the window in Figure 23 and choose the results you want to print in the window that opens (Figure 36).

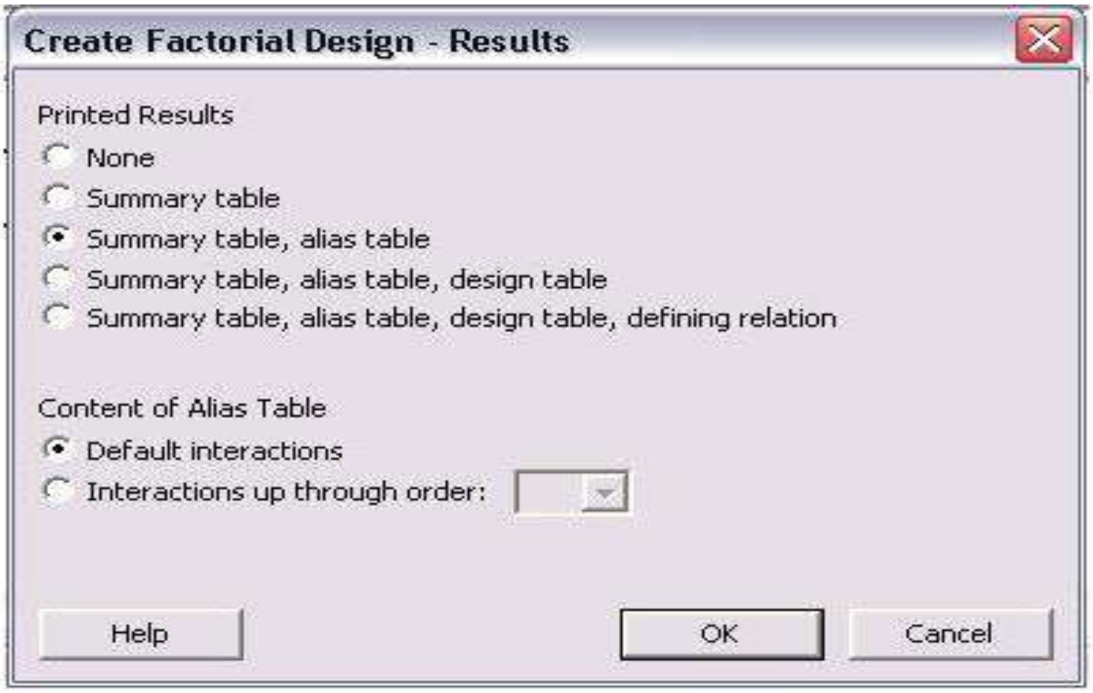


Figure 36: Window allowing to indicate the results of the mathematical analysis

- The information that is requested are saved in the two windows "Worksheet" (figure 37).

Worksheet

↓	C1	C2	C3	C4	C5	C6	C7
	OrdreStd	OrdEssai	PtCentr	Blocs	Speed	Addi weight	
1	1	1	1	1	-1	-1	
2	2	2	1	1	1	-1	
3	3	3	1	1	-1	1	
4	4	4	1	1	1	1	
5							

Figure 37: Minitab Worksheet for 2 factors

- **Statistical analysis of design**

When the experiments are carried out and we have all the responses. we can proceed to the mathematical or statistical analysis of the results. You must first enter the answers in a new column of the worksheet (figure 38).

worksheet

↓	C1	C2	C3	C4	C5	C6	C7	C8
	OrdreStd	OrdEssai	PtCentr	Blocs	Speed	Addi weight	Consumption	
1	1	1	1	1	-1	-1	8,3	
2	2	2	1	1	1	-1	10,7	
3	3	3	1	1	-1	1	9,7	
4	4	4	1	1	1	1	12,3	
5								

Figure 38: Responses are entered into the Minitab worksheet

- Click on ***Stat*** from the main menu. you get a drop-down menu (Figure 29) where you choose ***Analyze Factorial Design***. The window in Figure 39 appears.

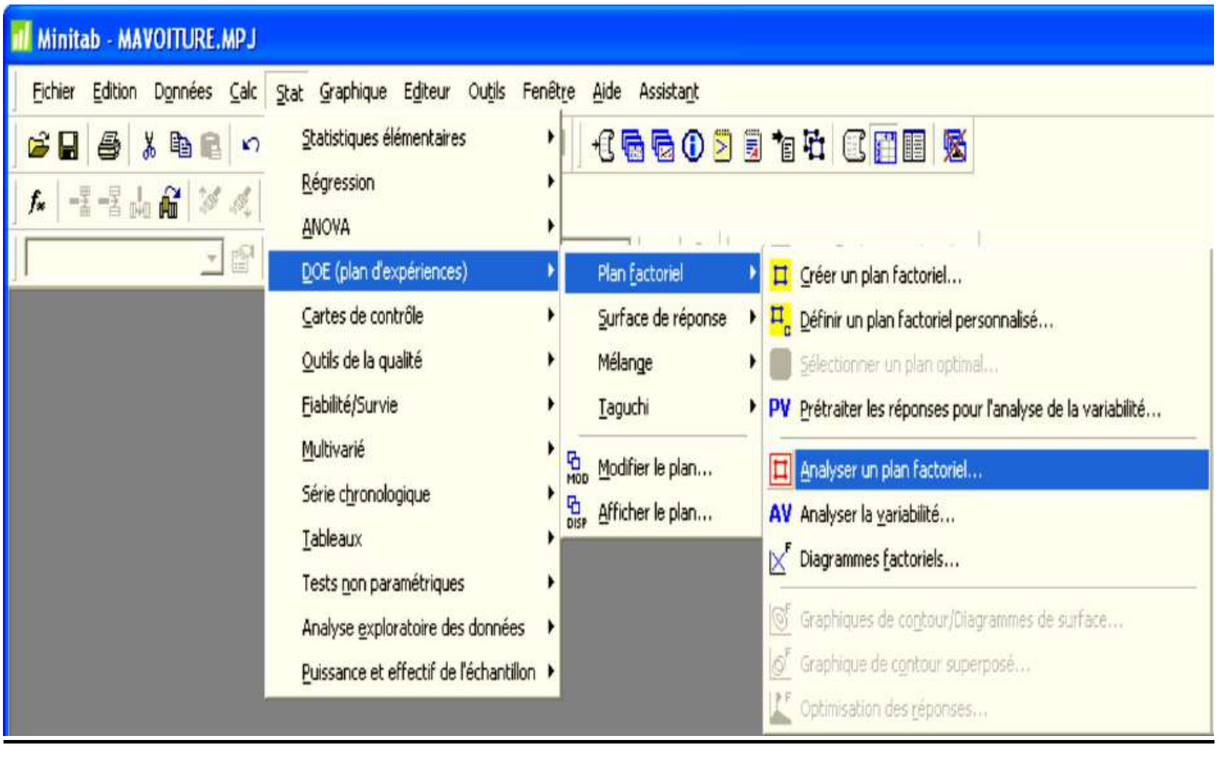


Figure 39: Drop-down menus allowing you to begin the analysis of the experimental design

- In the ***Analyze Factorial Design*** window you must choose the response to analyze. To do this, highlight **C7 Consumption** and click on the Select button (figure 40).

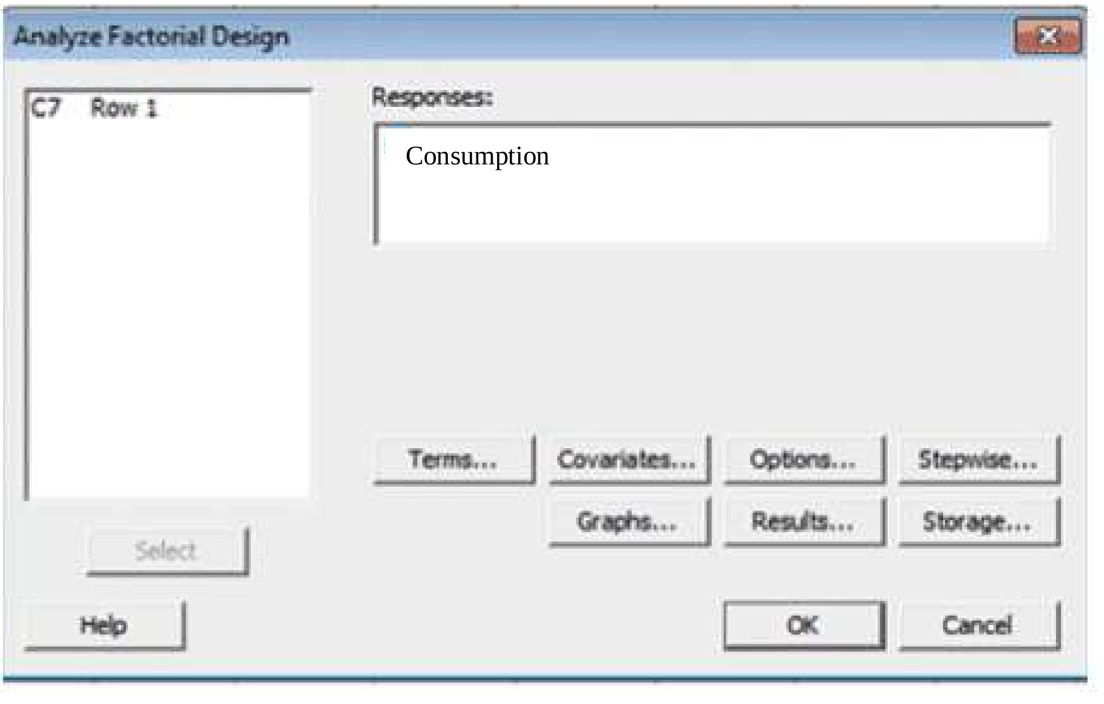


Figure 40: Window allowing you to specify the analysis methods

- The plan being very simple we will only use the **Terms**. **Results** and **Storage** buttons.

➢ **Terms**

Allows you to choose the mathematical model. Here we choose a polynomial model with a single interaction. click the ***OK*** button after that (figure 40).

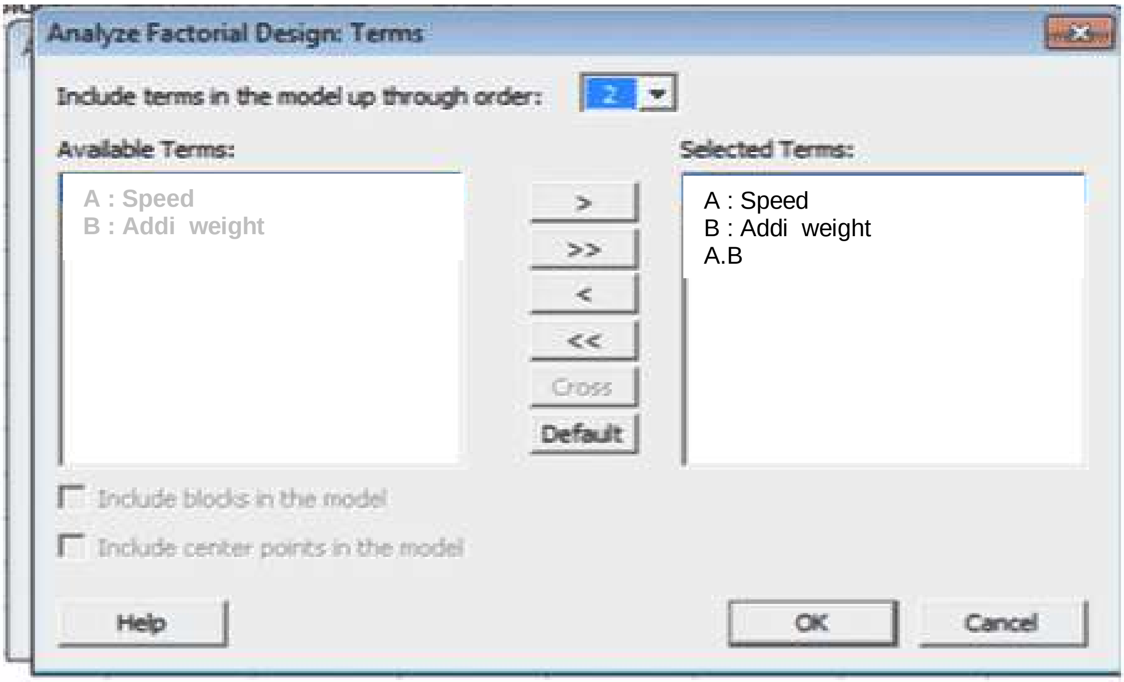


Figure 41: Window allowing you to choose the mathematical model

➢ **Results**

This window allows us to check all the useful results we want to display (like ANOVA. coefficients….) (figure 42).

- Click the **OK** button when you have finished making your choices. This window disappears.

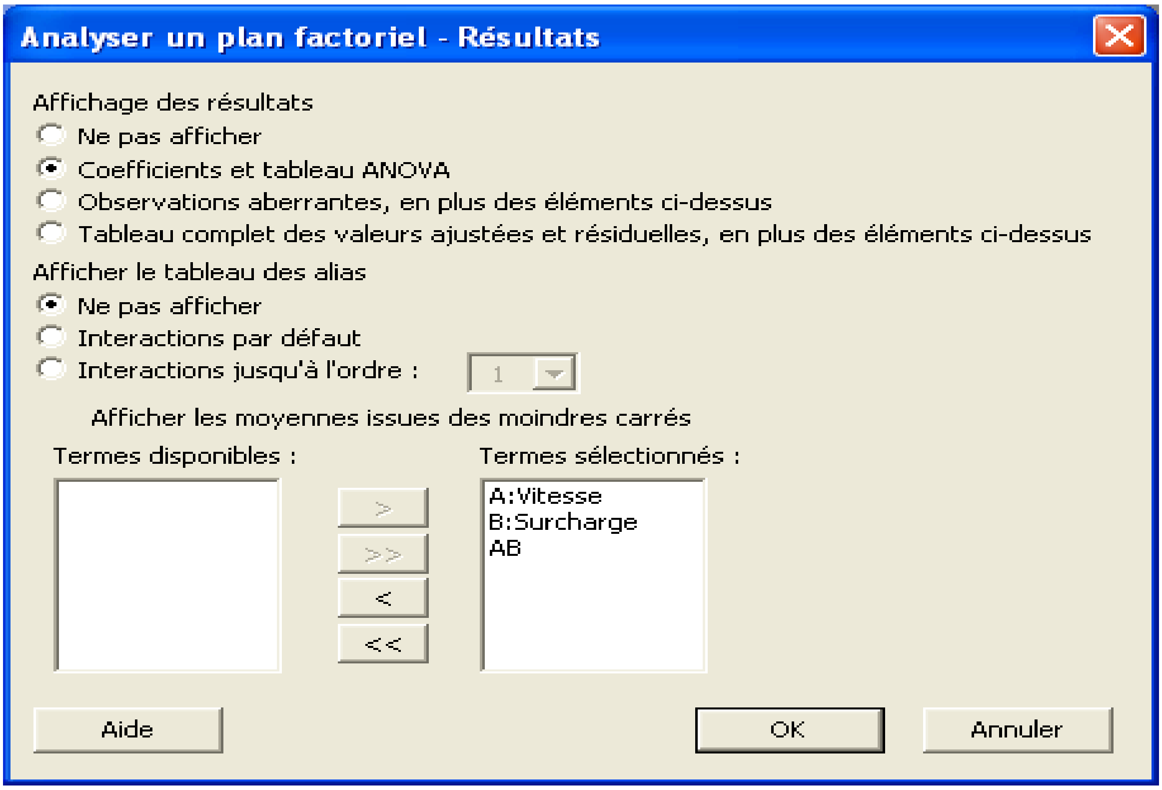


Figure 42: Window allowing you to select useful results

- **Storage**

We choose to keep the effects and coefficients (figure 43).

- Click the ***OK*** button when you have finished making your choices. This window disappears.

Figure 43: Outputs choice

- Click the ***OK*** button in Figure 42. The results are displayed in the session window (Figure 44) and in the worksheet (Figure 45).

Session

Plan factoriel : Consommation en fonction de Vitesse; Surcharge

```
Effets et coefficients estimés pour Consommation (unités codées)

Terme                  Effet     Coeff
Constante                      10,2500
Vitesse               2,5000    1,2500
Surcharge             1,5000    0,7500
Vitesse*Surcharge     0,1000    0,0500
```

Figure 44: The results analysis is in the session window

Feuille de travail 3 ***

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	Vitesse	Surcharge	Consommation	EFFETS1	COEFF1
1	1	1	1	1	-1	-1	8,3	2,5	10,25
2	2	2	1	1	1	-1	10,7	1,5	1,25
3	3	3	1	1	-1	1	9,7	0,1	0,75
4	4	4	1	1	1	1	12,3		0,05
5									

Figure 45: The analysis results are also displayed in the worksheet

The regression equation in coded units is then:

$$\hat{y} = 10.25 + 1.25\, x_1 + 0.75\, x_2 + 0.05\, x_1\, x_2$$

There is virtually no interaction between the two factors according to the model. Whether motorist drive fast or slowly, additional weight always leads to an increase in gas consumption. To reduce consumption, motorist must therefore minimize the speed and the additional weight as often as possible.

Example 2: 2^3 factorial design [4, 12-16]

An experimenter studies the deposition of gold by electrolysis on metal objects in order to give them a golden appearance. Three factors are studied and their values and levels are given in table 19. The experimenters decided to applied a full factorial design 2^3.

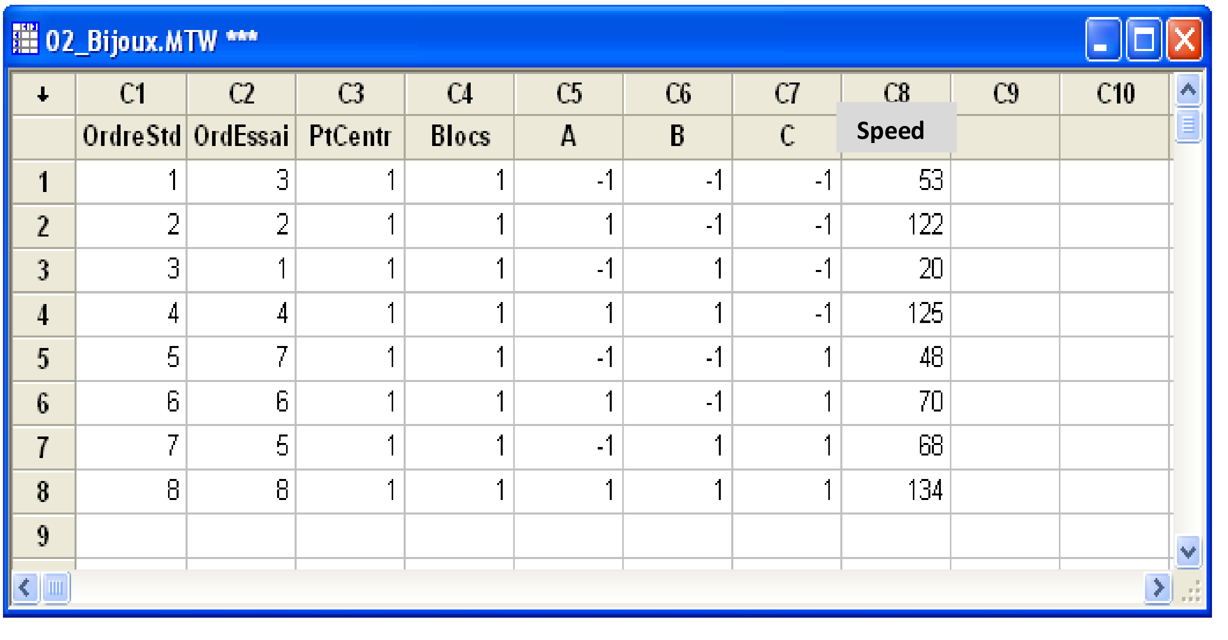
The response is the speed of deposition of the gold on the treated object.

Table 19: Factors values at different levels

Factor	**Low level (-1)**	**High level (+1)**
Gold concentration (A)	2 g/L	15 g/L
Current density (B)	5 A/dm^2	25 A/dm^2
Cobalt Concentration (C)	0.5 g/L	1.5 g/L

- **Design construction**

  - You should get a worksheet similar to figure 46 after introduce the response (**speed**).


02_Bijoux.MTW ***

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Speed		
1	1	3	1	1	-1	-1	-1	53		
2	2	2	1	1	1	-1	-1	122		
3	3	1	1	1	-1	1	-1	20		
4	4	4	1	1	1	1	-1	125		
5	5	7	1	1	-1	-1	1	48		
6	6	6	1	1	1	-1	1	70		
7	7	5	1	1	-1	1	1	68		
8	8	8	1	1	1	1	1	134		
9										

Figure 46: Worksheet with the example data

- **Design analysis**

The analysis of the design is carried out in the same way as previously and you must obtain the same information in the session window as that in figure 46.

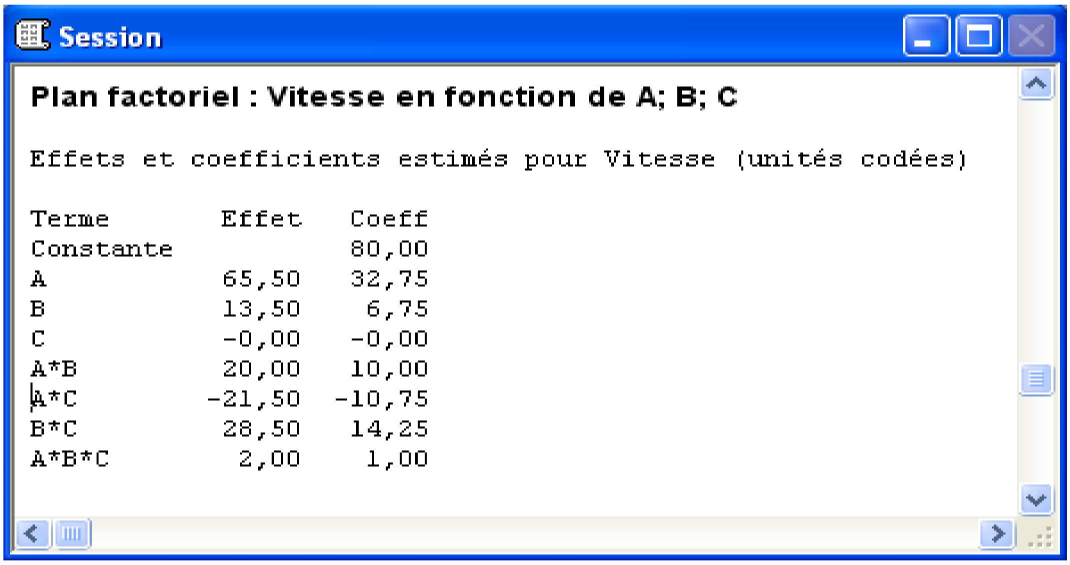


Figure 47: Session sheet with results: Effects and coefficients

These numerical results can be illustrated by the effects graph.

- Click on the ***Stat*** button on the main menu then choose ***DOE*** (design of experiments). ***Factorial design*** and ***Factorial diagrams*** (figure 22). You get a new window (figure 47) where you can choose the effects graph. the interactions graph and the cube graph.

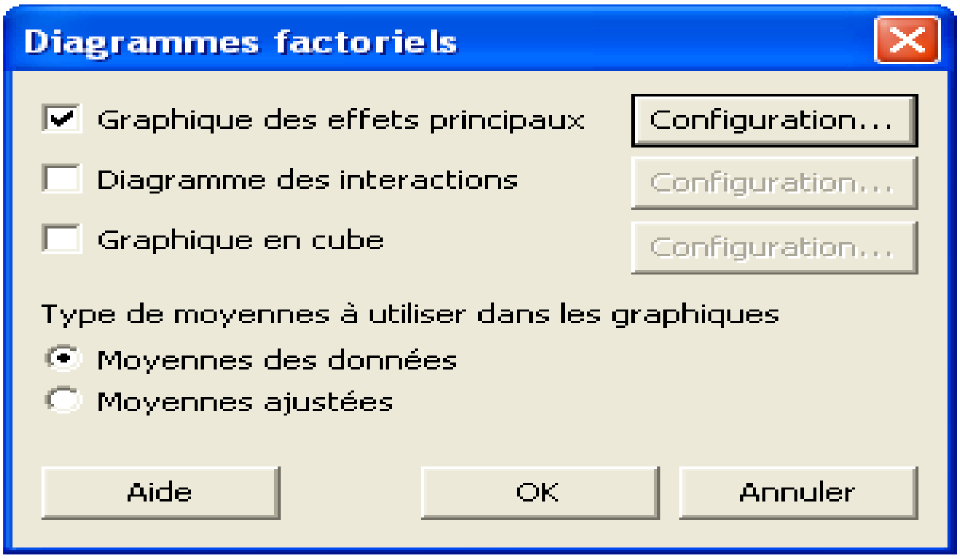


Figure 48: Window allowing us to choose the effects graph, the interactions graph and the cube graph

- **Effect graph**

  - Check ***Main Effects Plot*** in the window in figure 48.
  - Click on the Window ***Configuration*** button in Figure 48. You will get a window asking you for the response and the factors you want to illustrate (figure 49).

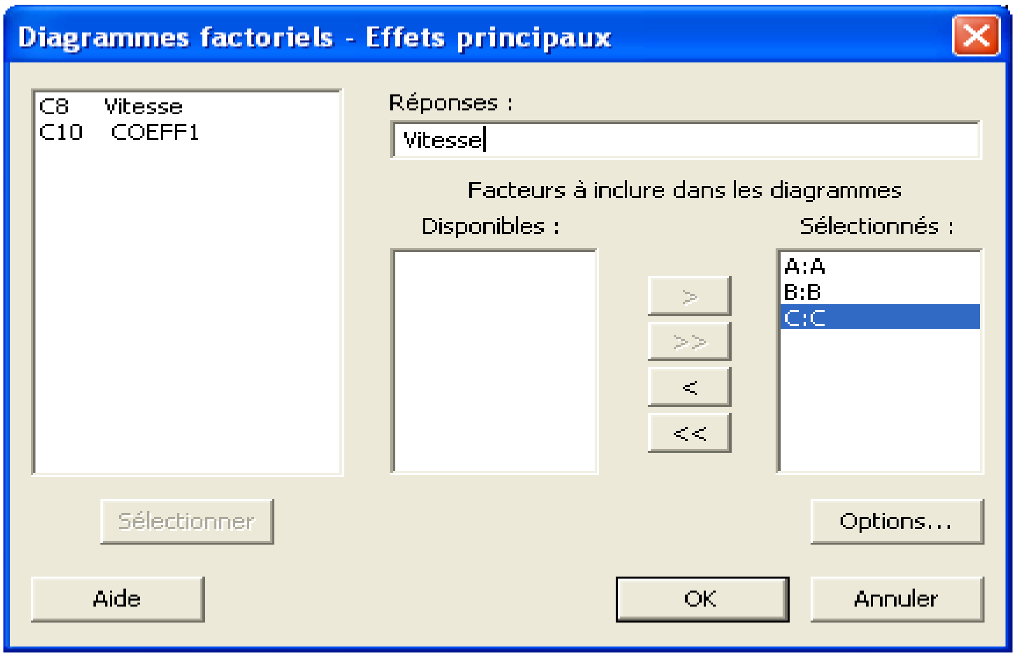


Figure 49: Definition of effects graphs

- Click on the ***OK*** button in the window in Figure 48. You return to Figure 48.
- Click on the OK button in the window in Figure 49. You obtain a window where the requested graphics are located (Figure 50).

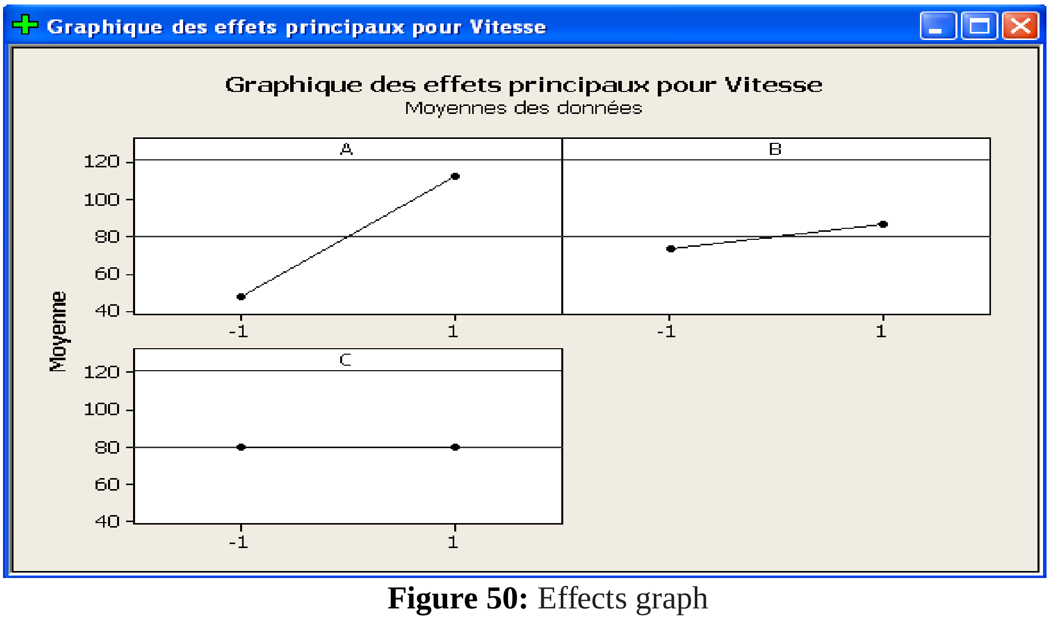


Figure 50: Effects graph

- **Interaction graph**

By an analogous process. we obtain the interaction graph (Figure 51).

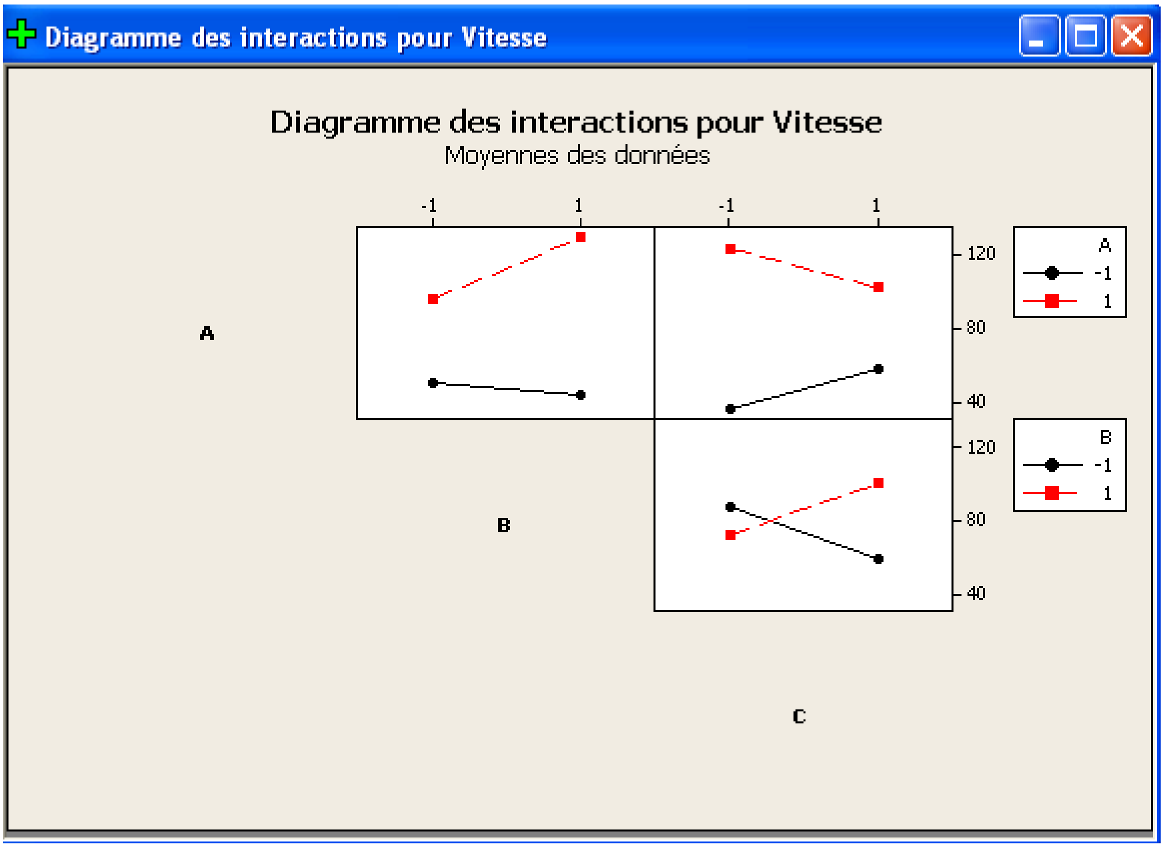


Figure 51: Interaction graph

- **Interpretation of plan analysis and conclusion**

All the precedent steps allow us to find the optimal conditions which maximize the speed of gold deposition.

  - The gold concentration (A) must be at level +1. i.e. 15 g/l
  - Current density (B) must be at level +1. i.e. 25 A/dm^2
  - The cobalt concentration (C) must be at level +1 or 1.5 g/L

Under these conditions we can expect a speed of 134 mg/min.

Example 3 Response surface design (CCD) [4, 12-16].

The objective of the foreman who conducts the study is to adjust a machine tool so that the surface condition of the machined parts is as close as possible to perfection. The **roughness of the surface** is the response studied**;** it is measured by a standardized method. The smallest possible value is desired.

The foreman uses two factors:

- Factor 1: forward speed of the grinding wheel (in meters/minute).
- Factor 2: tangential cutting speed (in meters/second).

The high and lower levels of each factor are given in table 20.

Table 20: Study domain of CCD

Factor	–1.21 Level	–1 Level	0 Level	+1 Level	+1.21 Level
Forward speed (1)	0.74	0.9	1.65	2.4	2.56
Cutting speed (2)	13.95	15	20	25	26.05

The foreman starts with a traditional factorial design. but he suspects that he will have to continue the study with a response surface design by CCD. Therefore. he plans to have two control points at the center of the domain.

- **Design construction**
- Click on ***Stat*** from the main menu. A drop-down menu appears in which is DOE.
- Click on ***DOE***. A new menu appears in which there are four choices (figure 51).

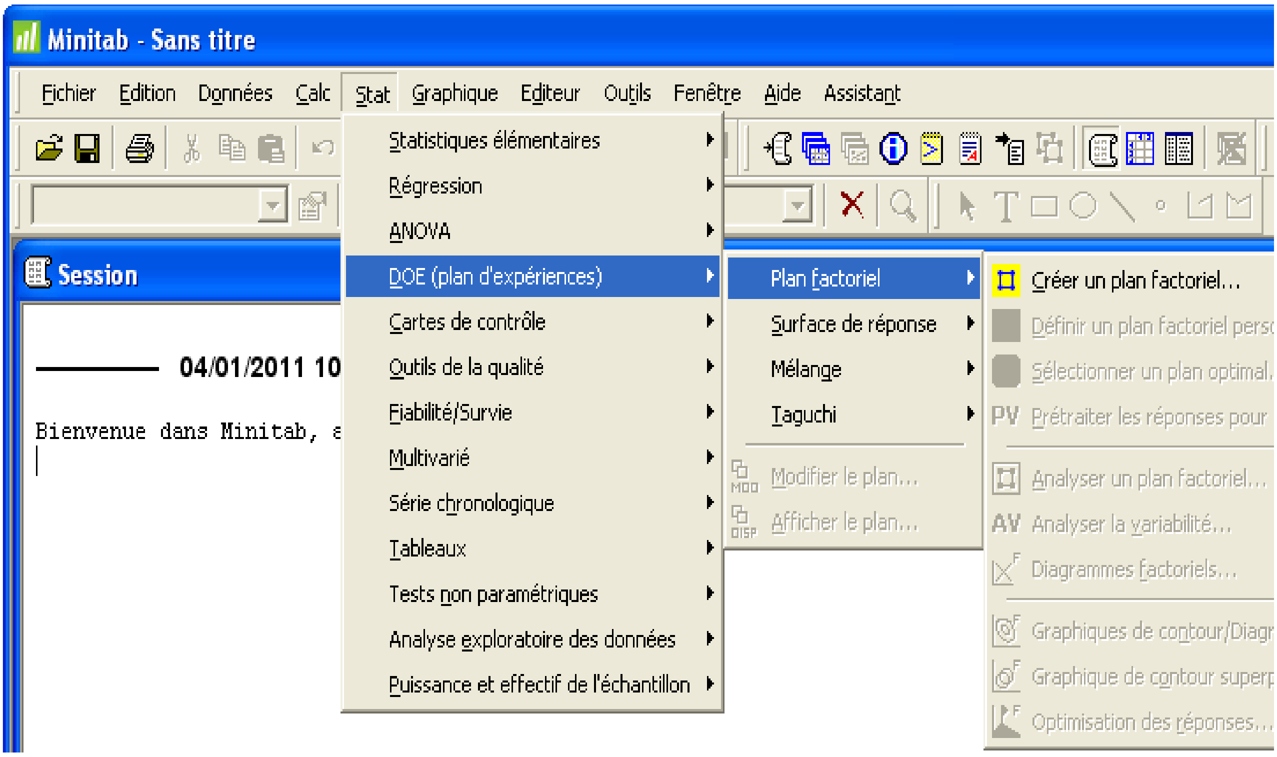


Figure 52: Menu allowing you to create a response surface design

- Click ***Create Response Surface design***. A new window appears which allows you to specify the design for the response surface (Figure 53). You choose a ***two-factor central composite design.***

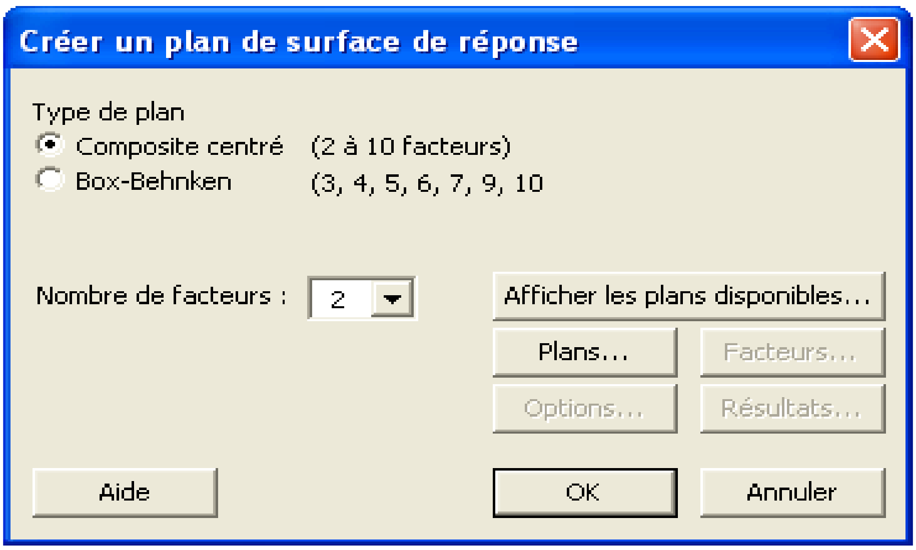


Figure 53: Menu allowing you to create a response surface design

- Click the ***Designs*** button. You get a new window where you can set the number of points in the center (figure 54).

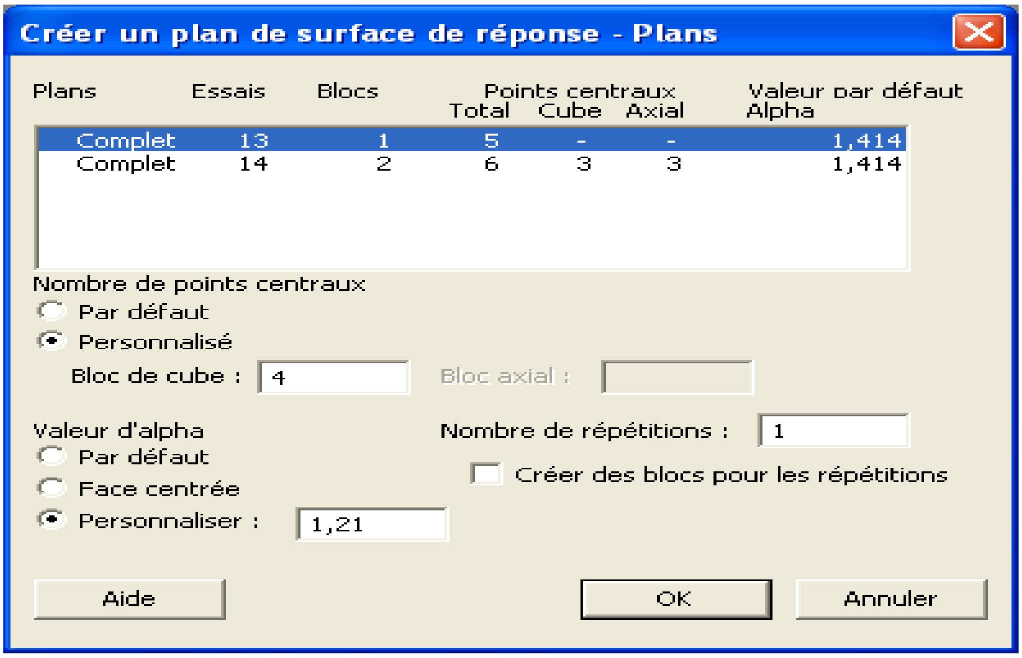
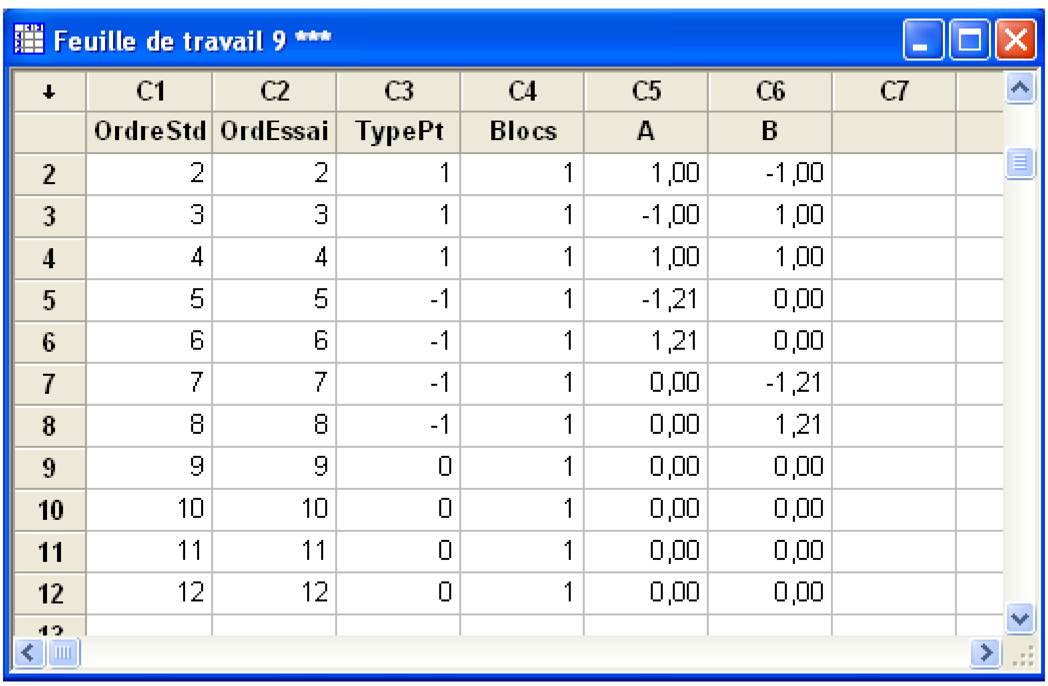


Figure 54: Menu allowing you to specify the characteristics of the response surfaces design

- ➢ Click the ***OK*** button. You return to the window in Figure 53.
- ➢ In the ***option,*** uncheck **randomization.**
- ➢ Click the ***OK*** button in Figure 53. You obtain the desired design (Figure 55).


Feuille de travail 9 ***

↓	C1	C2	C3	C4	C5	C6	C7
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	
2	2	2	1	1	1,00	-1,00	
3	3	3	1	1	-1,00	1,00	
4	4	4	1	1	1,00	1,00	
5	5	5	-1	1	-1,21	0,00	
6	6	6	-1	1	1,21	0,00	
7	7	7	-1	1	0,00	-1,21	
8	8	8	-1	1	0,00	1,21	
9	9	9	0	1	0,00	0,00	
10	10	10	0	1	0,00	0,00	
11	11	11	0	1	0,00	0,00	
12	12	12	0	1	0,00	0,00	

Figure 55: Minitab worksheet the CCD response surface design

➢ Design analysis

➢ You must copy the response values into the worksheet (figure 56).

12_Rectificationsurface_de_reponse.MTW ***

↓	C1	C2	C3	C4	C5	C6	C7	C8
	OrdreStd	OrdEssai	Blocs	TypePt	Avance	Coupe	Rugo	Pics
1	1	1	1	1	-1,00	-1,00	194	
2	2	2	1	1	1,00	-1,00	282	
3	3	3	1	1	-1,00	1,00	120	
4	4	4	1	1	1,00	1,00	91	
5	5	5	1	-1	-1,21	0,00	154	
6	6	6	1	-1	1,21	0,00	195	
7	7	7	1	-1	0,00	-1,21	278	
8	8	8	1	-1	0,00	1,21	122	
9	9	9	1	0	0,00	0,00	232	
10	10	10	1	0	0,00	0,00	230	
11	11	11	1	0	0,00	0,00	233	
12	12	12	1	0	0,00	0,00	235	
13								

Figure 56: Design matrix and responses of CCD

- To begin the statistical analysis:

➢ Click ***Stat*** from the main menu and choose ***DOE*** / ***Response Surface*** / Analyze Response Surface design (Figure 29).

➢ In the ***Analyze Response Surface design*** window that appears you enter the response (figure 57).

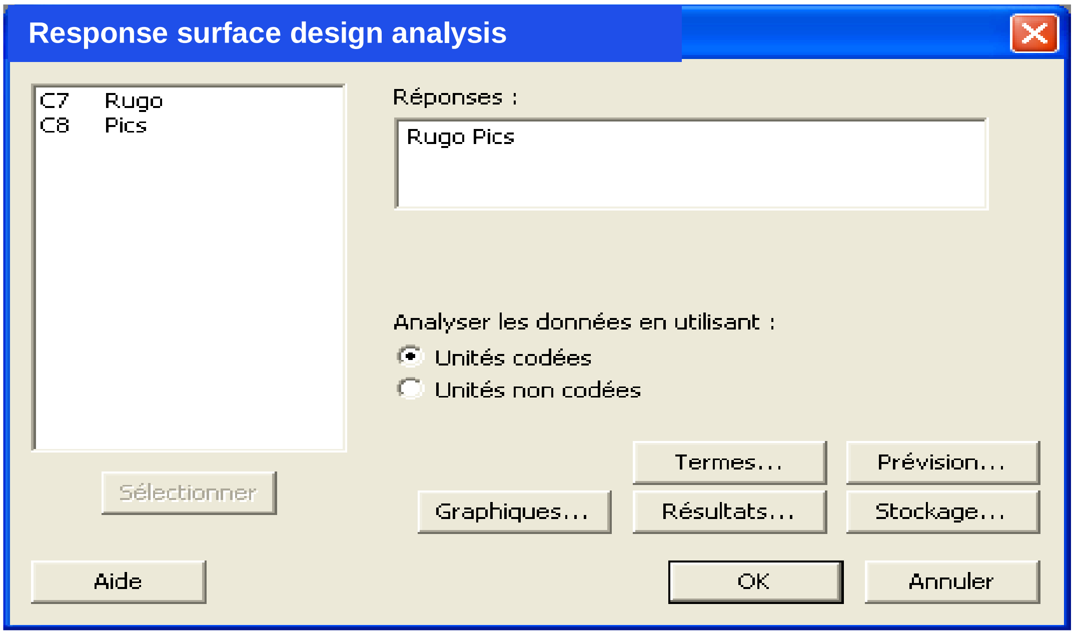


Figure 57: Analyze a response surface design window allowing you to define the desired analyses.

- Then. you click on the ***Terms*** button. The software offers a quadratic model. you accept by clicking on the ***OK*** button (Figure 58). You return to Figure 57.

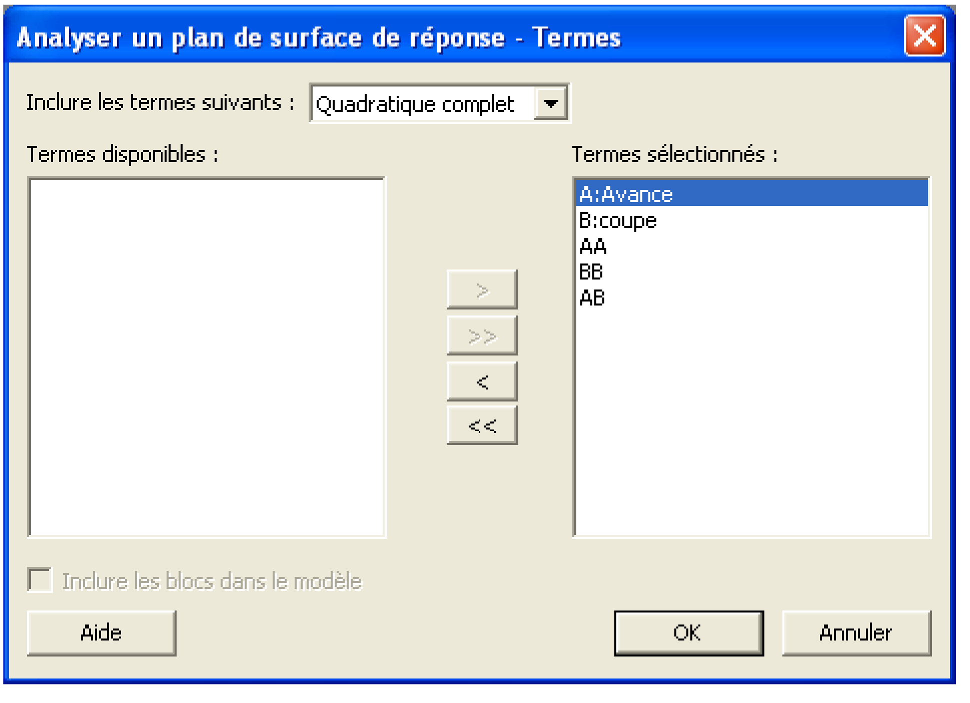


Figure 58: ***Analyze Response Surface design*** window - **Terms** for defining the mathematical model.

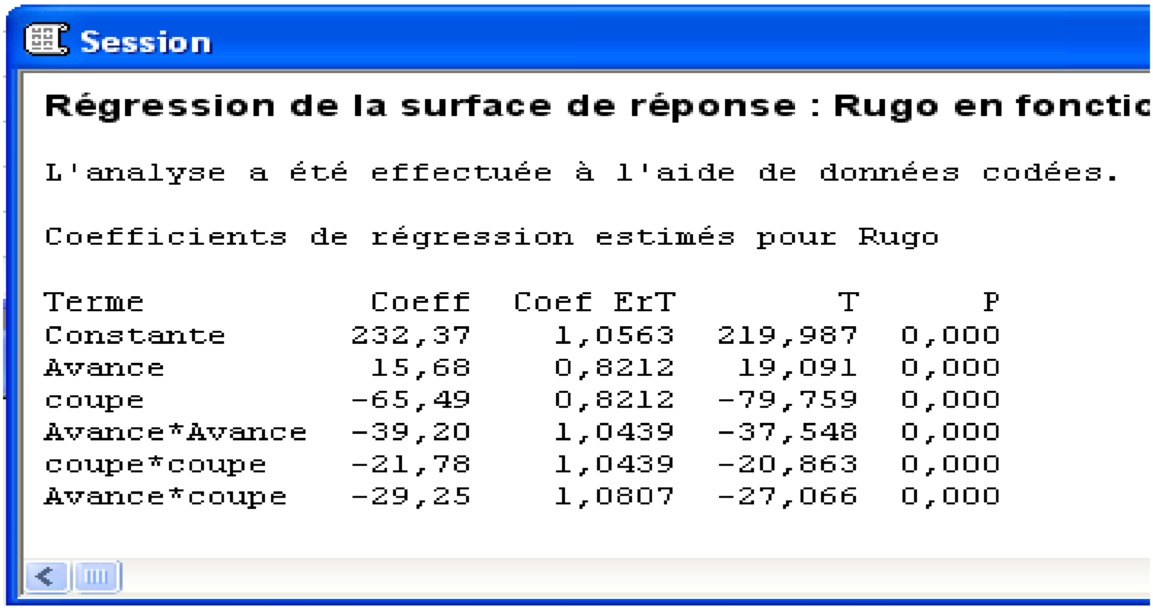
- You give your output instructions using the other buttons in Figure 58 and finish by clicking on the ***OK*** button.

The coefficients of the quadratic model are displayed in the worksheet (figure 59) and the statistical results are displayed in the Session window (figure 60).

Worksheet 3 ***

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	TypePT	Blocs	Avance	Coupe	Rugo	Pics	COEF1	COEF2	
1	1	1	1	1	-1,00	-1,00	194	77,8	232,370	62,0854	
2	2	2	1	1	1,00	-1,00	282	68,4	15,677	4,5035	
3	3	3	1	1	-1,00	1,00	120	65,3	-65,495	3,7134	
4	4	4	1	1	1,00	1,00	91	96,1	-39,196	-4,3263	
5	5	5	-1	1	-1,21	0,00	154	52,3	-21,779	19,5792	
6	6	6	-1	1	1,21	0,00	195	60,4	-29,250	10,0500	
7	7	7	-1	1	0,00	-1,21	278	87,0			
8	8	8	-1	1	0,00	1,21	122	95,7			
9	9	9	0	1	0,00	0,00	232	61,5			
10	10	10	0	1	0,00	0,00	230	60,5			
11	11	11	0	1	0,00	0,00	233	63,8			
12	12	12	0	1	0,00	0,00	235	61,9			
13											

Figure 59: Minitab Worksheet with responses and coefficients



Session

Régression de la surface de réponse : Rugo en fonctic

L'analyse a été effectuée à l'aide de données codées.

Coefficients de régression estimés pour Rugo

Terme	Coeff	Coef ErT	T	P
Constante	232,37	1,0563	219,987	0,000
Avance	15,68	0,8212	19,091	0,000
coupe	-65,49	0,8212	-79,759	0,000
Avance*Avance	-39,20	1,0439	-37,548	0,000
coupe*coupe	-21,78	1,0439	-20,863	0,000
Avance*coupe	-29,25	1,0807	-27,066	0,000

Figure 60: Statistical analysis results in session window

- **Interpretation of design analysis**

To interpret the previous results. the simple way is to draw the isoresponse graphs of the response. We thus determine the places in the studied domain where the objectives are or are not achieved.

- ➢ Click ***Stat*** from the main menu and choose ***DOE / Response Surface / Contour Plots / Surface Plots*** (Figure 61).

- ➢ In the ***Contour Plots/Surface Plots*** window that appears. you check ***Contour Plot*** and click the ***Setup*** button (Figure 61). You will see the **Contour Plots/Surface Plots-Contour** window (Figure 62).

Figure 61: Contour Plot/Surface Plot window

- To specify the characteristics of the graph you use the ***Contours. Configuration*** and ***Options*** buttons.

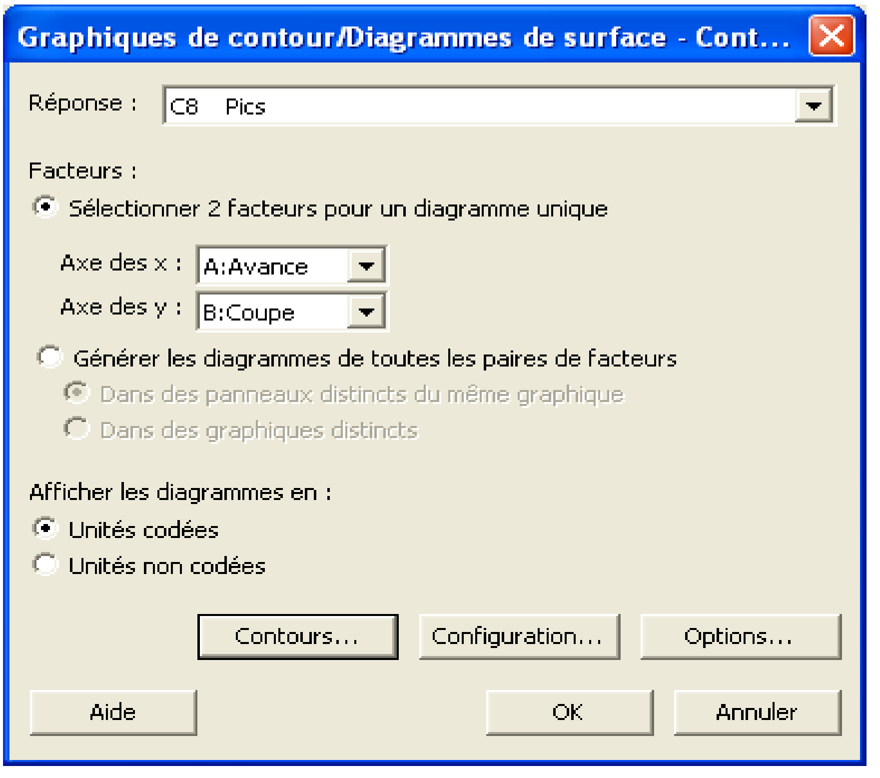


Figure 61: Contour Plot/Surface contour Plot window

- The ***Contours*** button gives access to a window (Figure 62) allowing you to indicate the number of isoresponse curves and the properties of the graph.

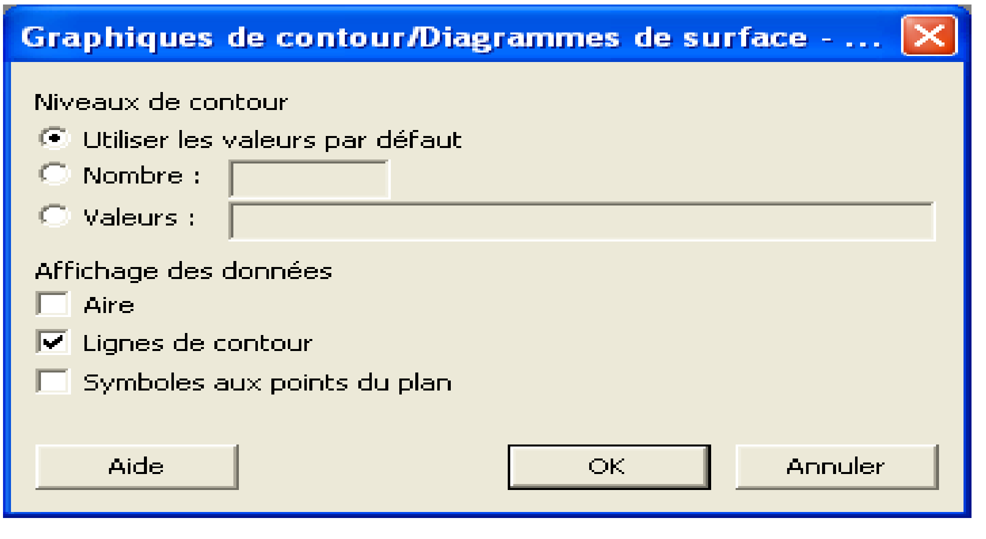


Figure 62: Window allowing you to specify the nature of the isoresponse graph

The objective of this study is to obtain a machined surface having a roughness less than 150. Figure 63 shows that this objective is achieved when the cutting speed of 0.1 to 0.4 and the level of forward speed equal approximately to (-1.2).

- -1.2 of forward speed correspond to 0.75 meters/minute.
- 0.1 to 0.4 level of the cutting speed approximately 20 to 25 meters/second.

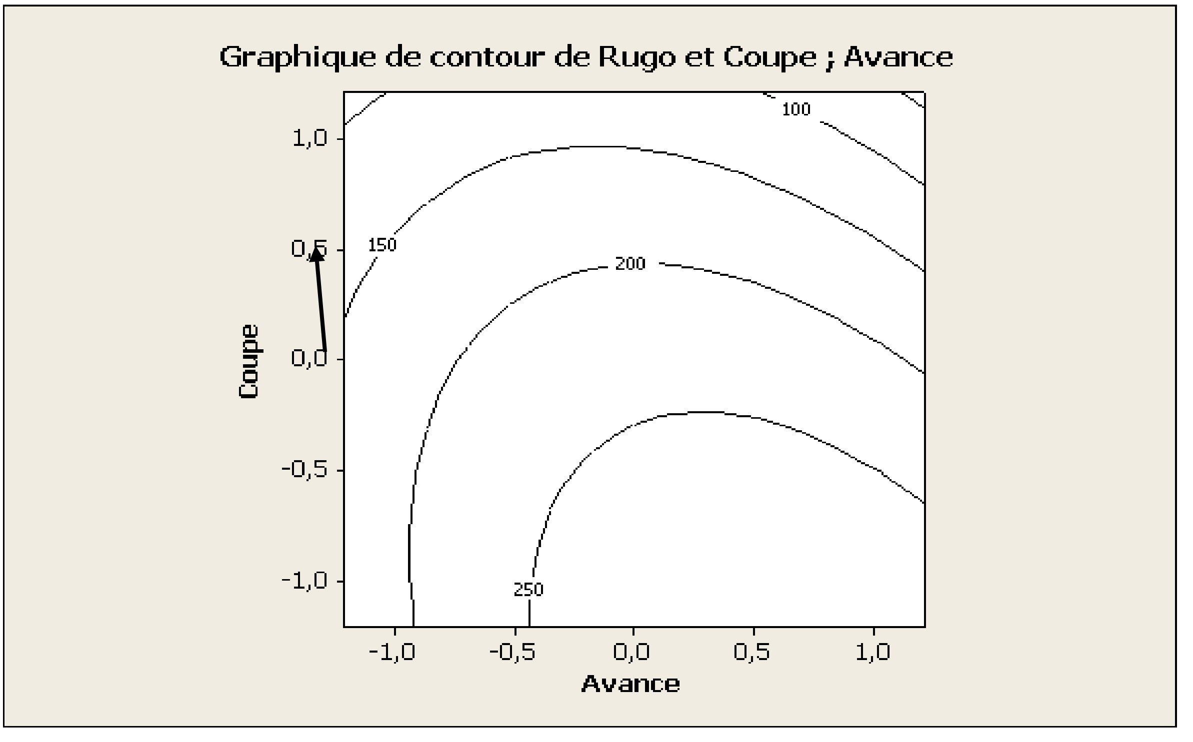


Figure 63: Contour plots of the surface roughness

Example 4: Response surface design by BBD [4. 12-16]

The aim of this study is to decrease the acidic taste of the fermented milk. For this reason. stabilized milk is produced from a natural stabilizer, which attenuates the variations of acid in the final product in spite of the presence of lactic leaveners.

- The response chosen by the researcher is acid loss. The goal is to get stabilized milk with an acid loss of at least 48.

The three factors used in this experiment are:

- Factor 1: Dilution ratio. This is the ratio of added water to raw milk.
- Factor 2: pH. which is related to the injected stabilizer.
- Factor 3: Milk ratio. This is the ratio of raw milk to stabilized milk.

The high and low levels of each factor are shown in table 21.

Table 21: Design domain of BBD

Factor	−1 Level	+1 Level
Dilution (1)	0.5	2
pH (2)	6	5
Concentration (3)	1.5	2.5

- **Design construction**
- Click on ***Stat*** from the main menu. A drop-down menu appears in which is DOE.
- Click on ***DOE***. A new menu appears in which there are four choices.
- Click ***Create Response Surface design***. A new window appears which allows you to specify the design for the response surface (Figure 38).
- Check Box-Behnken in the Create Response Surface design window (figure 64).

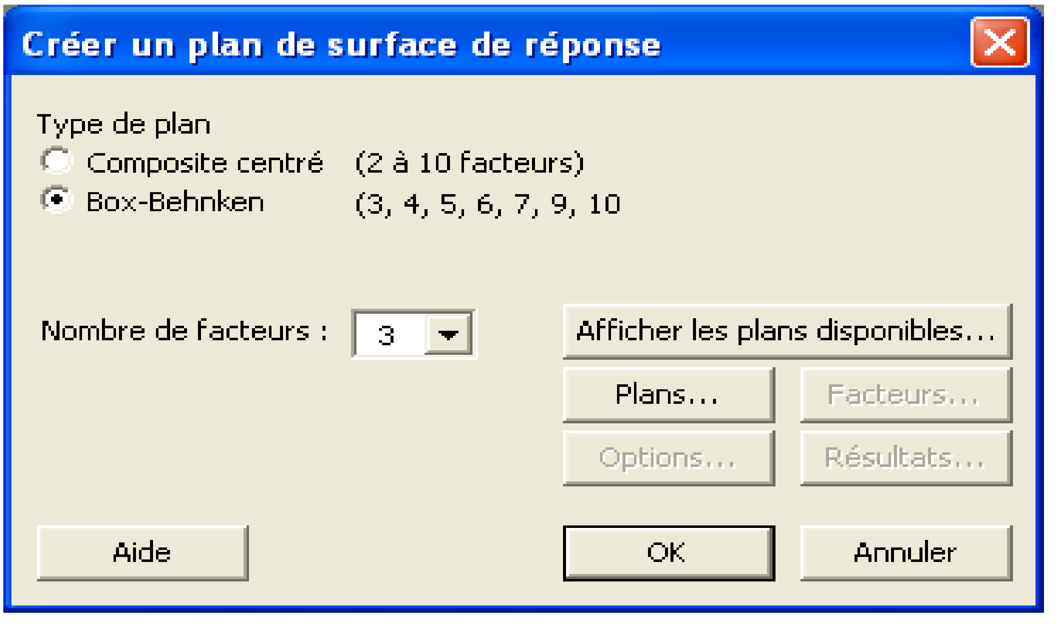


Figure 64: Menu allowing you to create a Box-Behnken design

- The software gives the desired plan (Figure 65).

Feuille de travail 3 ***

↓	C1 OrdreStd	C2 OrdEssai	C3 TypePt	C4 Blocs	C5 A	C6 B	C7 C	C8
1	1	1	2	1	-1	-1	0	
2	2	2	2	1	1	-1	0	
3	3	3	2	1	-1	1	0	
4	4	4	2	1	1	1	0	
5	5	5	2	1	-1	0	-1	
6	6	6	2	1	1	0	-1	
7	7	7	2	1	-1	0	1	
8	8	8	2	1	1	0	1	
9	9	9	2	1	0	-1	-1	
10	10	10	2	1	0	1	-1	
11	11	11	2	1	0	-1	1	
12	12	12	2	1	0	1	1	
13	13	13	0	1	0	0	0	
14	14	14	0	1	0	0	0	
15	15	15	0	1	0	0	0	
16								

Figure 65: Minitab worksheet with the Box-Behnken design

➢ **Box-Behnken design analysis**

When experiments are achieved. you must copy the response values into the worksheet (figure 66). and you carry out the calculations as in the previous example.

13_Yoghourt.MTW ***

↓	C1 OrdreStd	C2 OrdEssai	C3 TypePt	C4 Blocs	C5 Dilution (1)	C6 pH (2)	C7 Concentration (3)	C8 Appauvrissement	C
1	1	1	2	1	-1	-1	0	51,3	
2	2	2	2	1	1	-1	0	42,6	
3	3	3	2	1	-1	1	0	42,2	
4	4	4	2	1	1	1	0	50,4	
5	5	5	2	1	-1	0	-1	40,7	
6	6	6	2	1	1	0	-1	41,3	
7	7	7	2	1	-1	0	1	41,5	
8	8	8	2	1	1	0	1	40,8	
9	9	9	2	1	0	-1	-1	39,5	
10	10	10	2	1	0	1	-1	35,3	
11	11	11	2	1	0	-1	1	35,2	
12	12	12	2	1	0	1	1	39,8	
13	13	13	0	1	0	0	0	50,8	
14	14	14	0	1	0	0	0	50,1	
15	15	15	0	1	0	0	0	49,4	
16									

Figure 66: Minitab worksheet with the Box-Behnken design and responses

- **Statistical analysis**

➢ Click ***Stat*** from the main menu and choose ***DOE*** / ***Response Surface*** / **Analyze Response Surface design** (Figure 38).

➢ In the ***Analyze Response Surface design*** window that appears you enter the response.

➢ Click **OK.** The statistical analysis appears in the ***Session*** window (Figure 67).

Session

The regression coefficients estimation

Terme	Coeff	Coef ErT	T	P
Constante	50,1000	0,3014	166,232	0,000
Dilution (1)	-0,0750	0,1846	-0,406	0,701
pH (2)	-0,1125	0,1846	-0,610	0,569
Concentration (3)	0,0625	0,1846	0,339	0,749
Dilution (1)*Dilution (1)	0,0750	0,2717	0,276	0,794
pH (2)*pH (2)	-3,5500	0,2717	-13,068	0,000
Concentration (3)*Concentration (3)	-9,1000	0,2717	-33,497	0,000
Dilution (1)*pH (2)	4,2250	0,2610	16,187	0,000
Dilution (1)*Concentration (3)	-0,3250	0,2610	-1,245	0,268
pH (2)*Concentration (3)	2,2000	0,2610	8,429	0,000

Figure 67: Session sheet giving the coefficients of the quadratic model

- **Interpretation of design analysis**

The mathematical model (figure 67 results) indicates that we have an acid loss of 50.1 at the center domain. The objective of 48 will therefore be achieved and even exceeded.

From an economic point of view, it is advantageous to choose a low dilution (le effect of dilution is negative). For this reason. we choose a dilution level of (-1) (i.e. 0.5).

To choose the levels of the other two factors, let's plot the graph of the isoresponses as a function of pH and concentration.

- Click ***Stat*** from the main menu and choose ***DOE / Response Surface / Contour Plots / Surface Plots/*** the ***Contour Plots/Surface Plots***

- You obtain the window of figure 68.

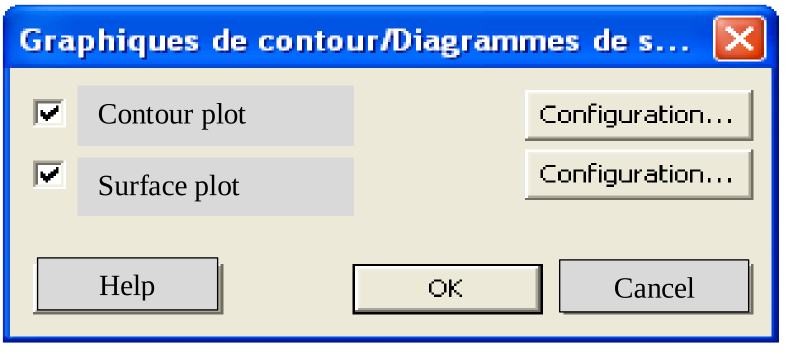


Figure 68: Contour Plot/Surface Plot window of BBD

- Click on the ***Configuration*** button to obtain the graphics definition window (Figure 69).

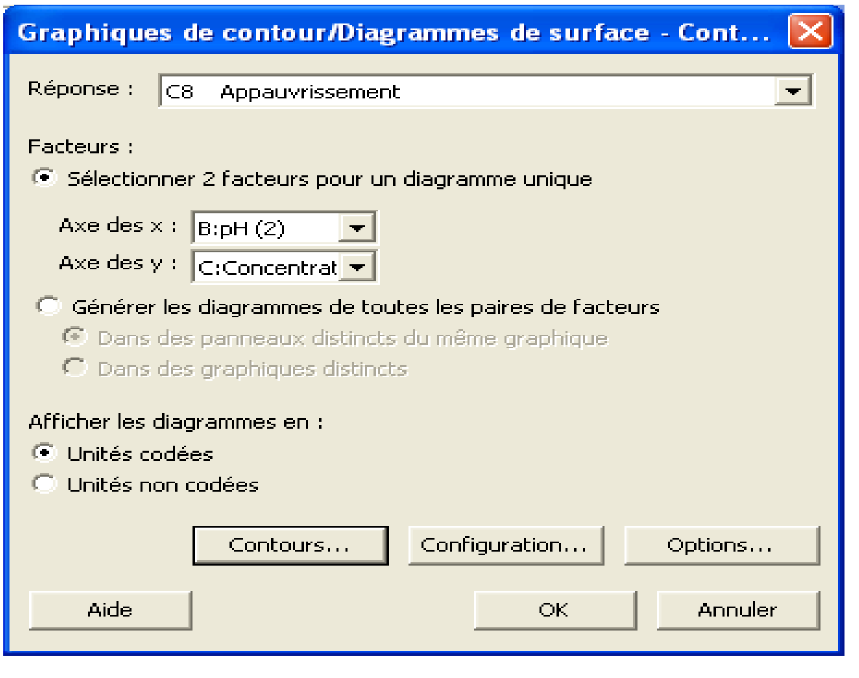


Figure 69: Response and factors choice to illustrate in the graph

- Click on the ***configuration*** button of the figure 69. The ***Contour Plots/Surface Plots-Contour-Configuration*** window appears (Figure 70). You indicate that the dilution is fixed at level (-1).

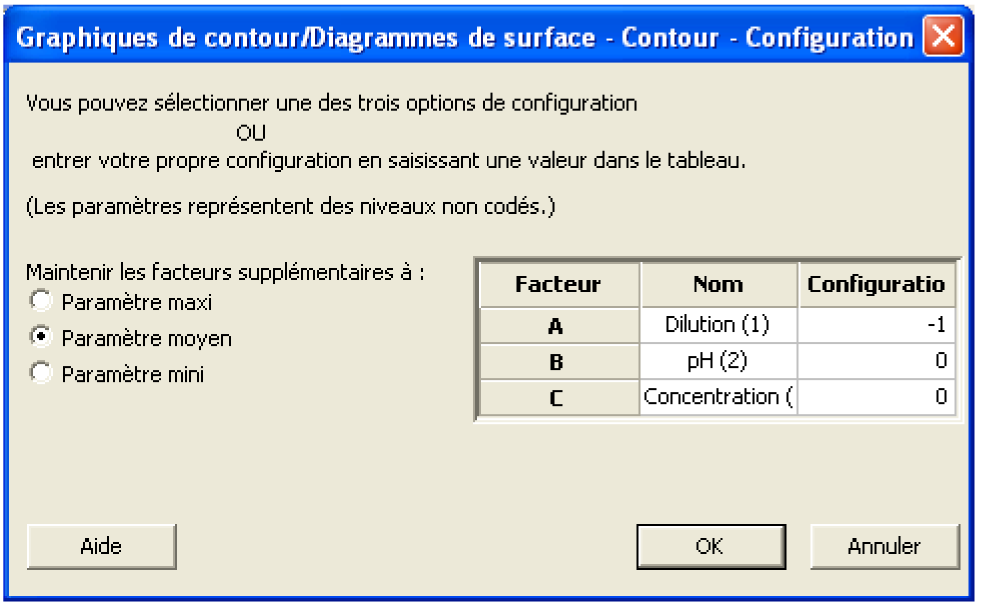


Figure 70: We impose a level to the factors which are not illustrated on the graph

- Click on the OK buttons and you get the desired graphics

You thus obtain the graph of the isoresponses (contour plot) of the response as a function of the concentration and the pH for a dilution of 0.5 (Figure 71).

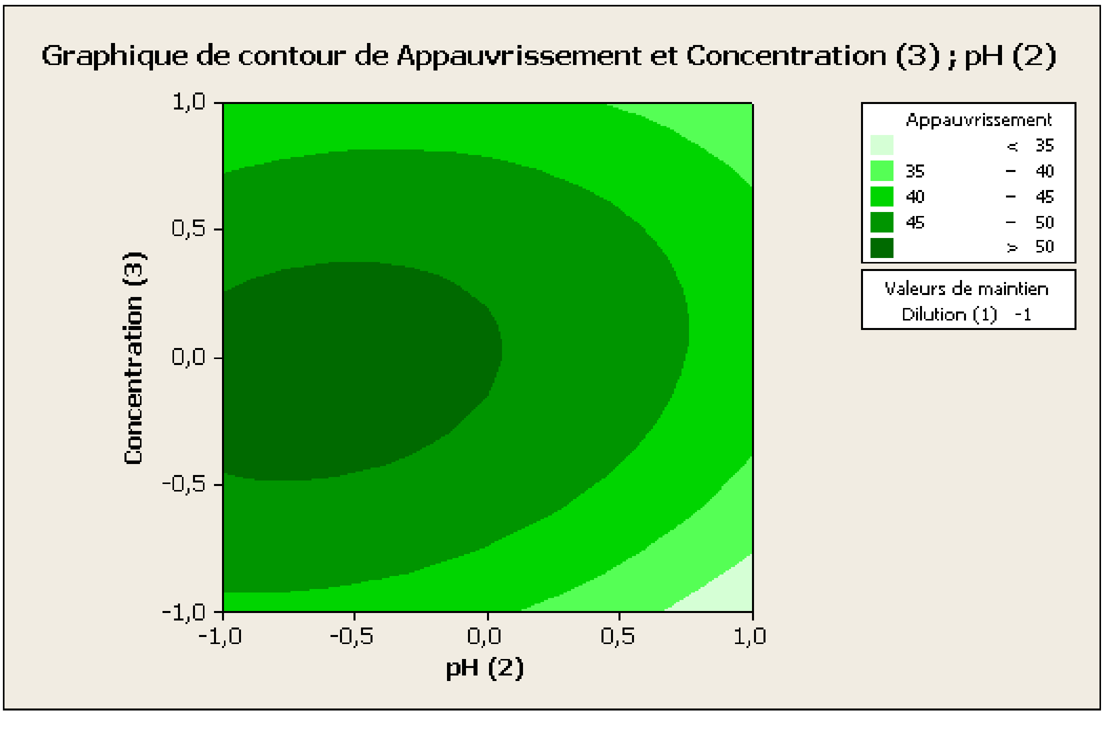


Figure 71: Acid loss as a function of the concentration and the pH for a dilution of 0.5

This graph shows that the zone of acid loss greater than 48 is located around level (-0.5) for pH (5.8 in natural units) and level zero (0) (2 in natural units) for concentration**.**

Validation tests are carried out which confirm the results of the plan and its analysis. We can now recommend the following operating conditions given by Minitab software.

- Dilution (1) : 0.5
- pH (2): 5.8
- Concentration (3): 2

The predicted response at this point is 51.6.

12. Application using Excel software

Example [18]

The effects of four experimental parameters on the cementation yield of copper by iron were investigated statistically. A statistical experimental design based on the second-order central composite rotatable design (CCRD) was planned fixing the cementation period at 2 h. The original values of each factor and their corresponding levels are given in the following table.

Values and levels of operating parameters

Operating factors	Levels				
	−2	−1	0	1	2
Z_1: $[Cu^{2+}]_0$ (mg/L)	10	32.5	55	77.5	100
Z_2: T (°C)	20	30	40	50	60
Z_3: pH	1	2	3	4	5
Z_4: Q_v (mL/s)	0.46	1.44	2.42	3.40	4.38

The cementation yield as dependent output response variable which is expressed as (%):

$$y(\%) = \frac{[Cu^{2+}]_0 - [Cu^{2+}]_t}{[Cu^{2+}]_0} \times 100$$

where y represents the copper cementation reaction yield $[Cu^{2+}]_0$ is the initial copper ions concentration (mg/L) and $[Cu^{2+}]_t$ is the copper ions concentration at time t (mg/L).

The experiments were performed according to the design matrix given in the following.

Experimental design and the results for copper cementation yield

Run no.	Natural values of parameters				Coded values of parameters					y (%)
	Z_1	Z_2	Z_3	Z_4	x_0	x_1	x_2	x_3	x_4	
1	32.5	30	2	1.44	1	−1	−1	−1	−1	67.938
2	77.5	30	2	1.44	1	1	−1	−1	−1	73.223
3	32.5	50	2	1.44	1	−1	1	−1	−1	75.622
4	77.5	50	2	1.44	1	1	1	−1	−1	84.261
5	32.5	30	4	1.44	1	−1	−1	1	−1	81.948
6	77.5	30	4	1.44	1	1	−1	1	−1	86.865
7	32.5	50	4	1.44	1	−1	1	1	−1	88.382
8	77.5	50	4	1.44	1	1	1	1	−1	90.302
9	32.5	30	2	3.4	1	−1	−1	−1	1	81.714
10	77.5	30	2	3.4	1	1	−1	−1	1	87.055
11	32.5	50	2	3.4	1	−1	1	−1	1	92.938
12	77.5	50	2	3.4	1	1	1	−1	1	95.182
13	32.5	30	4	3.4	1	−1	−1	1	1	89.114
14	77.5	30	4	3.4	1	1	−1	1	1	91.414
15	32.5	50	4	3.4	1	−1	1	1	1	91.203
16	77.5	50	4	3.4	1	1	1	1	1	96.528
17	10	40	3	2.42	1	−2	0	0	0	84.622
18	100	40	3	2.42	1	2	0	0	0	91.848
19	55	20	3	2.42	1	0	−2	0	0	84.045
20	55	60	3	2.42	1	0	2	0	0	95.599
21	55	40	1	2.42	1	0	0	−2	0	80.913
22	55	40	5	2.42	1	0	0	2	0	92.404
23	55	40	3	0.46	1	0	0	0	−2	73.973
24	55	40	3	4.38	1	0	0	0	2	93.684
25	55	40	3	2.42	1	0	0	0	0	92.495
26	55	40	3	2.42	1	0	0	0	0	90.384
27	55	40	3	2.42	1	0	0	0	0	91.536
28	55	40	3	2.42	1	0	0	0	0	92.896
29	55	40	3	2.42	1	0	0	0	0	92.367
30	55	40	3	2.42	1	0	0	0	0	90.855
31	55	40	3	2.42	1	0	0	0	0	92.871
32	55	40	3	2.42	1	0	0	0	0	92.324
33	55	40	3	2.42	1	0	0	0	0	92.371
34	55	40	3	2.42	1	0	0	0	0	92.911
35	55	40	3	2.42	1	0	0	0	0	91.511
36	55	40	3	2.42	1	0	0	0	0	92.102

The model coefficients are estimated by the following expression usng “EXCEL” software

$$B = \left[X^T.X\right]^{-1}.[X]^T.Y$$

where B is the column matrix of estimated coefficients; $[X^TX]^{-1}$ the dispersion matrix; $[X]^T$ the transpose matrix of experiments matrix [X] and Y is the column matrix of observations.

The results of this calculation are presented in the following tables:

Constant term	**Linear effects**			
b_0	b_1	b_2	b_3	b_4
92.052	2.101	3.261	3.367	4.835

Interaction and quadratic effects									
b_{12}	b_{13}	b_{14}	b_{23}	b_{24}	b_{34}	b_{11}	b_{22}	$b_{.33}$	b_{44}
0.018	-0.440	-0.347	-1.312	-0.127	-2.193	-1.168	-0.771	-1.562	-2.269

➢ Statistical analysis of the data

From statistical point of view, three tests are required to evaluate the adequacy of the model; Student's t-test which is about the significance of coefficients, R-square test and Fisher tests.

1. Test of coefficients significance

The estimated *t* values by Student test for particular process parameters can be calculated as follows:

$$t_j = \frac{|b_j|}{S_{bj}}$$

$$S_{\mathrm{bj}}^2 = C_{\mathrm{jj}}.S_{\mathrm{rep}}^2$$

where S_{bj}^2 is the coefficients variance; *Cjj* the diagonal terms of $[X^{\mathrm{T}}X]^{-1}$matrix and S_{rep}^2 is the reproducibility variance calculated at the center domain with 12 replicates ($S_{\mathrm{rep}}^2 = 0.67$).

The calculated *t* values are summarized in the following tables.

Constant term	**Linear effects**			
t_0	t_1	t_2	t_3	t_4
389.557	12.574	19.514	20.150	28.934

Interaction and quadratic effects									
t_{12}	t_{13}	t_{14}	t_{23}	t_{24}	t_{34}	t_{11}	t_{22}	t_{33}	t_{44}
0.087	2.152	1.695	6.413	0.622	10.715	8.069	5.327	10.793	15.682

The tabulate t value for 5% level of significance and 11 degrees of freedom (f=n_0-1=12−1 = 11) using the bilateral test of student (appendix 2) $t_{0.05}(11) = 2.201$.

If we compare this value to the calculated ones. we found that all individual effects are significant at 5% of significance level and only the interactions ($x_1.x_2$). ($x_1 .x_3$). ($x_1.x_4$). and ($x_2.x_4$) are not significant. Therefore. they are excluded from the regression equation.

The test of reliability for regression equation has been carried out by Fisher's variance ratio test known as F-test. The F-ratio is given by the following form:

$$F = \frac{S^2_{\text{res}}}{S^2_{\text{rep}}}$$

The following table gives the values S^2_{rep} . S^2_{res} and estimated F for regression equation.

Fisher test for cementation yield

Residual variance, σ^2_{res}	1.63
Replication variance, σ^2_{rep}	0.67
Estimated F value	2.432

The S^2_{res} degree of freedom (f_1 =N−l) and the S^2_{rep} degree of freedom (f = n_0 −1) are 25 and 11 respectively. The tabulated F value (appendix 3) for 5% level of significance is between 2.57 and 2.61. The estimated F value is less than this interval. Hence, it can be concluded that the two variances are equal and the most of the response variation can be explained by the regression. Furthermore, the test of significance of regression ($F_{\text{calculated}} = 98.45 > F_{\text{tabulated}}=2.24$) confirms that the established predicting equation gives an excellent fitting to observed data.

Finally. R^2 value is found to be 96.6% and the table bellow. shows that the difference between the measured and the predicted values do not exceed 3%. T Therefore. all those results indicate that the model can adequately represent the data.

Runs	y (%)	ŷ (%)	Absolut relative error (%)
1	67.938	69.213	1.876
2	73.223	73.415	0.262
3	75.622	78.359	2.620
4	84.261	82.561	2.017

5	81.948	82.957	1.232
6	86.865	87.159	0.339
7	88.382	86.855	1.727
8	90.302	91.057	0.836
9	81.714	83.269	1.903
10	87.055	87.471	0.477
11	92.938	92.415	0.563
12	95.182	96.617	1.508
13	89.114	88.241	0.979
14	91.414	92.443	1.125
15	91.203	92.139	1.026
16	96.528	96.341	0.193
17	92.495	92.052	0.478
18	90.384	92.052	1.846
19	91.536	92.052	0.563
20	92.896	92.052	0.909
21	92.367	92.052	0.341
22	90.855	92.052	1.318
23	92.871	92.052	0.882
24	92.324	92.052	0.294
25	92.371	92.052	0.345
26	92.911	92.052	0.924
27	91.511	92.052	0.591
28	92.102	92.052	0.054
29	84.622	83.178	1.706
30	91.848	91.582	0.290
31	84.045	82.446	1.903
32	95.599	95.49	0.114
33	80.913	79.07	2.277
34	92.404	92.538	0.145
35	73.973	73.306	0.902
36	93.684	92.646	1.108

The regression equation for copper cementation by iron obtained after performing 36 experiments and discarding the insignificant effects is as follows:

$$\hat{y} = 92.052 + 2.101x_1 + 3.261x_2 + 3.367x_3 + 4.835x_4 - 1.312x_2x_3 - 2.193x_3x_4 - \\ -1.168x_1^2 - 0.771x_2^2 - 1.562x_3^2 - 2.269x_4^2$$

- **Discussion**

The regression equation obtained above. shows that initial copper cementation. temperature. pH and flow rate all have an individual influence on the reaction yield of copper cementation. Flow rate (x_4) has the strongest effect on the response since coefficient of x_4 (b_4 = +4.835) is large than the coefficients of the other investigated factors. Positive sign of this coefficient indicates that there is a direct relation between flow rate and reaction yield; in other words. copper recovery increase with increasing flow rate.

The order for factors strength on cementation yield following flow rate was found as pH (x_3). temperature (x_2) and initial copper concentration (x_1); all being positive in sign.

The significance interactions found by the design of experiments for copper cementation yield are essentially between flow rate and solution pH ($x_3.x_4$) and between temperature and pH ($x_2 .x_3$).

➢ Optimization

In this work. the model equation is used to find the direction in which the variables should be changed in order to optimize cementation reaction yield. The corresponding contour plots of the quadratic model are shown in figures (a–c). The figures are drawn in pH–flow rate plan (the most important two factors affecting the response) for various level of temperature (−2. 0. +2) at optimal initial copper concentration ($[Cu^{2+}]_0$ = 75.25 mg/L) using "MATLAB 7.0" software.

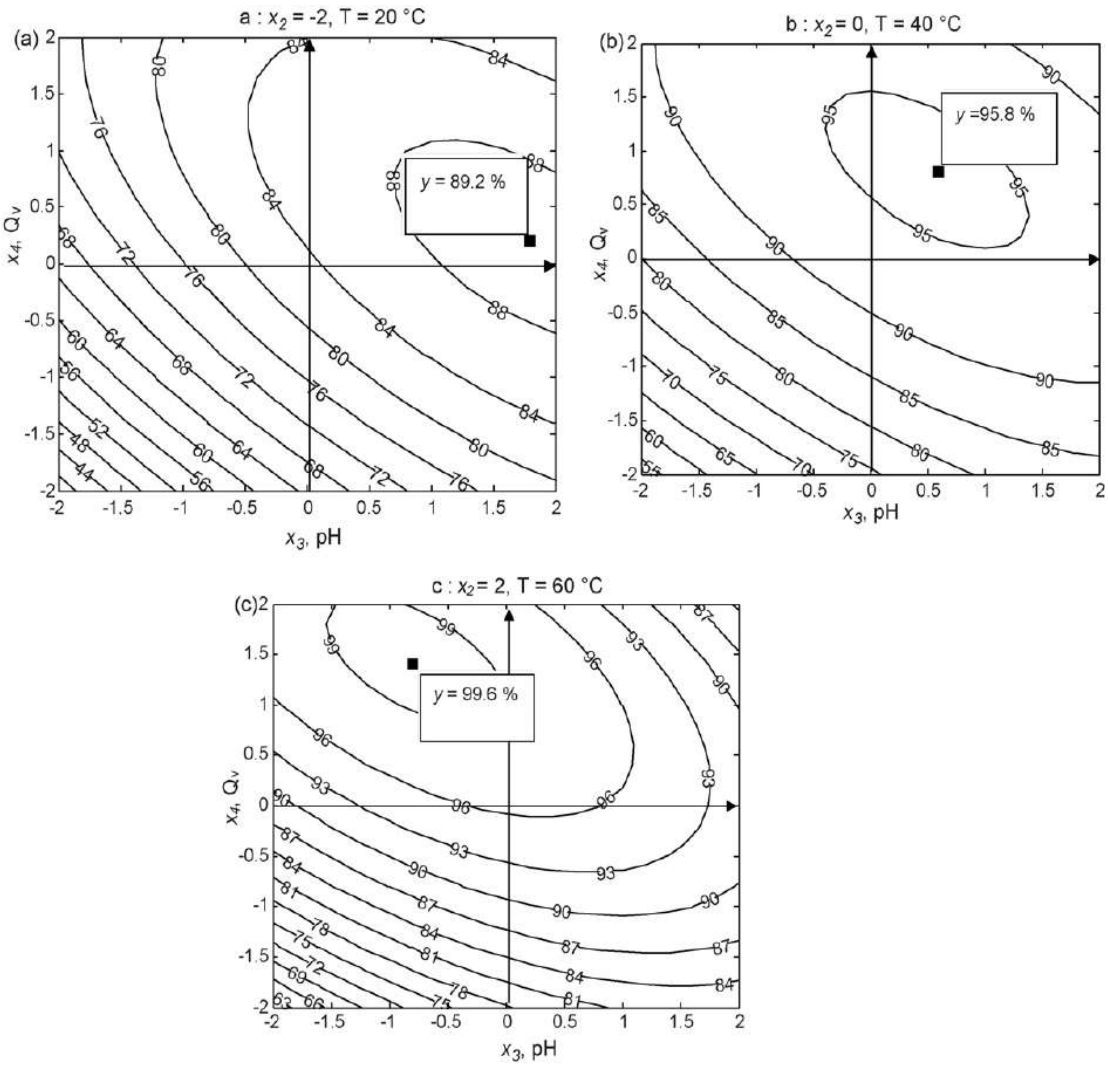
The surface contour plots of mutual interactions between the variables are found to be elliptical. The stationary point or central point is the point at which the slope of the contour is zero in all directions. The coordinates of the central point within the highest contour level in each of these figures will correspond to the optimum values of the respective parameters. The maximum predicted yield is indicated by the surface confined in the smallest curve of the contour diagram.

The analysis of these figures indicates clearly the significance influence of flow rate and its interaction with solution pH. The optimum cementation yield in all conditions (Figures a–c) increases in the direction of the increase in the temperature and it reaches 99.6% cementation yield at high flow rate and low pH values (figure c).

The corresponding conditions of the best cementation yield by deriving the model with respect to each factor are follows:

$$x_1 = 0.90, \text{ corresponding to } [Cu^{2+}]_0 = 75.25 \text{ mg/L};$$
$$x_2 = 2, \text{ corresponding to } T = 60\ ^\circ C;$$
$$x_3 = -0.8, \text{ corresponding to pH } = 2.2;$$
$$x_4 = 1.4, \text{ corresponding to } 3.79 \text{ mL/s}$$

Under economic considerations. 95.8% cementation yield (figure b) can be easily reached by working at moderate temperature (T = 30–40 ˚C) and average solution acidity (pH 3 or 4).



13. Exercises

Exercise 1

The influence of the temperature and the concentration C of a reagent on the yield of a chemical reaction y (in %) is studied.

It was decided to experiment with the temperature between 60°C and 80°C and the concentration between 10 $g.L^{-1}$ and 15$g.L^{-1}$ limiting itself to 2 levels per factor.

A/

- How many experiments should be carried out, knowing that no repetitions are planned?
- In a 2^2 factorial experiment, what are the experimental conditions to be carried out?.

B/ The temperature is called factor A and the concentration factor B.

Using the concept of centered and scaled variables:

- Give the values of A and B at the center of the experimental domain.
- Give the coordinates of A and B at the point (xA = + 0.5 ; xB = - 0.6).

C/ Construct (with EXCEL) the design matrix using coded and uncoded unit.

Results table:

Trials	**Temperature (°C)**	**Concentration ($g.L^{-1}$)**	**Yield Y (%)**
1	60	10	60
2	80	10	70
3	60	15	80
4	80	15	90

D/ Construct (with EXCEL) the effects matrix and calculate all the effects of such design.

Exercise 2

The same study is carried out as that of exercise 1. this time conducted in the presence of a catalyst; the other experimental conditions are unchanged.

Results table:

Trials	**1**	**2**	**3**	**4**
Yield Y (%)	60	70	80	95

1. Construct (with EXCEL) the effects matrix and calculate all the effects.
2. At a temperature of 70°C and a concentration of 12.5 g. L^{-1} (center of the studied domain). it was decided to carry out 6 additional tests. The 6 yields obtained are as follows (in %):

   77.3 – 79.1 – 77.8 – 77.0 – 77.7 – 79.1

   - Calculate the reproducibility variance at the center of the studied domain.
   - Which effects are significant at 5% of significance level (use a t test).

Exercise 3

In a solution usually manufactured at 30°C with stirring (200 rpm) a slight disturbance appears. The experimenter wants to know the cause(s) and thinks that 3 factors can have an influence on this problem.

- Temperature.
- Stirring speed.
- Concentration of an additive usually present at 0.30% (w/v).

The disorder is measured by an opacity index. this index is greater as the solution is cloudy.

It is decided to organize a 2^3 factorial design:

Factors	**Levels**	
	-1	**1**
Temperature	20°C	40°C
Stirring speed	100 rpm	300 rpm
Additive concentration	0.1%	0.5%

1. How many experimental conditions to do for this design?
2. Construct (with EXCEL) the effects matrix and calculate all the effects. knowing that the opacity measurements gave the following results:

N°	1	2	3	4	5	6	7	8
Opacity index	0	4.7	0	11.5	9	14.5	5.1	18.7

3. Give the expression of regression equation found.
4. Interpret the influence of each factor on the response and represent the effects diagrams and the interactions diagrams.

Exercise 4

In the formulation of a certain tablet three variables were considered to be important for the thickness of the tablets. These variables were investigated by a factorial design. The different variables were the amount of stearate lubricant, the amount of active substance and the amount of starch disintegrant.

The experimental domain is shown in Table 1. Experimental design and results are given in Table 2.

Table 1: Variables and experimental domain of the formulation

Variables	Experimental domain		
	(−)-level	0-level	(+)-level
x_1: Amount of stearate (mg)	0.5	1	1.5
x_2: Amount of active substance (mg)	60	90	120
x_3: Amount of starch (mg)	30	40	50

Table 2: Design and responses

Exp. no.	Variables			Thickness (mm)
	x_1	x_2	x_3	y
1	−	−	−	4.75
2	+	−	−	4.87
3	−	+	−	4.21
4	+	+	−	4.26
5	−	−	+	5.25
6	+	−	+	5.46
7	−	+	+	4.72
8	+	+	+	5.22
9	0	0	0	4.86

1. Estimate the effects coefficients of the experimental variables and evaluate their influence.
2. Determine a response model that contains only the probably significant terms. Use this model to estimate the amount of starch that has to be added to 100 mg of active substance to obtain tablets that are 5.00 mm thick.

Exercise 5

The modeling of the E. coli ST 131 activity (expressed by the inhibition zone in mm) of an actinobacteria strain by a centered composite design (CCD) is carried out in order to study the influence of four operating factors namely: starch concentration (X_1). casein concentration (X_2). the incubation time (X_3) and the pH (X_4). The different values of this factors at different levels are presented in the following table:

Factors	**Levels**				
	-2	**-1**	**0**	**+1**	**+2**
Starch (g/L)	2	6	10	14	18
Casein (g/L)	0.1	0.2	0.3	0.4	0.5
Incubation time (Days)	3	5	7	9	11
pH	3.2	5.2	7.2	9.2	11.2

1. Construct (with EXCEL) the design matrix. the effects matrix and find the second order regression equation with verification of its validity.
2. Find the optimum of the factors as well as the optimal antibacterial activity of the strain studied.

- The response results of factorial design trials ($N_{f} = 2^4 = 16$ trials):

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6.50	8.83	9.50	7.33	6	6.53	6.66	6	8.33	6.83	10.33	8.50	7.50	6.83	6	6.33

- The response results of the star design (N_{α}= 2.k= 2×4=8 trials):

17	18	19	20	21	22	23	24
6	6	13.66	8.16	6	6.83	9	9.5

- The responses at the center studied domain (N_0= 6):

25	26	27	28	29	30
9	8.5	8.5	10	6.33	6

Exercise 6

An experiment is performed with two levels of temperature: 25C and 35C. If these are the –1 and +1 levels of temperature. respectively. then:

- Find the coded value that corresponds to 28 °C.
- Determine the temperature that has a coded value of x = +0.6.
-

Exercise 7

To study the influence of temperature (*X_T*) (60°C to 80°C) and concentration of an additive (X_c) (10 $g.L^{-1}$ to 15 $g.L^{-1}$) on the yield of chemical reaction y (%), the predictive model found has a following expression:

$$\hat{Y} = 76{,}25 + 6{,}25x_T + 11{,}25x_c + 1{,}25x_Tx_c$$

- Give the expression of the response according to the natural variables
- What response can we predict for the following experimental condition

Exercise 8

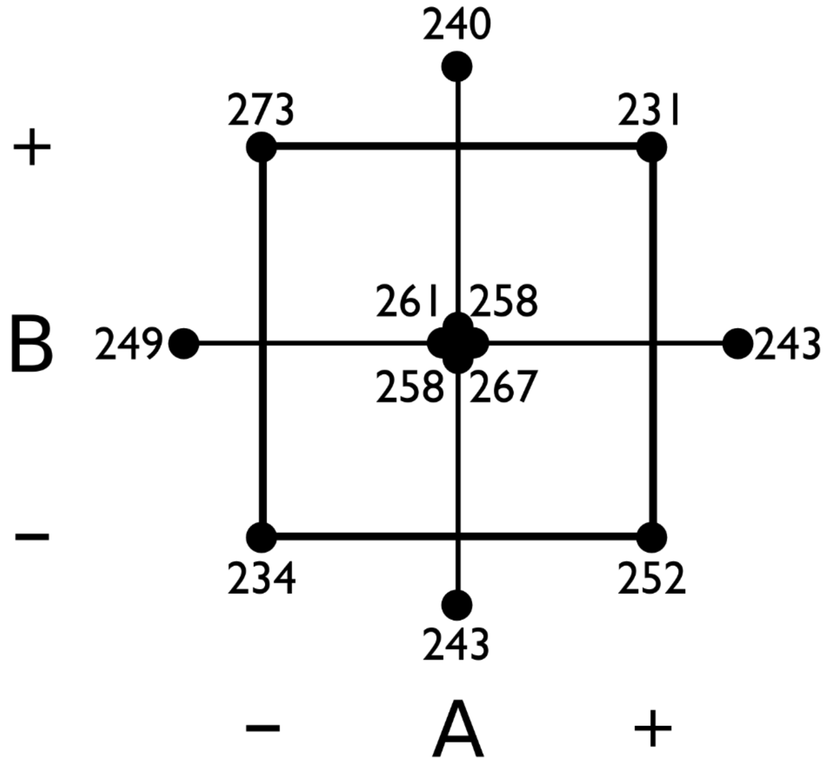
The following diagram shows data from a central composite design. The factors were run at their standard levels, and there were 4 runs at the center point.

1. Give the design matrix and response for this example.
2. Calculate the parameters for a suitable quadratic model in these factors using excel and Minitab software.
3. Draw a response surface plot of **A** *vs* **B** over a suitably wide range beyond the experimental region.
4. Where would you move **A** and **B** if your objective is to increase the response value?

    1. Report your answer in coded units.
    2. Report your answer in real-world units, if the full factorial portion of the experiments were ran at:

        A = *stirrer speed*, 200rpm and 340 rpm

        B = *stirring time*, 30 minutes and 40 minutes

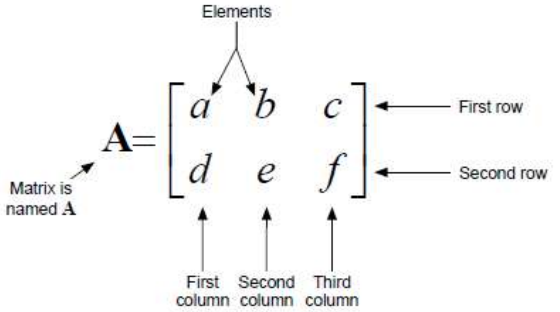

240
273
231
+
B
249
261
258
258
267
243
−
234
252
243
−
A
+

Appendix 1

Definitions of some matrix [4]

- **Matrix**

A matrix is an array made up of elements laid out in lines and columns.



- **Size of a matrix**

An $i \times j$ matrix is a matrix having i rows and j columns. The matrix **A** above is a 2×3 matrix.

- **Square matrix**

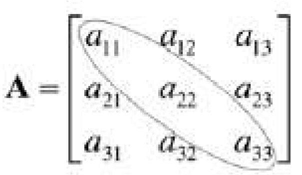
A square matrix has the same number of rows and columns. For example.

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 4 \\ 3 & 5 & 7 \end{bmatrix}$$

is a square matrix. since it is 3×3.

- **Main diagonal of square matrix**

The diagonal of a square matrix is formed by all the elements a_{ii}. where the row number and column number are the same. This main diagonal is shown in the following matrix.



- **Element notation**

The elements from the lines and columns of the matrix are designated by an index. For example. matrix **A** above could be written :

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

a_{13} is the element in the first row. third column. The element a_{22} is the element in the second row. second column.

- **Symmetric matrix**

symmetric matrix has a symmetry around the main diagonal. The elements a_{ij} are equal to the elements a_{ji}. That is. $a_{ij} = a_{ji}$.

For example.

$$\mathbf{S} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 3 & 4 & 1 \end{bmatrix}$$

Identity matrix

An identity matrix is a square matrix whose elements are zeros except for the main diagonal. whose elements are ones.

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- **Diagonal matrix**

A diagonal matrix has all its elements equal to zero except those on the main diagonal.

- **Transpose of a matrix**

The transpose of a matrix **A** is denoted as $\mathbf{A}^T$ and is obtained by inverting the rows and columns of **A**. That is. the first row of **A** becomes the first column of $\mathbf{A}^T$. The second row of **A** becomes the first column of $\mathbf{A}^T$. and so on.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$\mathbf{A}^{\mathbf{T}} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

- **Inverse matrix**

Only square matrices have inverses. The inverse of A is denoted A^{-1}.

The matrix A^{-1} is the inverse of A if their product is the identity matrix.

$$A^{-1}. A = I$$

- **Orthogonal matrices**

A matrix is orthogonal if the scalar product of its columns is all zero. The transpose of an orthogonal matrix is equal to its inverse.

$$\mathbf{A}^{T} = \mathbf{A}^{-1}$$

$\mathbf{A}^{T}$: Transpose matrix of A

$\mathbf{A}^{-1}$: Inverse matrix of A

- **Hadamard matrices**

There are some square matrices where the elements are either +1 or –1 such that:

$$\mathbf{X}^{\mathbf{T}}\mathbf{X} = n\mathbf{I}$$

I: Identity matrix

Appendix 2

In the t-test table, the significant values are determined for degrees of freedom(df) to the probabilities of t-distribution, *α*. [1]

t-test table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Appendix 3

Fisher Snedecor table for α = 0.05 [1]

γ_1 = degrees of freedom in numerator
γ_2 = degrees of freedom in denominator

v_1 → v_2 ↓	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.4	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.7	8.66	8.64	8.62	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.0	5.96	5.91	5.86	5.80	5.77	5.75	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.8	2.75	2.69	2.62	2.54	2.51	2.47	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.7	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.84	2.42	2.38	2.31	2.23	2.16	2.11	2.07	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.9	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.2	2.13	2.06	1.97	1.93	1.88	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.0

References

[1] Zivorad R. Lazic. (2004) **"Design of Experiments in Chemical Engineering".** Wiley-VCH edition. Weinheim ISBN: 3-527-31142-

[2] Jiju A**. (**2003) **"Design of experiments for engineers and scientists"** Butterworth-Heinemann. Linacre House. Jordan Hill. Oxford OX2 8DP

[3] Durakovic B. **(**2017) **"Design of Experiments Application. Concepts. Examples: State of the Art**". *Periodicals of Engineering and Natural Sciences ISSN 2303-4521 Vol 5. No 3. pp. 421–439.*

[4] Goupy J. Lee Creighton. (2007) **"Introduction to Design of Experiments with JMP Examples".** Third Edition. SAS Institute Inc.. Cary. NC. USA ISBN 978-1-59994-422-7.

[5] Vanaja K. Shobha Rani R.H. (2007) "**Design of experiments: Concept and applications of Plackett Burman design**". *Clinical Research and Regulatory Affairs.* 24(1):1–23.

[6] Dean A. Voss D. (1999) **"Design and Analysis of Experiments".** Springer- Verlag. New York. Inc.

[7] Florentinus Dika, Octa Riswanto. Abdul Rohman. Suwidjiyo Pramono. Sudibyo Martono. (2019) **"Application of response surface methodology as mathematical and statistical tools in natural product research"** *Journal of Applied Pharmaceutical Science Vol. 9(10). pp 125-133.*

[8] Box GEP. Hunter JS. Hunter WG. (2005) **"Factorial designs at two levels"**. Second edition. Hoboken. NJ: John Wiley and Sons. Inc. 173-234

[9] Lundstedt. Seifert E. Abramo L. Thelin Be. Nystrom A. Pettersen J. Bergman R. (1998) "**Experimental design and optimization**" *Chemometrics and Intelligent Laboratory Systems* 42. 3–40.

.

[10] Selvanathan L. Govindhan S. Dharumadurai Dhanasekaran. (2017) "**Response surface methodology: A nonconventional statistical tool to maximize the throughput of Streptomyces species biomass and their bioactive metabolites"**. C*ritical Reviews in Microbiology. VOL. 43. NO. 5. 567–582.*

[11] Bevilacqua A. Corbo M.R. Sinigaglia. M. (2010) **"Design of experiments: a powerful tool in food microbiology"** *Agricultural and Food Sciences.*

[12]: Goupy J. (2009) " **Plans d'expériences :Manuel d'utilisation du logiciel Minitab"**

[13]: Mathews P G.. 2005 "**Design of experiments with MINITAB"**. American Society for Quality. Quality Press. Milwaukee 53203.

[14] : Goupy J. (1999). « **Plans d'expériences pour surfaces de réponse** ». Dunod. Paris.

[15] : Goupy J. (2005) « **Pratiquer les Plans d'Expériences** ». Dunod. Paris.

[16]: Goupy J. Creighton L. (2009) « **Introduction aux Plans d'expériences** ». Dunod. Paris. Quatrième édition.

[17]: Djinni I, Djoudi W. Souagui S. Rabia F. Rahmouni S. Mancini I. Kecha.M (2018). « **Streptomyces thermoviolaceus SRC3 strain as a novel source of the antibiotic adjuvant streptazolin: A statistical approach toward the optimized production**». *Journal of Microbiological Methods 148, 161–168.*

[18]: Djoudi W. Aissani-Benissad F. Bourouina-Bacha S. (2007) «Optimization of copper cementation process by iron using central composite design experiments ». *Chemical Engineering Journal 133, 1–6.*