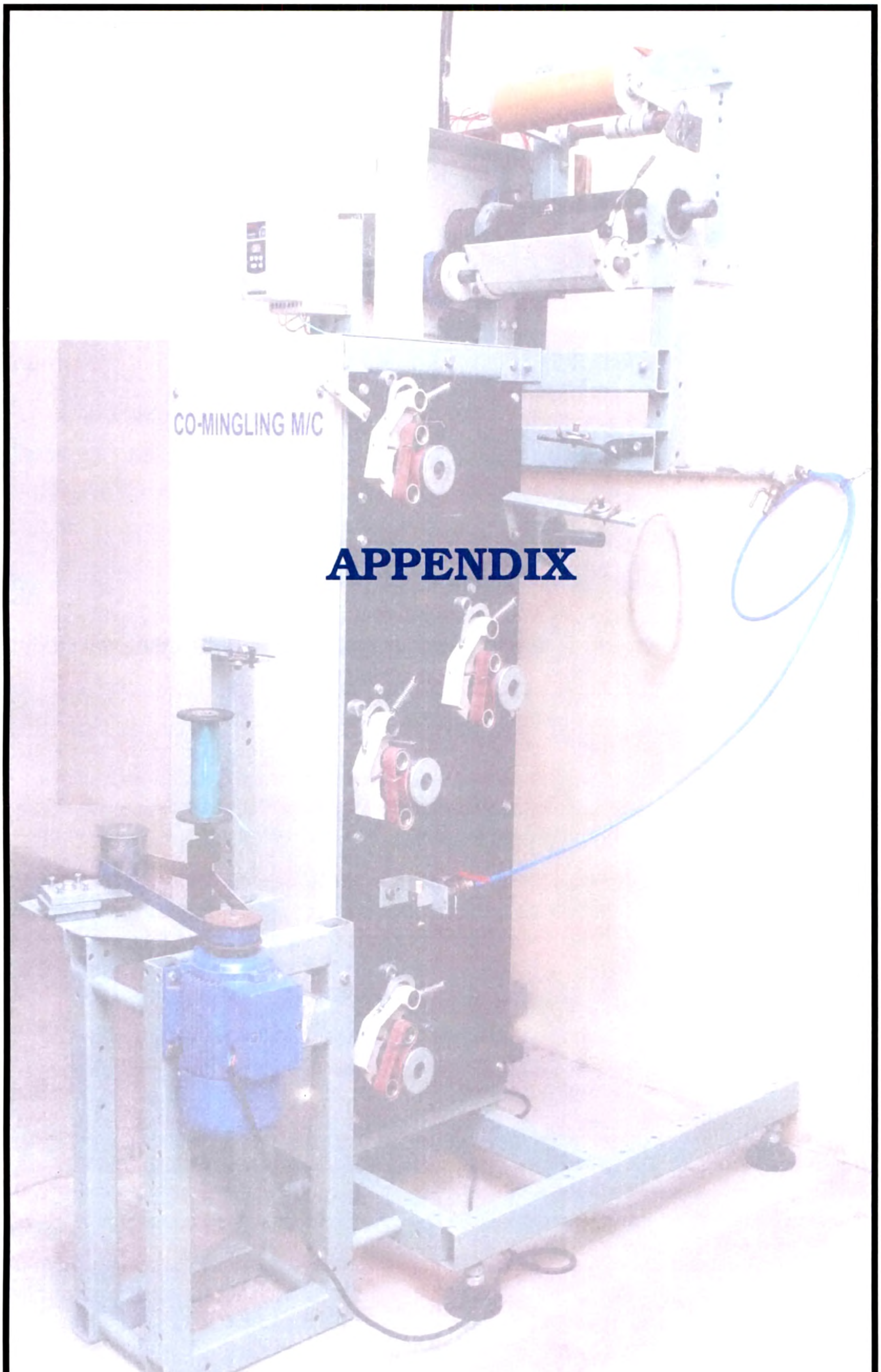
APPENDIX



APPENDIX

A.1 INTRODUTION

The classical approach to experimental planning (one-at-a-time designs) involves much effort and time and in some cases, it is inapplicable due to simultaneous interaction of so many factors. The most efficient way to enhance the value of research and to cut down the time in process development is through experimental designs. The statistical experiment designs widely used for optimization of experiments are termed "Response Surface Designs".

Response Surface Methodology (RSM) was first introduced by Box and Wilson in 1951 and later popularised by Montgomery[68].The response surface methodology can be defined as an empirical statistical technique employed for multiple regression analysis by using quantitative data obtained from properly designed experiments to solve multivariate equations simultaneously. The graphical representations of these equations are called response surfaces, which can be used to describe the individual and cumulative effect of the test variables on the response and to determine the mutual interactions between the test variables and their subsequent effect on the response.

Basically RSM is a combination of statistical experimental design fundamentals, regression modeling techniques, and optimisation methods. The RSM uses various Design of Experiments(DOE) techniques, such as Box-Behnkhen Design (BBD), Central Composite Design (CCD), Full and Fractional Factorial Design, as well as regression analysis methods. The DOE techniques are employed before, during and after the regression analysis to evaluate the accuracy of the model. The main idea of RSM is to replace a complicated response function with an approximate function by studying the relative significance of the effects of several factors supposed to have influence on the response of interest.

Hireni Mankodi
M S University of Baroda

- If the system response is rather well discovered, RSM techniques are used to find the best (optimum) value of the response.
- If discovering the best value is beyond the available resources of the experiment, then RSM techniques are used to at least gain a better response.
- Understanding of the overall response system and can be used to identify new operating conditions that produce demonstrated improvement in product quality.
- If obtaining the system response is necessary for complicated analysis that requires hours of run-time and advanced computational resources. The simplified equivalent response surface may be obtained by a few numbers of runs to replace the complicated analysis.

The RSM has some inherent advantages over other experimental designs, due to which this technique is highly acceptable among the researchers. It provides a way of rigorously choosing a few parameters in a design space to represent all the points efficiently thereby reducing the number of runs required to study the significance of different factors affecting the response of interest.

A.2 RESPONSE SURFACE METHODOLOGY : AN OVERVIEW

Assume that the true response, y, of a system depends on k number of controllable input variables (or factors) ξ_1, ξ_2, ξ_3... ξ_k. then their relationship can be represented as:

$$y = f\ (\xi_1, \xi_2, \xi_3, \ldots\ \xi_k) + \varepsilon \qquad [1]$$

The function **f** is called the true response function, form of which is unknown and usually complicated, and ε is a term representing sources of variability not accounted for in **f.** The term ε is treated as a statistical error. which ε includes effects such as measurement error on the response, background noise, the effect of other variables, and so on, and often assumed to have a normal distribution with mean and variance σ. The variables ξ_1, ξ_2, ... ξ_k in equation [1] are known as natural variables because they are expressed in the natural units of measurement, such as degrees celsius, pounds per square inch, etc. In most of the RSM experiment, the natural variables are transformed into dimensionless coded variables with mean zero and the standard deviation same as that of the natural variables. Usually coded variables are calculated using the equation:

$$X_i = \frac{X_{ij} - [(\max x_{ij} + \min x_{ij})/2}{[(\max x_{ij} - \min x_{ij})/2]} \quad [2]$$

where x_{ij}.. is the i^{th} natural variable for the j^{th} experimental run.

For two factors, (ie, **k** = 2), a second-order polynomial approximation of the true response function in terms of the coded variables will be written as:

$$\eta' = E(y) = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \beta_{jj} x_j^2 + \sum_{i<} \sum_{j=1}^{k} \beta_{ij} x_i x_j \quad [3]$$

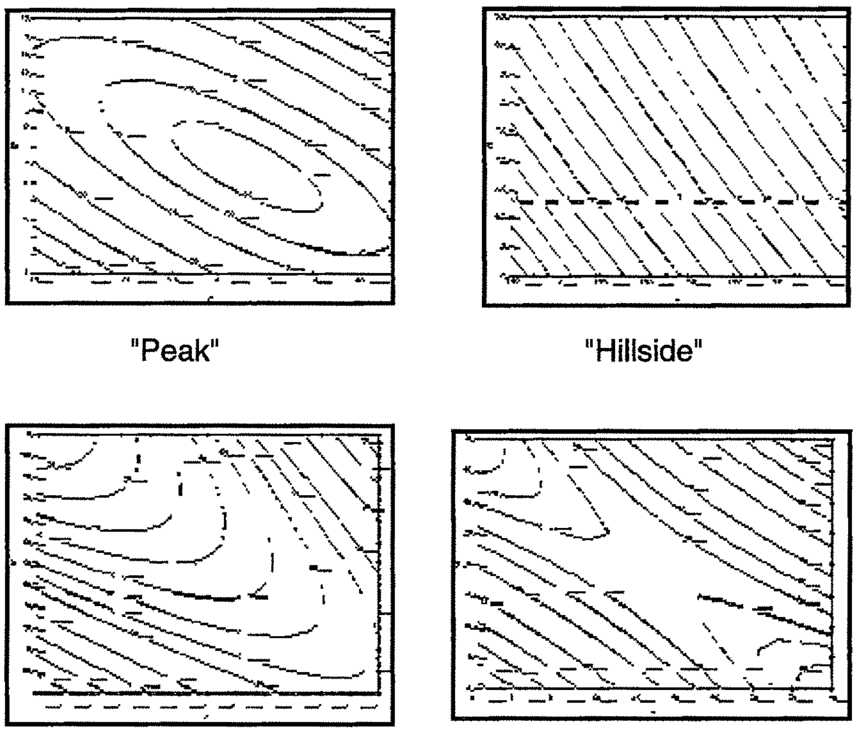
where, *x* are called 'coded variables' and β's are called regression coefficients.

The terms x_1, x_2 are main effects and the term x_1, x_2 are called interactions. Adding the interaction term introduces curvature into the response function. In most cases, the second-order model is adequate for well-behaved responses since it can take on a wide variety of functional forms, so it often works well as an approximation to the true response surface. Moreover the parameters in a second order model can be easily estimated using the least square method.

This empirical model is called a 'response surface model'. The surface represented by $f(x_i, x_j)$ is called a response surface. The response can be represented graphically, either in the three-dimensional space or as contour plots that help visualizes the shape of the response surface. Contours are curves of constant response drawn in the *xyz* plane keeping all other variable fixed. Each contour corresponds to a particular height of the response surface. A contour plot is formed by a series of horizontal and vertical lines. The horizontal axis plots are the most important factor in the experiment where as the vertical axis plots are the second most important factor in the experiment.

Phase 1: The first stage is a generic step for screening factors. The objective of factor screening is to reduce the list of candidate variables to a relatively few so that subsequent experiments will be more efficient and require fewer runs or tests. The screening is based on main effects estimation.

Phase 2: The object of phase 2 is to fit a second-order model for the factors identified from the screening experiment. A correct choice of design will ensure that the response surface is fit in the most efficient manner. The choice of a suitable design depends on the number of factors under consideration and the coverage of the chosen design over the region of interest on the response surface. The desirable features of a chosen design are orthogonality, i.e. main effects and block effects are estimable independently



"Peak" "Hillside"

Surface "Rising Ridge" Surface "Saddle"

Fig. A.1 Contour plots for different system response

Central Composite Design (CCD) and Box-Behnken Design (BBD). The CCD and rotatability, i.e. constant predictability at all points equidistant from the design centre. Based upon the desirable features, most preferred designs are is appropriate for evaluating linear or quadratic response surface models and is often recommended for sequential experimentation. The BBD on the other hand, can be used for performing non-sequential experiments. The CCD usually has axial points outside the design periphery.

Although these design points have significant contribution towards design accuracy, still they are not desired in many cases when these conditions are beyond the safe operating limits. The BBD design ensures that all factors are set within the experimental periphery, but has lower accuracy than CCD.

Phase 3: The objective of this phase is to identify the theoretical value of factor region that yields the optimal response. Some commonly used optimisation techniques are "Best corner", "Steepest ascent descent" and "Optimal plot" techniques. Among these techniques, "Control plot" technique provides the "best guess" as to where to run the experiment so as to obtain the desired optimal response.

Phase 4: Phase 4 begins when the process is near the optimum. The step-wise regression procedure is followed for adding or deleting model terms depending on probabilities (*p*values). The final model can be build up from the simplest models by adding and testing higher-order terms (the "forward" direction), or the final model can be reached starting with the full second-order model and eliminating terms until the most parsimonious, adequate model is obtained (the "backward" direction). Once an appropriate approximating model has been obtained, this model may be analysed to determine the optimum conditions for the process. The final model should have minimum residuals or error of prediction.

A.2.1 Comparisons of Response Surface Designs

The Box-Behnken design is an independent quadratic design which does not contain an embedded factorial or fractional factorial design. In this design the treatment combinations are at the midpoints of edges of the process space and at the center. These designs are rotatable (or near rotatable) and require

3 levels of each factor. The designs have limited capability for orthogonal blocking compared to the central composite designs. Fig A.2 illustrates a Box-Behnken design for three factors. Table A.1 gives comparison of the structures of four common quadratic designs that to investigate three factors. The table combines Central Composite Circumscribed (CCC) and central composite inscribed (CCI) designs because they are structurally identical. For three factors, the Box-Behnken design offers some advantage in requiring a fewer number of runs. For 4 or more factors, this advantage disappears.

Table A.2. illustrates the factor settings required for a central composite circumscribed (CCC) design and for a central composite inscribed (CCI) design (standard order), assuming three factors, each with low and high settings of 10 and 20, respectively. Because the CCC design generates new extremes for all factors, the investigator must inspect any worksheet generated for such a design to confirm that the factor settings called for are reasonable.

In Table A.2, treatments 1 to 8 in each case are the factorial points in the design; treatments 9 to 14 are the star points; and 15 to 20 are the system-recommended center points. If can been seen in the CCC design that how the

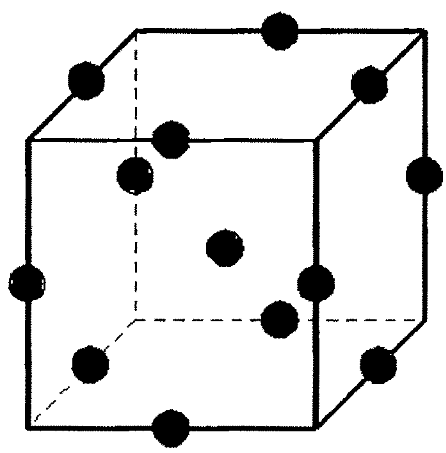


Fig. A.2 Box-Behnken design for three factors

Table A.1 Structural Comparisons of CCC (CCI), CCF, and BBD for three factors

CCC (CCI)				CCF				BBD			
Rep	*X*1	*X*2	*X*3	Rep	*X*1	*X*2	*X*3	Rep	*X*1	*X*2	*X*3
1	-1	-1	-1	1	-1	-1	-1	1	-1	-1	0
1	+1	-1	-1	1	+1	-1	-1	1	+1	-1	0
1	-1	+1	-1	1	-1	+1	-1	1	-1	+1	0
1	+1	+1	-1	1	+1	+1	-1	1	+1	+1	0
1	-1	-1	+1	1	-1	-1	+1	1	-1	0	-1
1	+1	-1	+1	1	+1	-1	+1	1	+1	0	-1
1	-1	+1	+1	1	-1	+1	+1	1	-1	0	+1
1	+1	+1	+1	1	+1	+1	+1	1	+1	0	+1
1	-1 682	0	0	1	-1	0	0	1	0	-1	-1
1	1.682	0	0	1	+1	0	0	1	0	+1	-1
1	0	-1.682	0	1	0	-1	0	1	0	-1	+1
1	0	1.682	0	1	0	+1	0	1	0	+1	+1
1	0	0	-1.682	1	0	0	-1	3	0	0	0
1	0	0	1.682	1	0	0	+1				
6	0	0	0	6	0	0	0				
Total Runs = 20				Total Runs = 20				Total Runs = 15			

low and high values of each factor have been extended to create the star points. In the CCI design, the specified low and high values become the star points, and the system computes appropriate settings for the factorial part of the design inside those boundaries.

Table A.3. illustrates the factor settings for the corresponding central composite face-centered (CCF) and Box-Behnken designs. Note that each of these designs provides three levels for each factor and that the Box-Behnken design requires fewer runs in the three-factor case. Table A.4. give comparisons of properties of various classical response surface designs.

Table A.2. Factor for CCC and CCI designs for three factors

Central Composite Circumscribed (CCC)					Central Composite Inscribed (CCI)			
Sr. No.	*X*1	*X*2	*X*3		Sr. No.	*X*1	*X*2	*X*3
1	10	10	10		1	12	12	12
2	20	10	10		2	18	12	12
3	10	20	10		3	12	18	12
4	20	20	10		4	18	18	12
5	10	10	20		5	12	12	18
6	20	10	20		6	18	12	18
7	10	20	20		7	12	12	18
8	20	20	20		8	18	18	18
9	6.6	15	15	*	9	10	15	15
10	23.4	15	15	*	10	20	15	15
11	15	6.6	15	*	11	15	10	15
12	15	23.4	15	*	12	15	20	15
13	15	15	6.6	*	13	15	15	10
14	15	15	23.4	*	14	15	15	20
15	15	15	15		15	15	15	15
16	15	15	15		16	15	15	15
17	15	15	15		17	15	15	15
18	15	15	15		18	15	15	15
19	15	15	15		19	15	15	15
20	15	15	15		20	15	15	15

* star points

Table A.3.Factors for CCF and BBD for three factors

Central Composite Face-centered (CCF)					Box-Behnken Design(BBD)			
Sr. No.	*X*1	*X*2	*X*3		Sr. No.	*X*1	*X*2	*X*3
1	10	10	10		1	10	10	10
2	20	10	10		2	20	10	15
3	10	20	10		3	10	20	15
4	20	20	10		4	20	20	15
5	10	10	20		5	10	15	10
6	20	10	20		6	20	15	10
7	10	20	20		7	10	15	20
8	20	20	20		8	20	15	20
9	10	15	15	*	9	15	10	10
10	20	15	15	*	10	15	20	10
11	15	10	15	*	11	15	10	20
12	15	20	15	*	12	15	20	20
13	15	15	10	*	13	15	15	15
14	15	15	20	*	14	15	15	15
15	15	15	15		15	15	15	15
16	15	15	15					
17	15	15	15					
18	15	15	15					
19	15	15	15					
20	15	15	15					

* are star points for the CCC design

Table A.4. Comparison of Properties of Classical Response Surface Designs

Design type	Comment
CCC	CCC designs provide high quality predictions over the entire design space, but require factor settings outside the range of the factors in the factorial part. Note: When the possibility of running a CCC design is recognized before starting a factorial experiment, factor spacings can be reduced to ensure that ± alfa for each coded factor corresponds to feasible (reasonable) levels. It requires 5 levels for each factor.
CCI	CCI designs use only points within the factor ranges originally specified, but do not provide the same high quality prediction over the entire space compared to the CCC. It requires 5 levels of each factor.
CCF	CCF designs provide relatively high quality predictions over the entire design space and do not require using points outside the original factor range. However, they give poor precision for estimating pure quadratic coefficients. It requires 3 levels for each factor
BBD	Box-Behnken design requires fewer treatment combinations than a central composite design in cases involving 3 or 4 factors.This design is rotatable (or nearly so) but it contains regions of poor prediction quality like the CCI design. Its "missing corners" may be useful when the experimenter should avoid combined factor extremes. This property prevents a potential loss of data in those cases. It requires 3 levels for each factor.

A.2.2 Desirable Features for Response Surface Designs

G. E. P. Box and N. R. Draper in "Empirical Model Building and Response Surfaces," John Wiley and Sons, New York, identify desirable properties for a response surface design:

- Satisfactory distribution of information across the experimental region *rotatability*
- Fitted values are as close as possible to observed values. Minimize residuals or error of prediction
- Good lack of fit detection.

- Internal estimate of error.
- Constant variance check.
- Transformations can be estimated.
- Suitability for blocking.
- Sequential construction of higher order designs from simpler designs
- Minimum number of treatment combinations.
- Good graphical analysis through simple data patterns.
- Good behaviour when errors in settings of input variables occur.

A.2.3 Applications and limitations of RSM

The application of RSM is aimed at reducing the cost of expensive analysis methods and their associated huge investment of resources and volumes of numerical data analysis. This particular advantage has paved the way for its successful application in different disciplines such as chemical and pharmaceutical processes, biological/biochemical processes, food science, production engineering, air quality analysis and toxicological research and computational and simulation studies.

Researchers from different fields in recent years have published several interesting applications of response surface methodology. In one such study, C J Stevens of NASA, United States, integrated RSM with computational fluid dynamics (CFD) for prediction of combined cycle propulsion components in hypersonic jet fighter. Neda and co-workers of Memorial University of Newfoundland, Canada, combined Monte Carlo simulation method with the RSM to compute permanent displacement of submarine slope under earth-quake loads. The results obtained from the experimental study were reported to be almost identical to that obtained from replicating the actual model. In another research, Matthew et al explored nonparametric version of RSM to estimate the location of the maximum AIDS incidence in California among two ethnic groups. The study was reported to be the true reflection of the underlying situation in comparison to its parametric counter part.

In spite of such inspiring experimental outcome, RSM suffer from some serious drawbacks. One is its sequential approach. This sequential approach of RSM can be considered as a disadvantage when the experimental preparation is time-consuming or its duration is long. Cheng and his co-workers suggested to integrate factor screening and response surface development on the same experiment and proposed a new approach, which can serve as a link between the two.

Another limitation of RSM is sensitivity to system noise. It is assumed in RSM that the experimental noise factors controllable during process development for purposes of a designed experiment This assumption reduces the robust-ness of the RSM models. Professor Taguchi modified RSM and developed a new approach known as robust parameter design methodology (RPD) that will make RSM models insensitive (or robust) to changes in a set of uncontrollable factors.