1 Introduction (Level 1)

Objective

In this chapter, we shall discuss where and how design of experiments, DOE, is used by industry today. Regardless of whether the experimental work takes place in the laboratory, the pilot plant, or the full-scale production plant, design of experiments is useful for three primary experimental objectives, screening, optimization and robustness testing. We shall discuss these three objectives and introduce one general example corresponding to each objective. These three examples will accompany us throughout the course.

We shall also describe another industrial application, the CakeMix application, which will help us to highlight some of the key elements involved in DOE. This example shows how changes in some important factors, that is, ingredients, can be linked to the changes in the response, that is, taste. In this way, an understanding is gained of how one must proceed to modify the amounts of the ingredients to improve the taste. The aim of this chapter is also to overview some of the most commonly employed DOE design families and point out when they are meaningful. The chapter ends with some arguments emphasizing the main benefits of adhering to the DOE methodology.

When and where DOE is useful

Design of experiments, DOE, is used in many industrial sectors, for instance, in the development and optimization of manufacturing processes. Typical examples are the production of wafers in the electronics industry, the manufacturing of engines in the car industry, and the synthesis of compounds in the pharmaceutical industry. Another main type of DOE-application is the optimization of analytical instruments. Many applications are found in the scientific literature describing the optimization of spectrophotometers and chromatographic equipment.

Usually, however, an experimenter does not jump directly into an optimization problem; rather initial screening experimental designs are used in order to locate the most fruitful part of the experimental region in question. Other main types of application where DOE is useful is robustness testing and mixture design. The key feature of the latter application type is that all factors sum to 100%.

Areas where DOE is used in industrial research, development and production:

- optimization of manufacturing processes
- optimization of analytical instruments
- screening and identification of important factors

- robustness testing of methods
- robustness testing of products
- formulation experiments

What is DOE?

One question which we might ask ourselves at this stage is *what is design of experiments*? DOE involves making a set of experiments representative with regards to a given question. The way to do this is, of course, problem dependent, and in reality the shape and complexity of a statistical experimental design may vary considerably. A common approach in DOE is to define an interesting standard reference experiment and then perform new, representative experiments around it (see Figure 1.1). These new experiments are laid out in a symmetrical fashion around the standard reference experiment. Hence, the standard reference experiment is usually called the *center-point*.
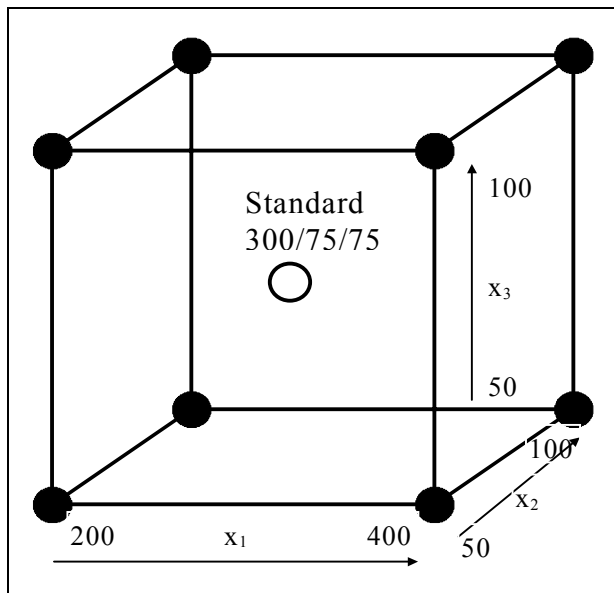

Figure 1.1: *A symmetrical distribution of experimental points around a center-point experiment.*

In the given illustration, the standard operating condition was used as the center-point. It prescribed that the first factor (x_1) should be set at the value 300, the second factor (x_2) at 75, and the third factor (x_3) at 75. In the next step, these three factors were varied according to the cubic pattern shown in Figure 1.1. This cubic pattern arises because the three factors are varied systematically around the center-point experiment. Thus, the first factor, x_1, is tested at a level slightly below the center-point, the value 200, and at a level slightly above the center-point, the value 400. A similar reasoning applies to factors x_2 and x_3.

Moreover, at a later stage in the experimental process, for instance, at an optimization step, already performed screening experiments may be used to predict a suitable reference experiment for an optimization design.

In the next three sections, we will introduce three representative DOE applications, which will accompany us in the course.

General Example 1: Screening

Screening is used at the beginning of the experimental procedure. The objective is (i) to explore many factors in order to reveal whether they have an influence on the responses, and (ii) to identify their appropriate ranges. Consider the laser welding material displayed in Figure 1.2. This is a cross-section of a plate heat-exchanger developed and manufactured by Alfa Laval Thermal.

Figure 1.2: A cross-section of a plate heat-exchanger developed and manufactured by Alfa Laval Thermal.

In this application, the influence of four factors on the shape and the quality of the laser weld was investigated. The four factors were *power of laser*, *speed of laser*, *gas flow at nozzle*, and *gas flow at root* (underside) of the welding. The units and settings of low and high levels of these factors are seen in Figure 1.3. The experimenter measured three responses to characterize the shape and the quality of the weld, namely *breakage of weld*, *width of weld*, and *skewness of weld*. These are summarized in Figure 1.4. The aim was to obtain a persistent weld (high value of breakage), of a well-defined width and low skewness.

	Name	Abbr.	Units	Type	Use	Settings
1	Power	Po	kW	Quantitative	Controlled	2.15 to 4.15
2	Speed	Sp	m/min	Quantitative	Controlled	1.88 to 5
3	NozzleGas	No	l/min	Quantitative	Controlled	27 to 36
4	RootGas	Ro	l/min	Quantitative	Controlled	27 to 42

Responses		
	Name	Abbr.	Units
1	Breakage	Br	MPa
2	Width	Wi	mm
3	Skewness	Sk	-

Figure 1.3: (left) The four varied factors of General Example 1.
Figure 1.4: (right) The three measured responses of General Example 1.

In the first stage, the investigator carried out eleven experiments. During the data analysis, however, it soon became apparent that it was necessary to upgrade the initial screening design with more experiments. Thus, the experimenter conducted another set of eleven experiments, selected to well supplement the first series. We will provide more details later.

In summary, with a screening design, the experimenter is able to extract a *yes* or *no* answer with regard to the influence of a particular factor. Information is also gained about how to modify the settings of the important factors, to possibly further enhance the result. Screening designs need few experiments in relation to the number of factors.

General Example 2: Optimization

Optimization is used after screening. The objective is (i) to predict the response values for all possible combinations of factors within the experimental region, and (ii) to identify an optimal experimental point. However, when several responses are treated at the same time, it is usually difficult to identify a single experimental point at which the goals for all responses are fulfilled, and therefore the final result often reflects a compromise between partially conflicting goals.

Our illustration of the optimization objective deals with the development of a new truck piston engine, studying the influence on fuel consumption of three factors, *air mass used in combustion*, *exhaust gas re-circulation*, and *timing of needle lift*. The settings of these factors are shown in Figure 1.5. Besides monitoring the *fuel consumption*, the investigator measured the levels of *NOx* and *Soot* in the exhaust gases. These responses are summarized in Figure 1.6. The goal was to minimize fuel consumption while at the same time not exceeding certain stipulated limits of NOx and Soot. The relationships between the three factors and the three responses were investigated with a standard 17 run optimization design. We will provide more details in Chapter 15.

	Factors					
	Name	Abbr.	Units	Type	Use	Settings
1	Air	Air	kg/h	Quantitative	Controlled	240 to 284
2	EGR%	EGR	%	Quantitative	Controlled	6 to 12
3	NeedleLift	NL	°BTDC	Quantitative	Controlled	-5.78 to 0

	Responses		
	Name	Abbr.	Units
1	Fuel	Fu	mg/st
2	NOx	NO	mg/s
3	Soot	So	mg/s

Figure 1.5: (left) The three varied factors of General Example 2.
Figure 1.6: (right) The three measured responses of General Example 2.

In summary, with an optimization design the experimenter is able to extract detailed information regarding how the factors combine to influence the responses. Optimization designs require many experiments in relation to the number of investigated factors.

General Example 3: Robustness testing

The third objective is robustness testing, and it is applied as the last test just before the release of a product or a method. When performing a robustness test of a method – as in the example cited below – the objective is (i) to ascertain that the method is robust to small fluctuations in the factor levels, and, if non-robustness is detected, (ii) to understand how to alter the bounds of the factors so that robustness may still be claimed.

To portray a typical robustness test of an analysis method, we have selected an application taken from the pharmaceutical industry, which deals with a high-performance liquid chromatography (HPLC) system. Five factors, of which four were quantitative and one qualitative, were examined. These factors were *amount of acetonitrile* in the mobile phase, *pH*, *temperature*, *amount of the OSA counterion* in the mobile phase, and *type of stationary phase* (column). These five factors, summarized in Figure 1.7, were investigated using a design of 12 experiments. To describe the chromatographic properties of the HPLC-system, three responses were recorded, that is, the *capacity factor k_1* of analyte 1, the *capacity factor k_2* of analyte 2, and the *resolution Res1* between these two analytes (Figure 1.8).

	Name	Abbr.	Units	Type	Use	Settings
1	AcN	Ac	%	Quantitative	Controlled	25 to 27
2	pH	pH		Quantitative	Controlled	3.8 to 4.2
3	Temp	Te	°C	Quantitative	Controlled	18 to 25
4	OSA	OS	mM	Quantitative	Controlled	0.09 to 0.11
5	Column	Co		Qualitative	Controlled	ColA,ColB

	Name	Abbr.	Units
1	k1	k1	
2	k2	k2	
3	Res1	Re1	

Figure 1.7: *(left) The five investigated factors of General Example 3.*
Figure 1.8: *(right) The three registered responses of General Example 3.*

In HPLC, capacity factors measure the retention of compounds, and resolution the separation between compounds. In the present case, the resolution response was the main interest and required to be robust. More information regarding this example will be given at a later stage (Chapter 17).

In summary, with a robustness testing design, it is possible to determine the sensitivity of the responses to small changes in the factors. Where such minor changes in the factor levels have little effect on the response values, the analytical system is determined to be robust.

The CakeMix application

We will now concentrate on the CakeMix application, which is helpful in illustrating the key elements of DOE. This is an industrial pilot plant application in which the goal was to map a process producing a cake mix to be sold in a box, for instance, at a supermarket or shopping mall. On the box there will be instructions on how to use the cake mix, and these will include recommendations regarding baking temperature and time.

There are many parameters which might affect the production of a cake mix, but in this particular investigation we will only be concerned with the recipe. The experimental objective was screening, to determine the impact of three cake mix ingredients on the taste of the resulting cake. The first varied factor (ingredient) was *Flour*, the second S*hortening (fat)*, and the third *Eggpowder*. In reality, the investigated cake mix contained other ingredients, like sugar and milk, but to keep things simple only three ingredients were varied.

Firstly, the standard operating condition, the center-point, for the three factors was defined, and to do this a recommended cake mix composition was used. The chosen center-point corresponded to 300g Flour, 75g Shortening, and 75g Eggpowder. Secondly, the low and the high levels of each factor were specified in relation to the center-point. It was decided to vary Flour between 200 and 400g, Shortening between 50 and 100g, and Eggpowder between 50 and 100g. Thirdly, a standard experimental plan with eleven experiments was

created. This experimental design is shown in Figure 1.9, and in this table each row corresponds to one cake.

Cake Mix Experimental Plan

Cake No	Flour	Shortening	Egg Powder	Taste
1	200	50	50	3.52
2	400	50	50	3.66
3	200	100	50	4.74
4	400	100	50	5.20
5	200	50	100	5.38
6	400	50	100	5.90
7	200	100	100	4.36
8	400	100	100	4.86
9	300	75	75	4.73
10	300	75	75	4.61
11	300	75	75	4.68

Factors		
	Levels (Low/High)	Standard condition
Flour	200 g / 400 g	300 g
Shortening	50 g / 100 g	75 g
Egg powder	50 g / 100 g	75 g
Response: Taste of the cake, obtained by averaging the judgment of a sensory panel.		

Figure 1.9: The experimental plan of the CakeMix application.

For each one of the eleven cakes, a sensory panel was used to determine how the cake tasted. The response value used was the average judgment of the members of the sensory panel. A high value corresponds to a good-tasting cake, and it was desired to get as high value as possible. Another interesting feature to observe is the repeated use of the standard cake mix composition in rows 9-11. Such repeated testing of the standard condition is very useful for determining the size of the experimental variation, known as the replicate error.

Apart from listing all the experiments of the design as a table, it is also instructive to make a graphical presentation of the design. In the CakeMix application, a cube is a good tool to visualize the design and thus better understand its geometry. This is shown in Figure 1.10.
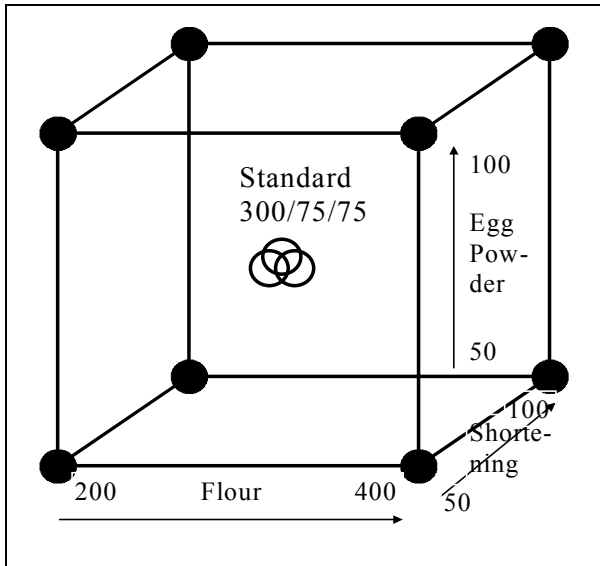
Figure 1.10: *A geometrical representation of the CakeMix experimental protocol.*

After the completion of an experimental plan, one must analyze the data to find out which factors influence the responses. Usually, this is done by fitting a polynomial model to the data. In the CakeMix application, the performed experimental design supports the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon,$$

where y is the response, x's the three ingredients, β_0 the constant term, β's the model parameters, and ε the residual response variation not explained by the model. The model concept, the philosophy of modelling, and model adequacy are further discussed in Chapter 3.

The aim of the data analysis is to estimate numerical values of the model parameters, the so called regression coefficients, and these values will indicate how the three factors influence the response. Such regression coefficients are easy to overview when plotted in a bar chart, and the results for the cake mix data are displayed in Figure 1.11. We see that the strongest term is the two-factor interaction between Shortening and Eggpowder.

Normally, one uses a regression coefficient plot to *detect* strong interactions, but response contour plots to *interpret* their meaning. The response contour plot displayed in Figure 1.12 shows how Taste varies as a function of Shortening and Eggpowder, while keeping the amount of Flour fixed at its high level. Apparently, to obtain a cake with as high "taste" as possible, we should stay in the upper left-hand corner, i.e., use much Flour, much Eggpowder and little Shortening.
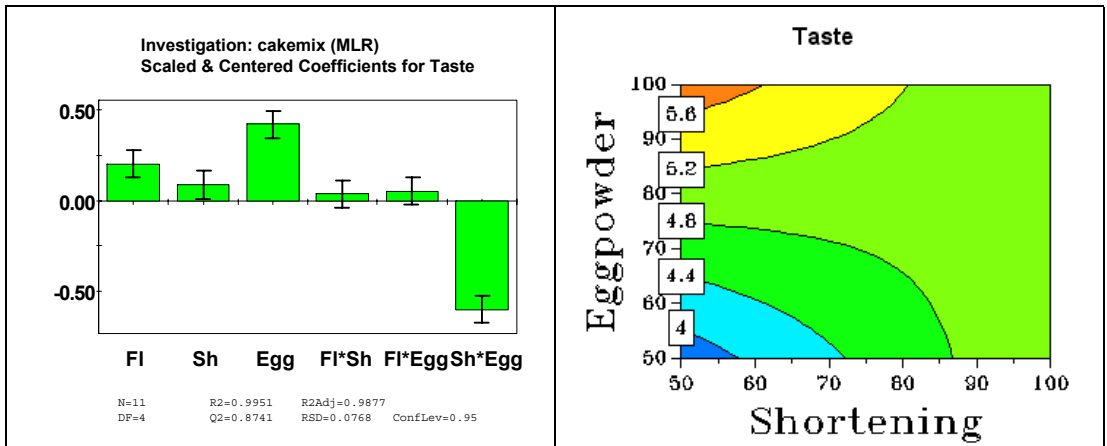
Figure 1.11: *(left) Regression coefficient plot of CakeMix regression model.*
Figure 1.12: *(right) Response contour plot of taste as a function of Eggpowder and Shortening.*

The type of response contour plot displayed in Figure 1.12 is useful for decision making – it suggests what to do next, that is, where to continue experimentally. Thus, we have converted the experimental data into an informative map with quantitative information about the modelled system. This is actually the essence of DOE, to plan informative experiments, to analyze the resulting data to get a good model, and from the model create meaningful maps of the system.

Examples of statistical designs

As we have seen, the DOE concept may be viewed as a framework for experimental planning. We shall here briefly overview a few basic designs of this framework, which are used to deal with the three major experimental objectives, and point out their common features and differences. Figure 1.13 provides a summary of the designs discussed.

The first row of Figure 1.13 shows complete, or full, factorial designs for the investigation of two and three factors. These are screening designs, and are called *full* because all possible corners are investigated. The snowflake in the interior part depicts replicated center-point experiments carried out to investigate the experimental error. Usually, between 3-5 replicates are made. The second row in the figure also shows a screening design, but one in which only a fraction of all possible corners have to be carried out. It belongs to the *fractional factorial* design family, and this family is extensively deployed in screening. Fractional factorial designs are also used a lot for robustness testing. The last row of Figure 1.13 displays designs originating from the *composite* design family, which are used for optimization. These are called composite designs because they consist of the building blocks, corner (factorial) experiments, replicated center-point experiments, and axial experiments, the latter of which are denoted with open circles.
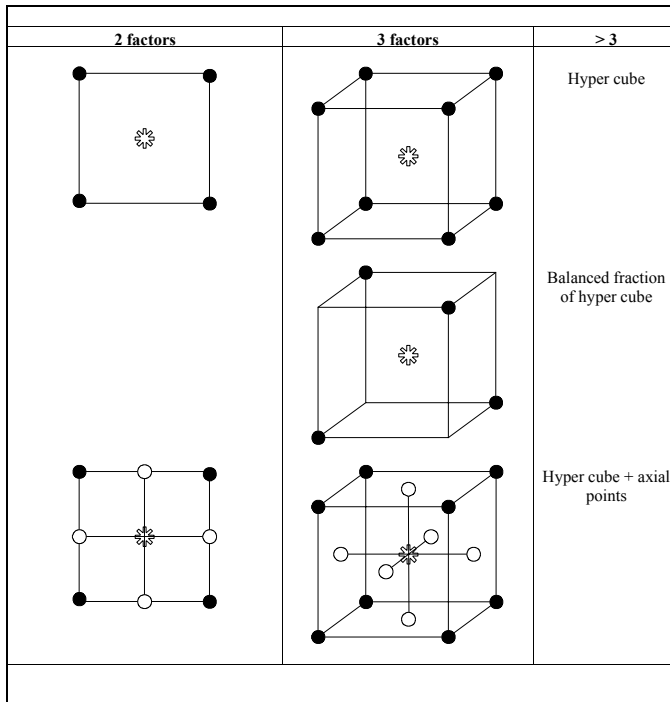
2 factors	3 factors	> 3
		Hyper cube
		Balanced fraction of hyper cube
		Hyper cube + axial points

Figure 1.13: Examples of full factorial, fractional factorial, and composite designs used in DOE.

Quiz

Please complete the following statements:

In a screening investigation, the idea is to investigate many factors and their influences on the responses. This is done by using comparatively ….. experiments in relation to the number of varied factors. (many/few)

In an optimization study, one wants to obtain detailed information about how a few factors combine in regulating the responses. This is accomplished by making comparatively ….. experiments in relation to the number of factors. (many/few)

In robustness testing of, for instance, an analytical method, the aim is to explore how sensitive the responses are to small changes in the factor settings. Ideally, a robustness test should show that the responses are *not* sensitive to small fluctuations in the factors, that is, the results are the same for all experiments. Since the expected result is similarity for all runs, a robustness testing design may well be done with …. experiments per varied factor. (very few/very many)

DOE consists of a few well-defined steps. First the experimenter has to select the … and define an …. which can be used to solve the problem. Then the selected experiments have to be carried out and the resulting data analyzed. The data analysis will give a model, which may be interpreted. To understand which factors are most important and to examine whether there are interactions, it is instructive to look at a …. plot. To even better understand the modelled system, one may convert the model information into a map, a … plot, which

explicitly shows where to continue experimentally. (experimental objective/experimental plan/coefficient/response contour)

Benefits of DOE

The great advantage of using DOE is that it provides an *organized* approach, with which it is possible to address both simple and tricky experimental problems. The experimenter is encouraged to select an appropriate experimental objective, and is then guided to devise and perform a *set* of experiments, which is adequate for the selected objective. Although the experimenter may feel some frustration about having to perform a series of experiments, experience shows that DOE requires fewer experiments than any other approach. Since these few experiments belong to an experimental plan, they are mutually connected and thereby linked in a logical and theoretically favorable manner. Thus, by means of DOE, one obtains more useful and more precise information about the studied system, because the joint influence of all factors is assessed. After checking the model adequacy, the importance of the factors is evaluated in terms of a plot of regression coefficients, and interpreted in a response contour plot. The latter type of plot constitutes a map of the system, with a familiar geometrical interpretation, and with which it is easy to decide what the next experimental step ought to be.

Summary

Design of experiments is useful in the laboratory, the pilot plant and full-scale production, and is used for any experimental objective, including screening, optimization, and robustness testing. We have introduced three general examples - the laser welding case, the truck engine study, and the HPLC robustness problem - which will be used to illustrate these three objectives. In addition, the CakeMix application was outlined for the purpose of overviewing some of the key elements involved in DOE. By conducting an informative set of eleven experiments, it was possible to create a meaningful response contour plot, showing how to modify the cake mix recipe to achieve even better tasting cakes. Finally, this chapter ended with a discussion of the main benefits of DOE.