CHAPTER 7
DESIGN OF EXPERIMENTS

The choice of an experimental design depends on the objectives of the experiment and the number of factors to be investigated. Types of designs are listed here according to the experimental objective they meet.

7.1 Comparative Objective

If there are several factors under investigation, but the primary goal of the experiment is to make a conclusion about one a-priori important factor, (in the presence of, and/or in spite of the existence of the other factors), and the question of interest is whether or not that factor is "significant", (i.e., whether or not there is a significant change in the response for different levels of that factor), then we have a comparative problem and we need a comparative design solution.

7.2 Screening Objective

The primary purpose of the experiment is to select or *screen out* the few important main effects from the many less important ones. These screening designs are also termed main effects designs.

7.3 Response Surface (method) Objective

The experiment is designed to allow us to estimate interaction between factors and even quadratic effects, and therefore give us an idea of the (local) shape of the response surface we are investigating. For this reason, they are termed response surface method (RSM) designs. RSM designs are used to:

1. Find improved or optimal process settings
2. Troubleshoot process problems and weak points
3. Make a product or process more robust against external and non-controllable influences. "Robust" means relatively insensitive to these influences.

7.3.1 Optimizing responses when factors are proportions of a mixture objective

If we have factors that are proportions of a mixture and we want to know what the "best" proportions of the factors are so as to maximize (or minimize) a response, then we need a mixture design.

7.3.2 Optimal fitting of a regression model objective

If we want to model a response as a mathematical function (either known or empirical) of a few continuous factors and we desire "good" model parameter estimates (i.e., unbiased and minimum variance), then we need a regression design. Selection of designs is summarized in the following table.

TABLE 7.1 Design Selection Guideline[7]

Number of Factors	Comparative Objective	Screening Objective	Response Surface Objective
1	1-factor completely randomized design	–	–
2 – 4	Randomized block design	Full or fractional factorial	Central composite or Box-Behnken
5 or more	Randomized block design	Fractional factorial or Plackett-Burman	Screen first to reduce number of factors

Choice of a design from within these various types depends on the amount of resources available and the degree of control over making wrong decisions that the experimenter desires. It is a good idea to choose a design that requires somewhat fewer runs than the budget permits, so that center point runs can be added to check for curvature in a 2-level screening design and backup resources are available to redo runs that have processing mishaps.

7.4 Response Surface Design

Earlier, we described the response surface method (RSM) objective. Under some circumstances, a model involving only main effects and interactions may be appropriate to describe a response surface when

1. Analysis of the results reveals no evidence of "pure quadratic" curvature in the response of interest (i.e., the response at the center approximately equals the average of the responses at the factorial runs).
2. The design matrix originally used includes the limits of the factor settings available to run the process.

In other circumstances, a complete description of the process behavior might require a quadratic or cubic model:

Quadratic

$$\begin{aligned}\hat{y} = b_0 &+ b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 \\ &+ b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2\end{aligned} \tag{7.1}$$

Cubic

$$\begin{aligned}\hat{y} = \text{quadratic model} &+ b_{123}x_1x_2x_3 + b_{112}x_1^2x_2 + b_{113}x_1^2x_3 \\ &+ b_{122}x_1x_2^2 + b_{133}x_1x_3^2 + b_{223}x_2^2x_3 + b_{233}x_2x_3^2 + b_{111}x_1^3 \\ &+ b_{222}x_2^3 + b_{333}x_3^3\end{aligned} \tag{7.2}$$

These are the full models, with all possible terms; rarely would all of the terms be needed in an application.

If the experimenter has defined factor limits appropriately and/or taken advantage of all the tools available in multiple regression analysis (transformations of responses and factors, for example), then finding an industrial process that requires a third-order model is highly unusual. Therefore, we will only focus on designs that are useful for fitting quadratic models.

Figures 7.1 to 7.4 identify the general quadratic surface types that an investigator might encounter.

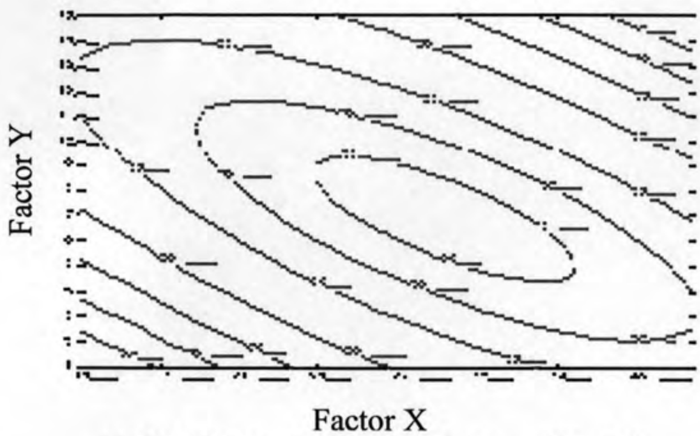

FIGURE 7.1 A response surface "Peak"

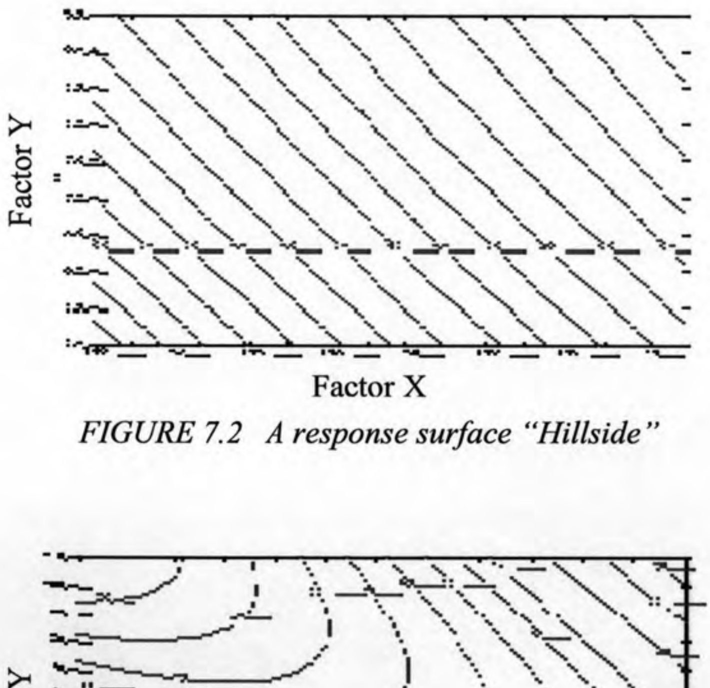

FIGURE 7.2 A response surface "Hillside"

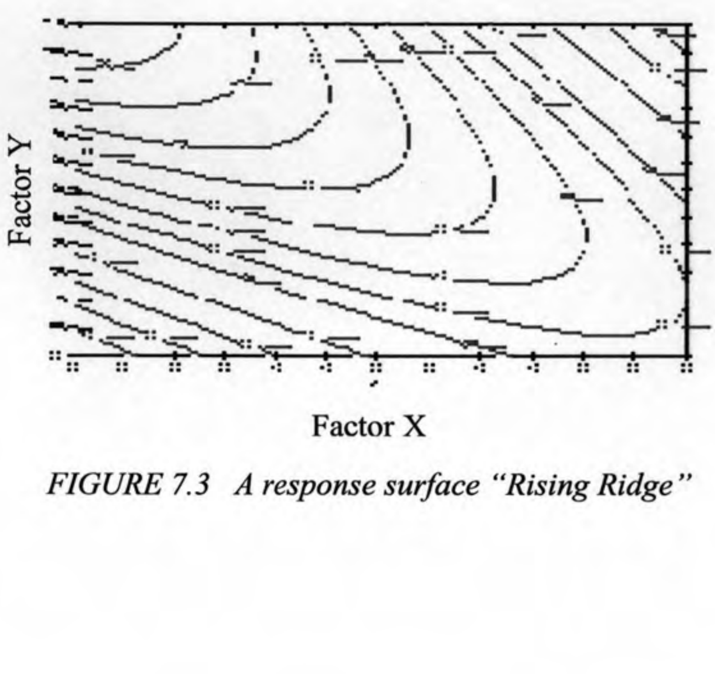

FIGURE 7.3 A response surface "Rising Ridge"

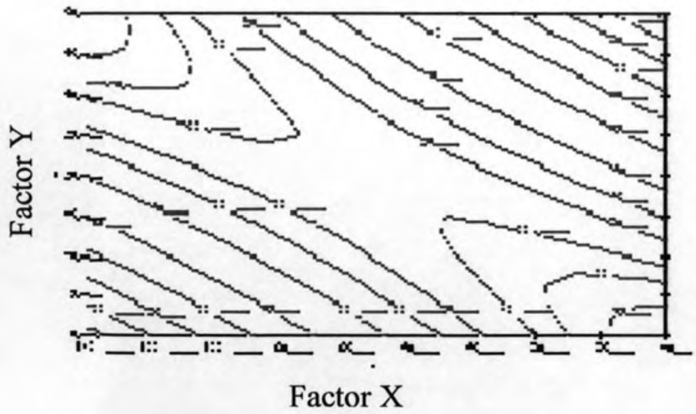

FIGURE 7.4 A response surface "Saddle"

7.4.1 Factor Levels for Higher-Order Designs

Figures 7.5 through 7.7 illustrate possible behaviors of responses as functions of factor settings. In each case, assume the value of the response increases from the bottom of the figure to the top and that the factor settings increase from left to right.

FIGURE 7.5 Linear function

FIGURE 7.6 Quadratic function

FIGURE 7.7 Cubic function

If a response behaves as in Figure 7.5, the design matrix to quantify that behavior needs only contain factors with two levels – low and high. This model is a basic assumption of simple two-level factorial and fractional factorial designs. If a response behaves as in Figure 7.6, the minimum number of levels required for a factor to quantify that behavior is three (low, medium, high). One might logically assume that adding center points to a two-level design would satisfy that requirement, but the arrangement of the treatments in such a matrix confounds all quadratic effects with each other.

While a two-level design with center points cannot estimate individual pure quadratic effects, it can detect them effectively. A solution to creating a design matrix that permits the estimation of simple curvature as shown in Figure 7.6 would be to use a three-level factorial design. Table 7.2 explores that possibility.

Finally, in more complex cases such as those illustrated in Figure 7.7, the design matrix must contain at least four levels of each factor to characterize the behavior of the response adequately.

Table 7.2 3-Level Factorial Designs[7]

Number of Factors	**Treatment Combinations 3^k Factorial**	**Number of Coefficients Quadratic Empirical Model**
2	9	6
3	27	10
4	81	15
5	243	21
6	729	28

Two-level factorial designs quickly become too large for practical application as the number of factors investigated increases. This problem was the motivation for creating `fractional factorial' designs. Table 7.2 shows that the number of runs required for a 3^k factorial becomes even more unacceptable than for 2^k designs. The last column in Table 7.2 shows the number of terms present in a quadratic model for each case.

With only a modest number of factors, the number of runs is very large, even an order of magnitude greater than the number of parameters to be estimated when k (number of factors) is not small. For example, the absolute minimum number of runs required to estimate all the terms present in a four-factor quadratic model is 15: the intercept term, 4 main effects, 6 two-factor interactions, and 4 quadratic terms. The corresponding 3^k design for $k = 4$ requires 81 runs.

Considering a fractional factorial designs at three levels is a logical step, given the success of fractional designs when applied to two-level designs. Unfortunately, the alias structure for the three-level fractional factorial designs is considerably more complex and harder to define than in the two-level case. Additionally, the three-level factorial designs suffer a major flaw in their lack of rotatability.

7.4.2 Rotatability of Designs

In a rotatable design, the variance of the predicted values of y is a function of the distance of a point from the center of the design and is not a function of the direction the point lies from the center. Before a study begins, little or no knowledge may exist about the region that contains the optimum response. Therefore, the experimental design matrix should not bias an investigation in any direction.

In a rotatable design, the contours associated with the variance of the predicted values are concentric circles. Figures 7.8 and 7.9 illustrate a three-dimensional plot and contour plot, respectively, of the `information function' associated with a 3^2 design. The information function is:

$$\frac{1}{V(\hat{y})}$$

with V denoting the variance (of the predicted value $\hat{y}$).

Each figure clearly shows that the information content of the design is not only a function of the distance from the center of the design space, but also a function of direction. Figures 7.10 and 7.11 are the corresponding graphs of the information function for a rotatable quadratic design. In each of these figures, the value of the information function depends only on the distance of a point from the center of the space.

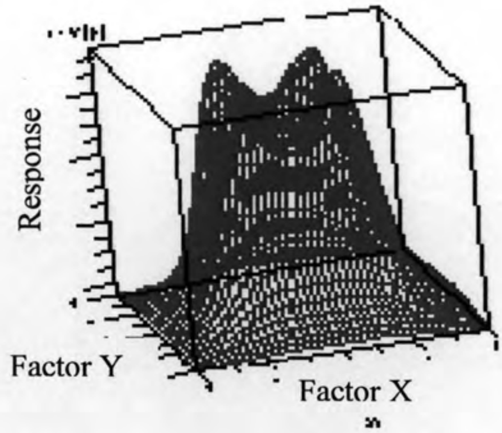

FIGURE 7.8 Three-dimensional illustration for the information function of a 3^2 design

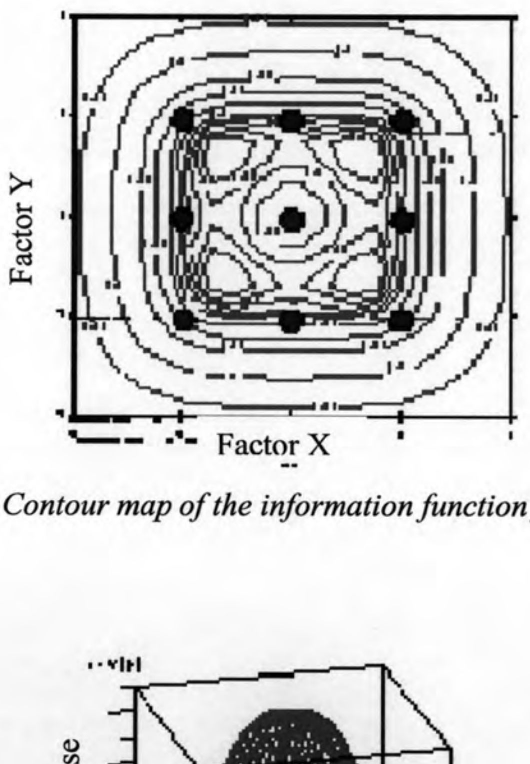

FIGURE 7.9 Contour map of the information function for a 3^2 design

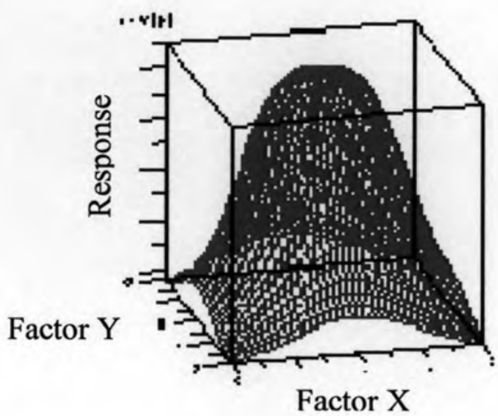

FIGURE 7.10 Three-dimensional illustration of the information function for a rotatable quadratic design for two factors

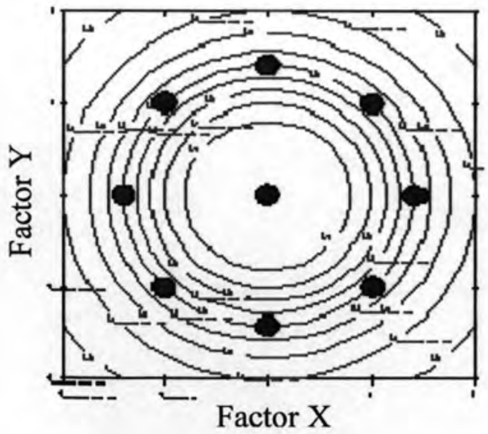

FIGURE 7.11 Contour map of the information function for a rotatable quadratic design for two factors

7.4.3 Box-Wilson Central Composite Designs

A Box-Wilson Central Composite Design, commonly called 'a central composite design,' contains an imbedded factorial or fractional factorial design with center points that is augmented with a group of 'star points' that allow estimation of curvature. If the distance from the center of the design space to a factorial point is ±1 unit for each factor, the distance from the center of the design space to a star point is $\pm\alpha$ with $|\alpha| > 1$. The precise value of α depends on certain properties desired for the design and on the number of factors involved.

Similarly, the number of center-point runs that the design should contain also depends on certain properties required for the design.

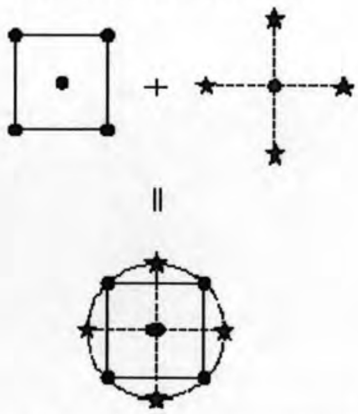

FIGURE 7.12 Generation of a central composite design for two factors

A central composite design always contains twice as many star points as there are factors in the design. The star points represent new extreme values (low and high) for each factor in the design. Table 7.3 summarizes the properties of the three varieties of central composite designs. Figure 7.13 illustrates the relationships among these varieties.

TABLE 7.3 Central Composite Designs[7]

CCD Type	Terminology	Comments
Circumscribed	CCC	CCC designs are the original form of the central composite design. The 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 star points establish new extremes for the low and high settings for all factors. Figure 5 illustrates a CCC design. These designs have circular, spherical, or hyperspherical symmetry and require 5 levels for each factor. Augmenting an existing factorial or resolution V fractional factorial design with star points can produce this design.
Inscribed	CCI	For those situations in which the limits specified for factor settings are truly limits, the CCI design uses the factor settings as the star points and creates a factorial or fractional factorial design within those limits (in other words, a CCI design is a scaled down CCC design with each factor level of the CCC design divided by α to generate the CCI design). This design also requires 5 levels of each factor.
Face Centered	CCF	In this design the star points are at the center of each face of the factorial space, so $\alpha = \pm 1$. This variety requires 3 levels of each factor. Augmenting an existing factorial or resolution V design with appropriate star points can also produce this design.

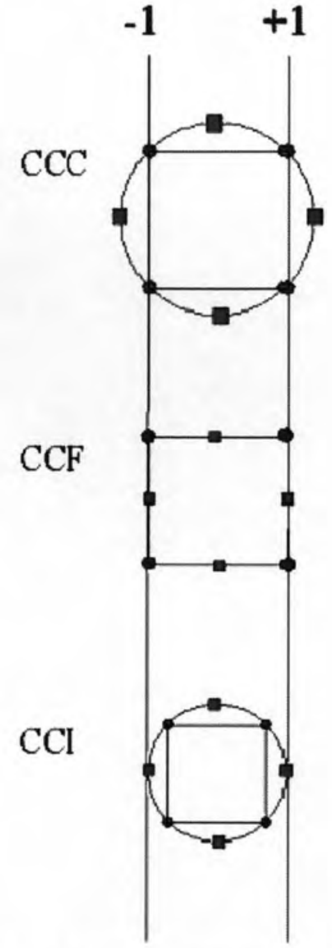

FIGURE 7.13 Comparison of the Three Types of Central Composite Designs

The diagrams in Figure 7.13 illustrate the three types of central composite designs for two factors. Note that the CCC explores the largest process space and the CCI explores the smallest process space. Both the CCC and CCI are rotatable designs, but the CCF is not. In the CCC design, the design points describe a circle circumscribed about the factorial square. For three factors, the CCC design points describe a sphere around the factorial cube.

7.4.3.1 Determining α in Central Composite Designs

To maintain rotatability, the value of α depends on the number of experimental runs in the factorial portion of the central composite design:

$$\alpha = [\text{number of factorial runs}]^{1/4} \qquad (7.3)$$

If the factorial is a full factorial, then

$$\alpha = [2^k]^{1/4} \tag{7.4}$$

However, the factorial portion can also be a fractional factorial design of resolution V (Variance). Table 7.4 illustrates some typical values of α as a function of the number of factors.

TABLE 7.4 Determining α for rotatability[7]

Number of Factors	Factorial Portion	Scaled Value for α Relative to ±1
2	2^2	$2^{2/4} = 1.414$
3	2^3	$2^{3/4} = 1.682$
4	2^4	$2^{4/4} = 2.000$
5	2^{5-1}	$2^{4/4} = 2.000$
5	2^5	$2^{5/4} = 2.378$
6	2^{6-1}	$2^{5/4} = 2.378$
6	2^6	$2^{6/4} = 2.828$

The value of α also depends on whether or not the design is orthogonally blocked. That is, the question is whether or not the design is divided into blocks such that the block effects do not affect the estimates of the coefficients in the 2nd order model. Under some circumstances, the value of α allows simultaneous rotatability and orthogonality.