

**RESPONSE SURFACE METHODOLOGY USING  
AN OPTIMIZATION TECHNIQUE**

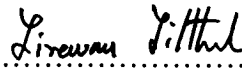
**Chantha Wongoutong**

**A Dissertation Submitted in Partial  
Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy (Statistics)  
School of Applied Statistics  
National Institute of Development Administration  
2016**

**RESPONSE SURFACE METHODOLOGY USING  
AN OPTIMIZATION TECHNIQUE**


**Chantha Wongoutong**  
**School of Applied Statistics**


---

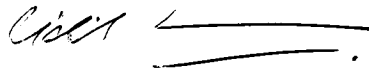
Associate Professor..........Major Advisor  
(Jirawan Jithavech, Ph.D.)


Associate Professor..........Co-Advisor  
(Vichit Lorchirachoonkul, Ph.D.)


The Examining Committee Approved This Dissertation Submitted in Partial  
Fulfillment of the Requirements for the Degree of Doctor of Philosophy (Statistics).

Professor..........Committee Chairperson  
(Prachoom Suwattee, Ph.D.)

Associate Professor..........Committee  
(Jirawan Jithavech, Ph.D.)

Associate Professor..........Committee  
(Vichit Lorchirachoonkul, Ph.D.)

Professor..........Committee  
(Samruam Chongcharoen, Ph.D.)

Assistant Professor..........Dean  
(Sutep Tongngam, Ph.D.)

March 2017

## ABSTRACT

<b>Title of Dissertation</b>	Response Surface Methodology Using an Optimization Technique
<b>Author</b>	Miss Chantha Wongoutong
<b>Degree</b>	Doctor of Philosophy (Statistics)
<b>Year</b>	2016

---

Response surface methodology (RSM) is techniques combine both of experimental designs and statistical techniques for empirical model building and optimization. The experimental design is considered by the objective is to optimize one or more response variables influenced by several independent variables. However, in real situation, we may not be able to identify the true model and so an approximated model, usually a central composite design for building a second-order polynomial model, this design is popular in RSM.

A novel method using the Nelder-Mead algorithm is proposed to be used instead of the first order model in moving the experiment in the response surface methodology toward the neighbor of the optimum. A second order model similar to the second order model in the CCD is constructed to estimate the optimum design point and the optimum response. From the simulation using five published test functions and five different normal generators, it can be concluded that the proposed method outperforms the traditional CCD in terms of the number of experiments, the MAPEs of the estimated.

## **ACKNOWLEDGEMENTS**

I would like to express my deepest thanks to everyone for their help in completing this dissertation. In particular, I wish to thank my advisor, Associate Professor Dr. Jirawan Jitthavech, for her invaluable and continual advice, encouragement, and constructive criticism throughout this research, which enabled me to complete this dissertation successfully. I also wish to extend thanks and appreciation to the other committee members: Associate Professor Dr. Vichit Lorchirachoonkul and Professor Dr. Samruam Chongcharoen for contributing both their time and helpful comments and suggestions. They have all provided me with whatever support I have asked from them. Finally, I would like to express my gratitude to my family and friends for their help, patience, great love, encouragement, and support throughout my doctorate study.

Chantha Wongoutong

June 2016

## **TABLE OF CONTENTS**

	<b>Page</b>
<b>ABSTRACT</b>	<b>iii</b>
<b>ACKNOWLEDGEMENTS</b>	<b>iv</b>
<b>TABLE OF CONTENTS</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>xi</b>
<b>LIST OF FIGURES</b>	<b>ix</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
1.1 Background	1
1.2 Objectives of the Study	4
1.3 Scope of the Study	4
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>5</b>
2.1 Response Surface Methodology (RSM)	5
2.2 Screening Design	10
2.3 The Method of Steepest Ascent	13
2.4 Second-order Models	18
2.5 Experimental Design Properties	21
2.6 Canonical Analysis	24
2.7 The NM Method	26
2.8 Coverage Probability	30
2.9 Mean Absolute Percentage Error (MAPE)	31
<b>CHAPTER 3 THE PROPOSED METHOD TO FIL A SECOND-ORDER MODEL</b>	<b>32</b>
3.1 The Seeding Step for the Proposed Method	32
3.2 The Second Step in the Proposed Method	33

<b>CHAPTER 4 SIMULATION STUDY</b>	<b>35</b>
4.1 Steps in the Simulation Study	37
4.2 Results of the Simulation Study	41
<b>CHAPTER 5 CONCLUSIONS AND FUTURE WORK</b>	<b>74</b>
5.1 Conclusions	74
5.2 Future Work	75
<b>BIBLIOGRAPHY</b>	<b>76</b>
<b>BIOGRAPHY</b>	<b>80</b>

## LIST OF TABLES

Tables	Page
2.1 Relationship between the Response and Independent Variables	8
2.2 Number of Runs for $2^k$ Full Factorial Designs	11
2.3 $2^4$ Full Factorial Design	12
2.4 $2^{4+1}$ Fractional Factorial Design	13
2.5 CCD Types	20
2.6 Structural Comparison of CCC (CCI), CCF, and BBD for Three Factors	21
4.1 The Minima and the Minimum Response of the Test Functions	37
4.2 The Average Number of Required Experiments for RSM (CCD) and the Proposed Method	41
4.3 Comparison of the Coverage Probability of RSM(CCD) to the Proposed Method (NM(1),NM(2),NM(3),NM(4)) for $f_1$	45
4.4 Comparison of the Coverage Probability of the Classical RSM(CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function $f_2$	49
4.5 Comparison of the Coverage Probability of the Classical RSM(CCD) with the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function $f_3$	53
4.6 Comparison of the Coverage Probability of the Classical RSM(CCD) and the Proposed Method forms NM(1), NM(2), NM(3), and NM(4) for Function $f_4$	57
4.7 Comparison of the Coverage Probabilities of the Classical RSM(CCD) and the Proposed Method Forms NM(1), NM(2), NM(3),and NM(4) for Function $f_5$	61

4.8 Comparison of the MAPE of the classical RSM(CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function $f_1$	63
4.9 Comparison of the MAPE of the Classical RSM (CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for function $f_2$	65
4.10 Comparison of the MAPE of the Classical RSM (CCD) and the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function $f_3$	67
4.11 Comparison of the MAPE of the classical RSM(CCD) and the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function $f_4$	69
4.12 Comparison of the MAPE of the Classical RSM(CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function $f_5$	71
4.13 The True Value, Mean and Standard Deviation of the Estimated Values of the Minimum Design Points and the Minimum Responses in the Simulation by the Classical RSM(CCD) and the Proposed Method	72



## LIST OF FIGURES

Figures	Page
2.1 $2^3$ Full Factorial Design	11
2.2 Path of Steepest Ascent Improvement in the Region of the Optimum	14
2.3 Flow Chart showing the Construction of the Response Surface	14
2.4 First-order Response Surface and Path of Steepest Ascent	17
2.5 Generation of a CCD for Two Factors	19
2.6 The Three CCD Types with Three Factors	21
2.7 Types of Response Surface	26
2.8 The Sequence of Triangles $\{ T_k \}$ converging to the Point (3,2) for Minimization of a Function of <i>Two</i> Variables using the NM Method	27
2.9 Possible Operations Performed on a Simplex in $\Re^2$	30
3.1 The $2^2$ Factorial Design as a Starting Point	32
3.2 Four Possible Starting Points in the First Phase of the Proposed Method	33
3.3 Flow Chart of the NM Algorithm	34
4.1 The Response Surface and Contour of $f_1$	35
4.2 The Response Surface and Contour of $f_2$	36
4.3 The Response Surface and Contour of $f_3$	36
4.4 The Response Surface and Contour of $f_4$	36
4.5 The Response Surface and Contour of $f_5$	37
4.6 $2^2$ Factorial Design	39
4.7 Four Possible Starting Points in the First Phase of the Proposed Method	40
4.8 Comparison of Average Points for CCD and the Proposed Method for Functions $f_1$ to $f_5$	41

- 4.9 One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_1$  Using (CCD) and Propose 42  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.10 One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_1$  Using (CCD) and Propose 43  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.11 One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_1$  Using (CCD) and Propose 44  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.12 One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_2$  Using (CCD) and Propose 46  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.13 One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_2$  Using (CCD) and Propose 47  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.14 One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_2$  Using (CCD) and Propose 48  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.15 One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_3$  Using (CCD) and Propose 50  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.16 One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_3$  Using (CCD) and Propose 51  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.17 One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_3$  Using (CCD) and Propose 52  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)
- 4.18 One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_4$  Using (CCD) and Propose 54  
Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)

4.19	One Hundred CIs of Average $\hat{x}_{\min_2}$ of $f_4$ Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)	55
4.20	One Hundred CIs of Average $\hat{y}_{\min}$ of $f_4$ Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)	56
4.21	One Hundred CIs of Average $\hat{x}_{\min_1}$ of $f_5$ Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)	58
4.22	One Hundred CIs of Average $\hat{x}_{\min_2}$ of $f_5$ Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)	59
4.23	One Hundred CIs of Average $\hat{x}_{\min_2}$ of $f_5$ Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)	60
4.24	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_1$ of $f_1$	62
4.25	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_2$ of $f_1$	62
4.26	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $y$ of $f_1$	62
4.27	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_1$ of $f_2$	64
4.28	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_2$ of $f_2$	64
4.29	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $y$ of $f_2$	64
4.30	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_1$ of $f_3$	66

4.31	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_2$ of $f_3$	66
4.32	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $y$ of $f_3$	66
4.33	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_1$ of $f_4$	68
4.34	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_2$ of $f_4$	68
4.35	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $y$ of $f_4$	68
4.36	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $x_1$ and $f_5$	70
4.37	Comparison of the MAPE of RSM (CCD) to the proposed method for the True $x_2$ and $f_5$	70
4.38	Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True $y$ and $f_5$	70

# **CHAPTER 1**

## **INTRODUCTION**

### **1.1 Background**

It is common knowledge that the design of experiments (DOEs) is a statistical technique used to find the optimum performance of a system when the independent variables (factors) are known. The first step using this technique is to screen the experimental design used to identify the important factors suspected of affecting the system's performance (response). When the number of independent variables (factors) is large, there is a need to reduce this number by screening them using the experimental design. After that, another designed experiment test plan needs to be carried out with the objective of optimizing the system's performance. In general, the initial and final optimized DOEs are called the screening design and the response surface method (RSM), respectively. Response surface methodology, developed by Box and Wilson (1951), is usually employed to achieve optimum efficiency in manufacturing by establishing the best settings for the control process parameters (factors) to achieve the optimum process response. For a recent review of response surface methodology, see Myers et al. (2004), and there are also many books devoted exclusively to various aspects of the process, such as Khuri and Cornell (1996), Box and Draper (2007), and Myers et al. (2009).

Myers and Montgomery (2002) introduced three phases for conducting an RSM: phase 0, the screening of independent variables (input factors) is considered to identify those which have a significant effect on the response(s); phase 1) the location of the optimum operating conditions is determined by conducting a sequence of suitable experiments; and phase 2) the fitting of an appropriate empirical model, usually a second-order polynomial model, is carried out to examine the nature of the response surface in the vicinity of the optimum. There are many types of second-order design with the central composite design (CCD) introduced by Box and Wilson (1951) being the most popular.

Most of the problems in response surface methodology occur when searching for the relationship between the response and independent variables, which is usually solved by using a low order polynomial in some region of the independent variables. However, for some regions that should be in the optimum region (refer to phase 1 above), the steepest ascent/descent search approach is used to move them to the optimum region.

The traditional RSM is based on initially conducting steepest ascent or descent searches until a significant curvature is detected. Box and Wilson (1951) used this method to maximize the response based on experiments conducted on the direction defined by the gradient of an estimated main effects model. The observed responses along the steepest ascent direction are used to locate the neighborhood of the maximum. This method can theoretically locate the maximum through numerous iterations as long as it exists. However, if it is used on a badly scaled system, the rate of convergence may become too slow and the use of the method is impractical. Normally, the step size is estimated by using the coefficient of regression in the first order model based on the results from the experiments. In other words, the effectiveness of the traditional RSM depends on the step size given by the first order model.

An alternative to solving the step size problem in the traditional RSM is to only obtain the values experimentally without using the steepest direction when searching for the optimum of the response surface. The best-known methods in the direct search class include genetic algorithms (Holland, 1992), simulated annealing (Brooks and Morgan, 1995), Tabu search (Glover and Laguna, 1997), neural networks (Sexton et al., 1998) and the Nelder-Mead (NM) algorithm (Nelder and Mead, 1965). The NM algorithm is quite simple to understand and very easy to use (Gavin, 2013). These studies have led to its widespread application in many fields of science and technology, especially in chemistry and medicine.

The NM method is a derivative-free process for searching for the local optimum of a function. In this optimization process, the initial simplex adapts itself iteratively to the local surface landscape by varying its size and orientation. The NM algorithm is especially suitable for exploring “unwieldy” terrains, and has been widely accepted as the most robust and efficient of the current sequential techniques

for unconstrained optimization (Lagarias et al., 1998). Numerous software packages include the NM algorithm as an optimization solver, such as Mathematica (2002), MultiSimplex (2001), PROC IML in SAS (1998), among others. The idea behind the NM algorithm is to “crawl” to the optimum in the selected direction by moving one vertex of the simplex during each iteration. The vertices are moved by performing four basic operations: reflection, expansion, contraction, and multiple contractions (shrink).

Aimed at having better convergence, several variants of the simplex method have been proposed (Torczon, 1989; Dennis and Torczon, 1991; Byatt, 2000; Kelley, 2000; Tseng, 2000; Price et al., 2002). The NM algorithm generally performs well for solving small dimensional real-life problems and continuously remains as one of the most popular direct search methods (Wright, 1996; Lagarias et al., 1998; Kolda et al., 2003). However, many researchers have observed that the NM algorithm may become inefficient for large dimensional problems (Parkinson and Hutchinson, 1972; Torczon, 1989; Byatt, 2000). Nevertheless, for the majority of problems arising from response surface methodology practice, the number of influential process factors included in the final model is rarely larger than six (Olsson and Nelson, 1975; Myers and Montgomery, 2002). Typically, a “pre-experiment” via a fractional factorial is carried out in an earlier phase to eliminate irrelevant factors, leaving only a small number of relevant ones.

The NM algorithm has demonstrated its wide versatility, accuracy, and ease of use for solving various types of optimization problem in noise-free environments in the area of applied statistics (Olsson, 1974; Olsson and Nelson, 1975; Khuri and Cornell, 1987; Copeland and Nelson, 1996). However, the application of the NM algorithm in response surface optimization in the presence of errors has seldom been reported in the response surface methodology literature. In this dissertation, the problem of moving non-stationary to the optimum region is addressed. The NM algorithm is used for searching for the optimum region instead of the steepest direction in traditional response surface methodology.

## **1.2 Objectives of the Study**

- 1) To develop an experiment capable of constructing a second-order model based on the NM algorithm.
- 2) To compare the number of experimental points between the classical RSM (CCD) and the proposed method.
- 3) To compare the coverage probability and mean absolute percentage error (MAPE) between the classical RSM (CCD) method and the proposed method.

## **1.3 Scope of the Study**

In this study, the proposed method is developed from the NM algorithm. For finding the relationship between the response and independent variables, a second-order model is used. The efficiency of the proposed method is considered by comparison with the classical RSM (CCD) in terms of the number of experimental points, MAPE, and coverage probability.



## **CHAPTER 2**

### **LITERATURE REVIEW**

In this chapter, a review of the literature on the following topics is presented. Sections 2.1-2.6 contain details of response surface methodology, screening factors, methods of steepest ascent/descent, second-order models, canonical analysis, and the NM algorithm.

#### **2.1 Response Surface Methodology (RSM)**

Response surface methodology is a method consisting of an experimental design and statistical techniques for empirical model building and optimization. The theoretical approach to response surface methodology consists of three phases, as suggested by Mayers and Montgomery (2002): the first phase is the experimental strategy for exploring the space of the process or independent variables (factors), the second phase is the use of an empirical statistical model useful in developing an appropriate approximate relationship between the response and the independent variables, and the last phase consists of optimization methods useful for finding the values of the independent variables that produce the desired response variable values. Box and Wilson (1951) introduced the use of the first-order model, and, for the second-order models, many scientists and engineers have a working knowledge of central composite designs (CCDs) and the three-level designs of Box and Behnken (1960).

An appropriate approximating model is developed by statistical modeling of the response  $y$  and the independent variables (factors)  $\xi_1, \xi_2, \dots, \xi_k$ . In general, the relationship between the response  $y$  and the independent variables  $\xi_1, \xi_2, \dots, \xi_k$  can be expressed as

$$y = f(\xi_1, \xi_2, \dots, \xi_k) + \varepsilon. \quad (2.1)$$

In a real-life situation, the form of the true response function  $f$  is unknown and may be very complicated. The error term ( $\varepsilon$ ) represents other sources of variability not measured in  $f$ , usually random error assumed to have a normal distribution with zero mean and variance  $\sigma^2$ .

$$E(y) = \eta = E[f(\xi_1, \xi_2, \dots, \xi_k)] + E(\varepsilon) = f(\xi_1, \xi_2, \dots, \xi_k), \quad (2.2)$$

where variables  $\xi_1, \xi_2, \dots, \xi_k$  are usually called natural variables because they are expressed in natural units of measurement. The natural variables can be transformed to coded variables  $x_1, x_2, \dots, x_k$ , which are usually defined to be dimensionless with zero mean and the same standard deviation for covariance.

Russell (2009) mentioned that an appropriate coding transformation of the data is an important aspect in the response surface analysis. In this approach, the coded data affect the results of canonical and steepest ascent analysis. Using a coding method that makes all coded variables in the experiment vary over the same range is a way of giving each predictor an equal share in potentially determining the steepest ascent path. Thus, coding is an important step in response surface analysis.

Coded variables are usually calculated using the following equation:

$$x_{ij} = \frac{\xi_{ij} - [(\max \xi_{ij} + \min \xi_{ij}) / 2]}{[(\max \xi_{ij} - \min \xi_{ij}) / 2]}, \quad (2.3)$$

where  $\xi_{ij}$  is the  $i^{\text{th}}$  natural variable for the  $j^{\text{th}}$  experimental run,  $i = 1, 2, \dots, k$  and  $j = 1, 2, \dots, n$ .

In terms of the coded variables, the response in equation (2.2) can be written as

$$\eta = f(x_1, x_2, \dots, x_k). \quad (2.4)$$

In general, such a relationship is unknown but can be approximated by a low-order polynomial model of the form

$$\mathbf{y} = \mathbf{f}^T(x_1, x_2, \dots, x_k) \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2.5)$$

where  $\mathbf{f}(x_1, x_2, \dots, x_k)$  is a vector function of  $p$  elements consisting of powers and cross products of powers  $x_1, x_2, \dots, x_k$  up to a certain degree denoted by  $d(\geq 1)$ ,  $\boldsymbol{\beta}$  is a vector of  $p$  unknown constant coefficients referred to as parameters, and  $\boldsymbol{\varepsilon}$  is a random experimental error assumed to have a zero mean. This is conditioned on the understanding that equation (2.5) provides an adequate representation of the response. In this case, the quantity  $\mathbf{f}^T(x_1, x_2, \dots, x_k)$  represents the mean response, that is, the expected value of  $y$ , and is denoted by  $\mu(\mathbf{x})$ .

Two important models are commonly used in response surface methodology. These are special cases of model (2.5), when  $(d = 1)$  the first-degree model is obtained as

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \varepsilon. \quad (2.6)$$

If curvature is detected, a higher order polynomial model is required, such as a quadratic model  $(d = 2)$ ;

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \varepsilon, \quad (2.7)$$

where  $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$  and  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ .

Typically, the parameters estimate is carried out using the least squares method. The layout between response and independent variables are shown in Table 2.1 below.

**Table 2.1** Relationship between the Response and Independent Variables

$\mathbf{y}$	$\mathbf{X}_1$	$\mathbf{X}_2$	....	$\mathbf{X}_k$
$y_1$	$x_{11}$	$x_{12}$	....	$x_{1k}$
$y_1$	$x_{21}$	$x_{22}$	....	$x_{2k}$
$\vdots$	$\vdots$	$\vdots$	....	$\vdots$
$y_n$	$x_{n1}$	$x_{n2}$	....	$x_{nk}$

Equation (2.6) can be rewritten in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

$$\text{where } \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \text{ and } \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Let  $L = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$  be the least squares estimator  $L$  to be minimized, then  $L$  may be expressed as

$$\begin{aligned} L &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \end{aligned}$$

$L$  is minimized by taking derivatives with respect to the model parameters and setting them to zero, thus

$$\begin{aligned} \left. \frac{\partial L}{\partial \boldsymbol{\beta}} \right|_{\hat{\boldsymbol{\beta}}} &= -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = 0 \\ \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} &= \mathbf{X}'\mathbf{y} \\ \hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \end{aligned}$$

The fitted regression model can be presented as  $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ .

Properties of LSE  $\hat{\boldsymbol{\beta}}$ :

1.  $E(\hat{\beta}) = E[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y})]$   
 $= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{y})$   
 $= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta)$   
 $= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\beta$   
 $= \beta$
2.  $V(\hat{\beta}) = V[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{y})]$   
 $= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'V(\mathbf{y})[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']'$   
 $= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']'$   
 $= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']'$   
 $= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}$   
 $= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$
3.  $\hat{\beta} \sim N(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$ .

### 2.1.1 The Important Steps in Response Surface Methodology

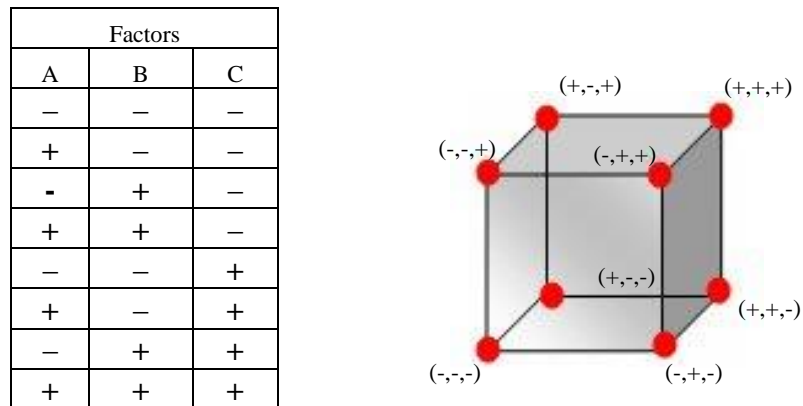
- 1) A screening factor is used when there are a lot of possible independent variables (factors), the main purpose being to eliminate the less important factors. After that, the experiment should be conducted with the independent variables varied around the present operating point.
- 2) Regression methods are usually used in this step for fitting the equation with experimental data. A linear (first-order) model usually represents the model sufficiently well.
- 3) If curvature is not found, it is necessary to move the experimental point in the direction of steepest ascent (or descent for a minimum).
- 4) If curvature is found, this indicates that the experiments are near the optimum.
- 5) Conduct a five-level factorial experiment around this point.
- 6) Regression analysis is used to obtain a fitted quadratic model.
- 7) Based on this quadratic model, determine the optimum by canonical analysis.

## 2.2 Screening Design

In many process development and manufacturing applications, there are a large number of independent variables (factors). Screening is used to reduce the number of independent variables by identifying the most important ones or process conditions that affect product quality. This reduction allows the focus to be on process improvement efforts on the few really important variables, or the "vital few". Screening may also suggest the optimal settings for these factors, and indicate whether or not curvature exists in the responses. The best settings of these factors can be determined by optimization experiments and defining the nature of the curvature. In industry, many designs are used to screen for the really important factors such as two-level full and fractional factorial designs, and Plackett-Burman (1946) designs; these are useful for fitting first-order models (which detect linear effects), and can be extended for second-order effects (curvature) when the design includes center points.

### 2.2.1 Full Factorial Designs

A full factorial design is a screening process which considers every possible combination of treatment levels for the different factors. In general, when using a full factorial design, responses for all combinations of design variable levels are evaluated. Therefore, all possible effects and interactions are included in the process, which means that for cases with large numbers of design variables and levels, the total number of runs becomes large. Hence, it is desirable to reduce the size of the runs, after which full factorial designs on two levels become appropriate. This design includes all input factors on two levels: 'high' and 'low', and can be denoted by '+1' and '-1', respectively. Generally, when  $k$  factors are considered, each at two levels, a full factorial design has  $2^k$  runs. Figure 2.1 shows the  $2^3$  full factorial design, where  $k = 3$ .



**Figure 2.1**  $2^3$  Full Factorial Design

**Table 2.2** Number of Runs for  $2^k$  Full Factorial Designs

Number of Factors (k)	Number of Runs ( $2^k$ )
2	$2^2=4$
3	$2^3=8$
4	$2^4=16$
5	$2^5=32$
6	$2^6=64$
7	$2^7=128$

From Table 2.2, when the number of factors is greater than five, a  $2^k$  full factorial design requires a large number of runs, and so is not very efficient. In these cases, other approaches such as a fractional factorial design or a Plackett-Burman design is a better choice.

### 2.2.2 Fractional Factorial Designs

Resources for the process are often insufficient when constructing a full factorial design, and so, for this reason, a fractional factorial design is a reasonable choice. Fractional factorial designs, especially two-level fractional factorial designs

$(2^{k-p})$ , are the most commonly used experimental plan. This is a very efficient screening design provided that the effects of interest can be estimated. Box and Hunter (1978) described a useful fractional factorial design for reducing the number of runs to be executed in an experiment by choosing a subset (fraction) of experimental runs of a full factorial design. This is a special category of two-level designs where not all factor level combinations are considered, and the experimenter can choose which combinations are to be excluded. Table 2.3 shows four factors in an experiment (A, B, C, and D) at each of two levels.

**Table 2.3**  $2^4$  Full Factorial Design

Factors			
A	B	C	D
–	–	–	–
+	–	–	–
–	+	–	–
+	+	–	–
–	–	+	–
+	–	+	–
–	+	+	–
+	+	+	–
–	–	–	+
+	–	–	+
–	+	–	+
+	+	–	+
–	–	+	+
+	–	+	+
–	+	+	+
+	+	+	+

A  $2^4$  full factorial design is design with four factors that consist of all 16 level combinations, but the available resources are only sufficient to conduct eight



experimental runs. Therefore, the requirement is to choose half of them. The chosen half is called a  $2^{4-1}$  fractional factorial design, as shown in Table 2.4.

**Table 2.4**  $2^{4-1}$  Fractional Factorial Design

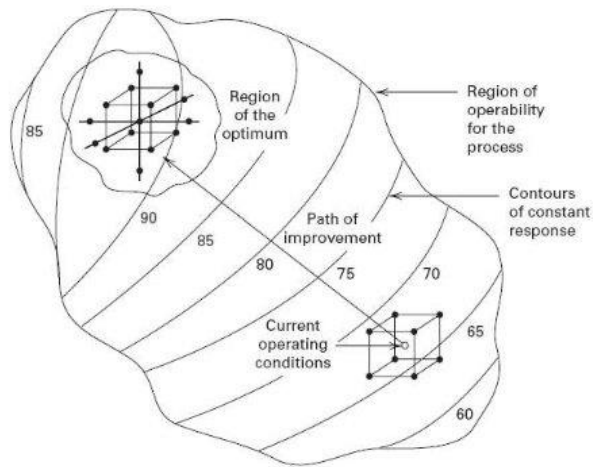
<b>I</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>AB</b>	<b>AC</b>	<b>BC</b>	<b>ABC=D</b>
1	-1	-1	-1	1	1	1	-1
1	1	-1	-1	-1	-1	1	1
1	-1	1	-1	-1	1	-1	1
1	1	1	-1	1	-1	-1	-1
1	-1	-1	1	1	-1	-1	1
1	1	-1	1	-1	1	-1	-1
1	-1	1	1	-1	-1	1	-1
1	1	1	1	1	1	1	1

When the number of factors  $k = 4$ , the fraction index  $p = 1$  can be used to construct a  $2^{4-1}$  design, and so the number of runs (level combinations)  $N = 2^4/2^1 = 8$ .

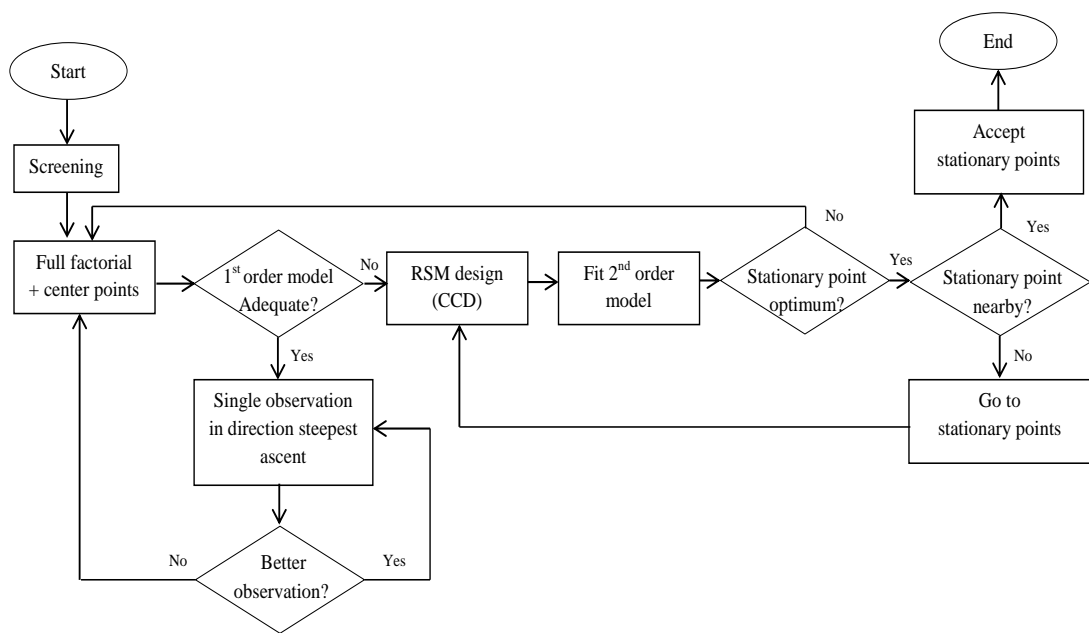
In this case, three factors (A, B, and C) are used to form a  $2^3$  full factorial design where factor D is confounded by high order interactions between A, B, and C.

### 2.3 The Method of Steepest Ascent

Traditionally, RSM are initially used to conduct the steepest ascending or descending searches until significant curvature is detected (Box and Wilson, 1951). Steepest ascent is the method which considers the direction defined by the gradient of an estimated main effects model for used in maximizing the response based on the experiments being conducted by means of the observed responses along the steepest ascent to estimate when the experiment reaches the maximum in that direction. Using the path of steepest ascent improves the region of the optimum and a flow chart of the response surface is constructed, as illustrated in Figures 2.2 and 2.3, respectively.



**Figure 2.2** Path of Steepest Ascent Improvement in the Region of the Optimum



**Figure 2.3** Flow Chart showing the Construction of the Response Surface

The method of steepest ascent is a procedure useful for moving sequentially along the path of steepest ascent. Using this method, the direction of potential improvement is essentially a "path" to find the region or neighborhood of the optimum. Consider the fitted first-order response surface model

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i, \quad (2.8)$$

where  $\hat{y}$  is the predicted response,  $x_i$  represents the  $i^{\text{th}}$  independent factor,  $\hat{\beta}_0$  is the estimated intercept, and the individual  $\hat{\beta}_i$ 's are the estimated coefficients for the  $i^{\text{th}}$  independent variable. The method of steepest ascent attempts to fit the points of a set independent variables that results in the maximum estimated response over all points that are a fixed distance  $r$  from the center of the design. As a result, the optimization problem involves the use of a Lagrange multiplier ( $\lambda$ ) subject to the restraint given by

$$\sum_{i=1}^k x_i^2 = r^2. \text{ Take partial derivatives with respect to } x_j,$$

$$L = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k - \lambda \left( \sum_{i=1}^k x_i^2 - r^2 \right), \quad (2.9)$$

and set them equal to zero,  $\frac{\partial L}{\partial x_j} = \hat{\beta}_j - 2\lambda x_j = 0$ . Next, obtain the coordinate of  $x_j$

along the path of steepest ascent as  $x_j = \frac{\hat{\beta}_j}{2\lambda}$ .

The expression  $\frac{1}{2\lambda}$  is set to be the constant of proportionality  $\rho$  in

$$x_1 = \rho \hat{\beta}_1, x_2 = \rho \hat{\beta}_2, \dots, x_k = \rho \hat{\beta}_k. \quad (2.10)$$

By considering the step size in one of the variables (selecting the largest estimated coefficient of the independent variables), the step size of steepest ascent can be found in other variables using the relationship in (2.9); The positive (negative) value of  $\rho$  in (2.10) is used when searching for the minimum (maximum) on the response surface. The origin of independent variables is moved to the point given by (2.9) which is the point where the largest absolute change of (2.8) occurs on the hyper-sphere of radius  $r$ . The value of  $x_j$  in (2.9) can be considered as an increment of  $x_j$ ,  $\Delta x_j$ , moving away from the origin. Let  $|\beta_i| = \max_{1 \leq j \leq k} |\beta_j|$ . The increment of  $x_j$ ,  $\Delta x_j$ , in term of the largest absolute increment  $\Delta x_j$  can be written as

$$\Delta x_j = \frac{\hat{\beta}_j}{\hat{\beta}_i / \Delta x_i} \quad j = 1, 2, \dots, k, i \neq j. \quad (2.11)$$

The incremental  $\Delta x_j$  in (2.11) is converted back to the increment of the corresponding natural variable before running the next experiment. After the experiment moves toward the neighborhood of the optimum along the steepest direction, step by step given by (2.11), until a curvature is detected by direct comparison of the successive responses, the  $2^k$  vertices of the current simplex with a number of center points are performed as in the initial experiment to test the hypothesis of the existence of a curvature. If the test cannot reject the null hypothesis, a number of axial points by using the central composite design (CCD) are added in the experiment to construct the second order model,  $d = 2$ , for estimating the optimum design point. The second order model consists of  $k$  first order terms,  $k$  second order terms and  $k(k-1)/2$  interaction terms and can be written as

$$\Delta x_j = \frac{\hat{\beta}_j}{\frac{\hat{\beta}_i}{\Delta x_i}} \quad j = 1, 2, \dots, k, i = j. \quad (2.12)$$

which is used to approximate the response surface in the neighborhood of the optimum. After the parameters of the second order model are estimated, the optimum design point in terms of the coded variables can be written as (Montgomery 2009)

$$\mathbf{x}_{opt} = -\frac{1}{2} \mathbf{B}^{-1} \mathbf{b}, \quad (2.13)$$

where

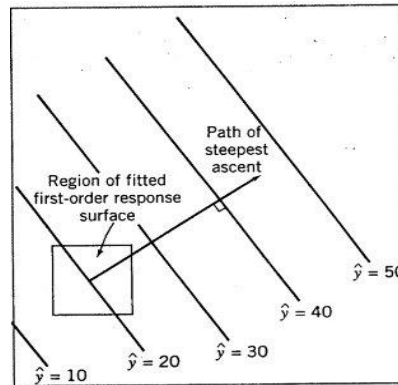
$$\mathbf{b} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \cdots & \hat{\beta}_{1k}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} & \cdots & \hat{\beta}_{2k}/2 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\beta}_{1k}/2 & \hat{\beta}_{2k}/2 & \cdots & \hat{\beta}_{kk} \end{bmatrix},$$

and the predicted response at the optimum design point is given by

$$\hat{y}_{opt} = \hat{\beta}_0 + \frac{1}{2} \mathbf{x}'_{opt} \mathbf{b}, \quad (2.14)$$

which is to be converted back in terms of the natural variables.

Typically, the steepest ascent is efficient when the initial experiment is far from the region with optimum conditions (see Figure 2.4). It is in these beginning regions that the first-order approximations are most reasonable. When the experiment has nearly reached the optimum, interactions and pure quadratics begin to dominate, and, in this situation, the steepest ascent method becomes less effective. At this point, the response surface model requires a fitted second-order method for finding the optimum conditions in the response variable.



**Figure 2.4** First-order Response Surface and Path of Steepest Ascent

### 2.3.1 Steps in the Steepest Ascent Method

1) Use an orthogonal design to fit a first-order model. Designs with two levels are appropriate, and center runs are often recommended.

2) Compute the path of steepest ascent /descent so one may expect the maximum increase or decrease in response. The path is considered to be proportional to the regression coefficients (  $\hat{\beta}_i$  ).

3) Experimental runs are conducted along the path, the results of which should show improving response values. At some region along the path, the improvement will decline and eventually disappear.

4) At some location where an approximation of the maximum response is located on the path, choose a base for the second experiment. It is reasonable that center runs for testing curvature and degrees of freedom for interaction-type lack of fit are important at this point.

5) Conduct a second experiment with the fitted first-order model. Lack of fit is tested to find the point where it is no longer significant, at which point a second path based on the new model is computed.

## 2.4 Second-order Models

It is well known that response surface methodology is an experimental technique invented to find the optimal response within specified ranges of factors. If curvature is detected (lack of fit to the first-order model), then we need to add points in order to obtain a second-order design near the optimum. The most popular second-order designs are CCD and the Box-Behnken design, which are used to fit a second-order prediction equation for the response, and the quadratic terms in these equations model the curvature of the path on the true response function. If there exists a maximum or minimum inside the factor region, then response surface methodology can estimate it reasonably.

The second-order model can be expressed as follows:

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j + \varepsilon ,$$

where  $y \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$  and  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ .

### 2.4.1 Central Composite Design (CCD)

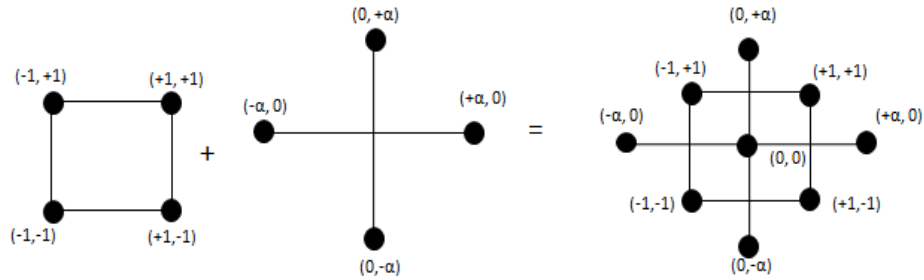
Myers et al. (1989) came up with an orthogonal design motivated by Box and Wilson (1951) in the case of the first-order model. For second-order models, many subject matter scientists and engineers have a working knowledge of CCD and the three-level designs of Box and Behnken (1960). In addition, the same researchers

stated that another important contribution came from Hartley (1959) who made an effort to create a more economical or compact composite design.

Box and Wilson's (1951) CCD has become the most popular choice for fit a second-order model. This design contains a factorial or fractional factorial design with center points that are augmented with a group of star points (axial points) that allow estimation of curvature. If the distance from the center of the design space to a factorial point is  $\pm 1$  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  $\alpha$  depends on certain properties desired for the design and on the number of factors involved. Figure 2.5 shows the generation of a CCD for two factors.

Assuming  $k \geq 2$  design variables, the CCD consists of:

- 1) An  $f = 2^{k-p}$  full ( $p=0$ ) or fractional ( $p>0$ ) factorial design; each point is of the form  $(x_1, \dots, x_k) = (\pm 1, \dots, \pm 1)$ .
- 2)  $2k$  axial or star points of the form  $(x_1, \dots, x_i, \dots, x_k) = (0, \dots, 0, \pm\alpha, 0, \dots, 0)$  for  $1 \leq i \leq k$ .
- 3)  $n_c$  center points  $(x_1, \dots, x_k) = (0, 0, \dots, 0)$ .



**Figure 2.5** Generation of a CCD for Two Factors

The CCD contains  $N = f + 2k + n_c$  points, and the  $\binom{k+2}{n}$  parameters are estimated in the second-order model. Each of the three types of points in a CCD plays a different role:

1) Estimation of the first-order and interaction terms is used by the factorial points.

2) Estimation of the squared terms is used by the axial points.

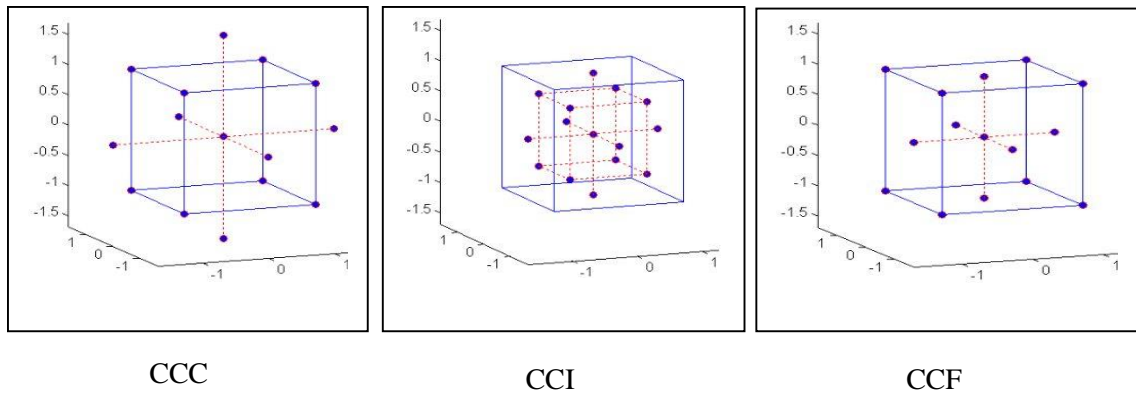
3) The center points provide an internal estimate of pure error used to test for lack of fit and also contribute towards the estimation of the squared terms.

The different CCD types are shown in Table 2.5 and Figure 2.6.

**Table 2.5** CCD Types

CCD Type	Terminology	Comments
Circumscribed	CCC	The original form of CCD is a central composite circumscribed (CCC) design. The distance $\alpha$ from the center based on the properties desired for the design results in a star point for each factor. New extremes for the low and high settings for each factor are established using star points. These designs require five levels for each factor and have circular, spherical, or hyper spherical symmetry. Augmenting the star points can result in an existing factorial or resolution V fractional factorial design.
Inscribed	CCI	This design also requires five levels for each factor, but has specific limits for factor settings which are true limits in a specific situation. The central composite inscribed (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 latter divided by $\alpha$ to generate the CCI design).
Face-centered	CCF	The central composite face-centered (CCF) design requires three levels for each factor. The star points $\alpha = \pm 1$ are produced at the center of each face of the factorial space. Augmenting an existing factorial or resolution V design with appropriate star points can also produce this design.





**Figure 2.6** The Three CCD Types with Three Factors

**Table 2.6** Structural Comparison of CCC (CCI), CCF, and BBD for Three Factors

CCC (CCI)				CCF				BBD			
Rep	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	Rep	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	Rep	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
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	0	1
	1	1	1	1	1	1	1	1	1	0	1
	-1.682	0	0	1	-1	0	0	1	0	-1	-1
	1.682	0	0	1	1	0	0	1	0	1	-1
	0	0.682	0	1	0	-1	0	1	0	-1	1
	0	1.682	0	1	0	1	0	1	0	1	1
	0	0	0.682	1	0	0		3	0	0	0
1	0	0	1.682	1	0	0	1				
6	0	0	0	6							
Total Runs = 20				Total Runs = 20				Total Runs = 15			

## 2.5 Experimental Design Properties

Box and Hunter (1957) suggested that a desirable experimental design should consist of two parts as follows: the first judged partly on the precision of the estimates of the regression coefficients, and the second partly on the magnitude of the bias of

those estimates. The researchers list the following qualities as desirable in the experimental design:

- 1) The design should estimate the assumed model within the region of interest.
- 2) The design should have a built-in check on the assumed model.
- 3) The design should not have a large number of experimental points.
- 4) The design should be blockable.
- 5) The design should be easily expanded.
- 6) The design should have properties such as orthogonality and rotatability.

### 2.5.1 Orthogonality

The important concept in a DOE is orthogonality because it can allude to independence. At the beginning of creating a DOE, each column corresponds to a different factor. Therefore, if every single column in the design is orthogonal, we ensure that each factor is independently estimated with regard to every other factor. Consider a  $2^3$  full factorial with eight runs:

Factors		
A	B	C
–	–	–
+	–	–
–	+	–
+	+	–
–	–	+
+	–	+
–	+	+
+	+	+

First, three factors are multiplied:  $A \cdot B$ ,  $A \cdot C$ , and  $B \cdot C$  to ensure that each column (vector) is orthogonal to every other column:

$$\begin{aligned}
A \cdot B &= 1(-1) + 1(-1) - 1(1) + 1(1) - 1(-1) + 1(-1) - 1(1) + 1(1) = 0, \\
A \cdot C &= -1(-1) + 1(-1) - 1(-1) + 1(-1) - 1(1) + 1(1) - 1(1) + 1(1) = 0, \\
B \cdot C &= -1(-1) - 1(-1) + 1(-1) + 1(-1) - 1(1) - 1(1) + 1(1) + 1(1) = 0.
\end{aligned}$$

Therefore, factors A, B and C are estimated independently.

### 2.5.2 Rotatability

An important property for second-order designs is rotatability. This property requires that the scaled prediction variance is equal for all points  $x$  that are the same distance from the center of the design. Box and Hunter (1957) found that first-order orthogonal designs, such as factorial designs, are rotatable. For second-order designs, two conditions are required for rotatability:

- 1) All odd moments through order four are zero.
- 2) The ratio of moments  $[iiii]/[iijj] = 3 \ (i \neq j)$ .

Therefore, a rotatable CCD requires that

$$\frac{[iiii]}{[iijj]} = \frac{F + 2\alpha^4}{F} = 3 \quad \text{or} \quad \alpha = \sqrt[4]{F}.$$

The second-order model

$$\eta(x) = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \beta_{ij} x_i x_j$$

may be written in matrix notation as

$$\eta(x) = x_s' \beta,$$

where

$$x' = (x_1, x_2, \dots, x_k),$$

$$x_s' = (1, x_1, x_2, \dots, x_k, x_1^2, x_2^2, \dots, x_k^2, x_1 x_2, \dots, x_{k-1} x_k),$$

$\beta$  is an  $rn \times 1$  column vector,

$\hat{y}(x)$  is the predicted response value at a particular point  $x' = (x_1, x_2, \dots, x_k)$ ,

$X$  is the  $N \times m$  matrix of values of the elements of  $x_s$ 's taken at the design points, and

$y$  is the  $N \times 1$  matrix of observations.

If the prediction variance  $Var[\hat{y}(x)] = x_s'(X'X)^{-1}x_s\sigma^2$  is constant at all points, then that design matrix is said to be rotatable equidistant from the design center, which, by proper coding of the control variables, can be chosen to be the point at the origin of the  $k$ -dimensional coordinates system. We can say that if the design is rotatable,  $Var[\hat{y}(x)]$  is constant at all points that fall on the surface of a hypersphere centered at the origin. This property is advantageous in that the prediction variance remains unchanged under any rotation of the coordinate axes. In addition, if optimization of  $\hat{y}(x)$  is desired on concentric hyperspheres, such as in the application of ridge analysis, then it is desirable for the design to be rotatable, which can easily be determined by comparing the values of  $\hat{y}(x)$  on a given hypersphere since all such values have the same variance.

## 2.6 Canonical Analysis

A contour plot and canonical analysis can be useful to study the shape of response when a process has only two or three process variables, which is easiest to consider with a contour plot. Canonical analysis is helpful first to transform the model into a new coordinate system with the origin at the stationary point  $x$  and then to rotate the axes of this system until they are parallel to the principal axes of the fitted response surface. In addition, canonical analysis is used to investigate the overall shape of the curvature and determine whether the stationary point is a maximal, minimal, or saddle point (see Figure 2.9). The eigenvalues and eigenvectors indicate the shape of the response surface, and the fitted second-order model, in matrix notation, is written as

$$\hat{y} = \beta_0 + x'b + x'\hat{B}x. \quad (2.12)$$

The intercept, linear, and second-order coefficients are estimated and denoted by  $\beta_0$ ,  $b$ , and  $\hat{B}$ , respectively. For the stationary point, we can differentiate  $\hat{y}$  in equation (2.12) with respect to  $x$  by setting the derivative equal to zero, that is

$$\frac{\partial \hat{y}}{\partial x} = b + 2\hat{B}x = 0. \quad (2.13)$$

The stationary point can be obtained in the form

$$x_s = -\frac{1}{2}\hat{B}^{-1}b, \quad (2.14)$$

where  $\Lambda$  is a diagonal matrix containing the eigenvalues of  $\hat{B}$ .

In factor analysis, rotation of canonical weights can improve interpretability of the stationary point solution. If the second-order model is transformed to a new center, i.e. the stationary point,  $z = x - x_s$ , and its axis rotated with respect to  $C$ , which is a  $k \times k$  matrix whose columns are normalized eigenvectors associated with the eigenvalues  $w = C'z$ , then we obtain

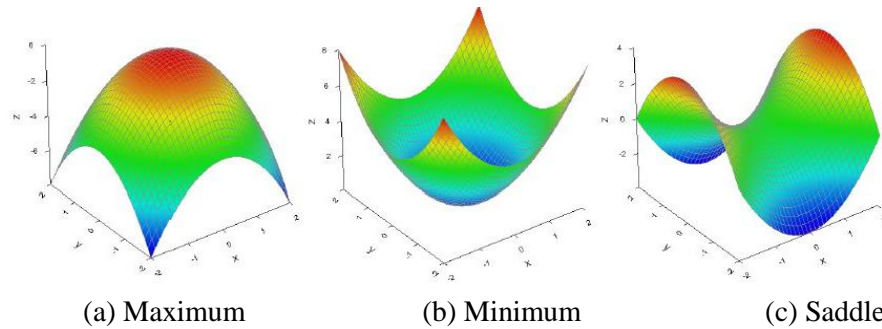
$$\begin{aligned} \hat{y} &= \beta_0 + (z + x_s)'b + (z + x_s)'\hat{B}(z + x_s) \\ &= \hat{y}_s + z'\hat{B}z \end{aligned}$$

The rotation is given by

$$\begin{aligned} \hat{y} &= \hat{y}_s + w'P'\hat{B}Pw \\ &= \hat{y}_s + w'\Lambda w \\ &= \hat{y}_s + \sum_{i=1}^k \lambda_i w_i^2 \end{aligned}$$

where  $\hat{y}_s$  is the estimated response at the stationary point and  $\lambda_1, \lambda_2, \dots, \lambda_k$  are the eigenvalues of  $\hat{B}$ . The eigenvalue decomposition is used to describe the nature of the stationary point by canonical analysis. The signs of the eigenvalues give the different types of the response system as follows:

- 1) If all eigenvalues are negative, the stationary point yields a maximum response.
- 2) If all eigenvalues are positive, the stationary point yields a minimum response.
- 3) If the eigenvalues have mixed signs, the stationary point yields a saddle point.

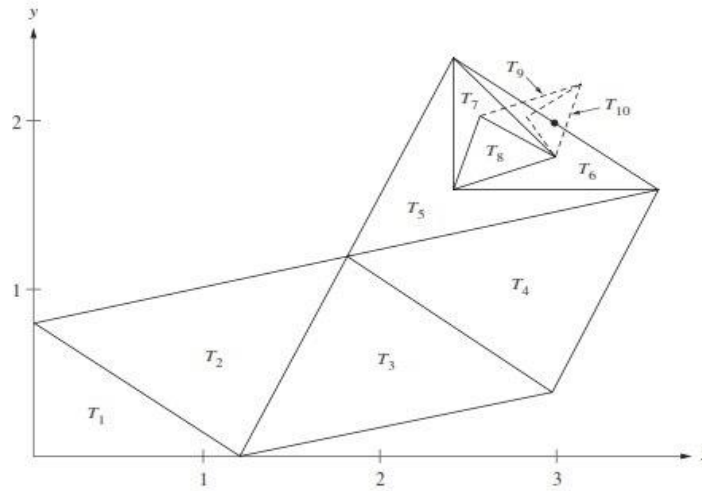


**Figure 2.7** Types of Response Surface

## 2.7 The NM Method

Nelder and Mead (1965) introduced the NM method, which is a heuristic algorithm for multidimensional unconstrained optimization problems. This method is a very popular derivative-free technique useful for finding the local minimum of a function. The minimization of a function of  $n$  variables by this method is dependent on a comparison of function values at the  $(n+1)$  vertices of a general simplex. The simplex adapts itself to the local landscape and contracts to the final minimum. For a two-dimensional problem, a simplex is a triangle, and, with this method, the function values at the three vertices of the triangle are compared using a pattern search. The worst vertex, where  $f(x, y)$  is largest, is rejected and replaced with a new vertex. A new triangle is formed and the search for a better outcome is continued. A sequence of triangles in the process is generated for which the function values at the vertices get smaller and smaller. The sizes of the triangles are iteratively reduced and the coordinates of the minimum point found (see Figure 2.8). When the response at a reflected point is not improved, the algorithm has special rules for cases where an

additional expanded reflection gives improvement; these special rules cause the simplex to either shrink or expand. Therefore, it is also referred to as the flexible simplex algorithm. The simplex algorithm can easily be extended to higher dimensions (Nelder and Mead, 1965).



**Figure 2.8** The Sequence of Triangles  $\{T_k\}$  converging to the Point (3,2) for Minimization of a Function of Two Variables using the NM Method

Box (1966), and Parkinson and Hutchinson (1972), have suggested using the Nelder-Mead simplex search (NMSS), although they stated that this method is less effective in optimum seeking as the number of dimensions increases. Myers and Montgomery (2002), and Olsson and Nelson (1975), mentioned that for the usual problems arising from response surface methodology practice, the number of influential process factors included in the final model is rarely larger than six. The earlier phase eliminates the unimportant factors, so the remaining number of significant ones is small. By using a “pre-experiment” such as a fractional factorial design, if properly devised, the NMSS can be expected to work well for stochastic response surface optimization. This particular case holds when the response surface has not been fitted because it is too complex to provide an adequate physical model, so the gradient information is unavailable. The alternative is to use NMSS to sequentially optimize the “actual” process response.

Spendley, Hext and Himsworth (1962) propounded the first simplex search scheme later developed by Nelder and Mead (1965). One of most popular “direct-search” techniques is the NMSS method for unconstrained optimization that requires no gradient computation of the function to be minimized. In essence, this procedure is of the steepest descent kind. First,  $k + 1$  vertices of the response function are evaluated under exploration in a procedure called an initial simplex, which is a polyhedron in the factor space of  $k$  (independent variables). Its purpose is to minimize the function by using a descent direction to move the search course away from the vertex with the worst (or highest) function value, yielding a newly reflected, possibly improved point that is located in the neighborhood. When optimizing the NMSS process, the simplex adapts itself to the local surface landscape by varying its size and orientation continuously, so NMSS is especially suitable for exploring “unwieldy” terrains. The current sequential technique for unconstrained optimization is NMSS, which has been widely accepted as the most robust and efficient.

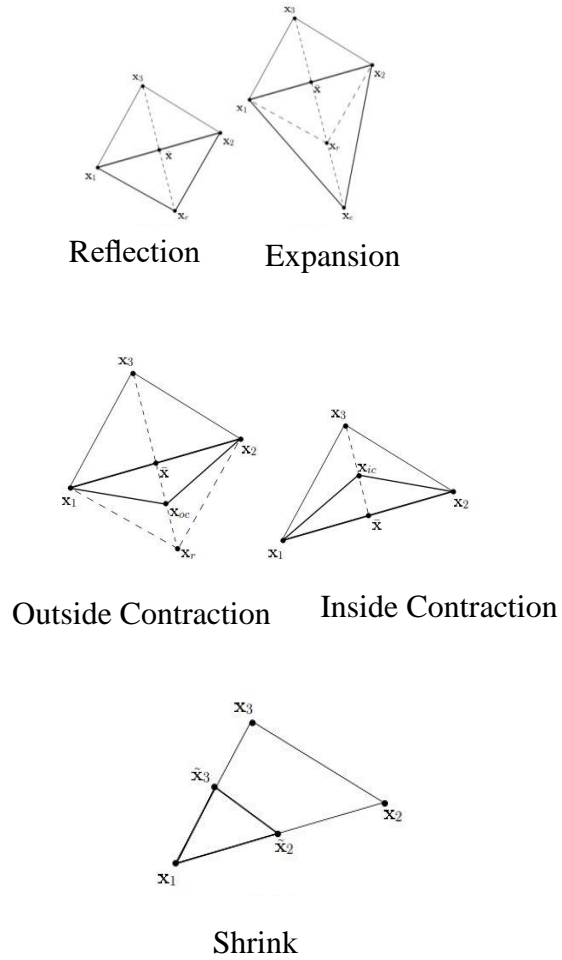
The NMSS method is a technique for solving the unconstrained optimization problem  $\min f(x)$ , where  $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$  is called the objective function with  $n$  dimensions. A simplex is described as a geometric figure with  $n$  dimensions that is the convex hull of  $n+1$  vertices, i.e. a simplex with vertices of  $x_1, x_2, \dots, x_{n+1}$  denoted by  $\Delta$ . The method iteratively generates a sequence of simplexes to approximate an optimal point of  $\min f(x)$ ; the simplexes are ordered according to the objective function values with  $f(X_1) \leq f(X_2) \leq \dots \leq f(X_{n+1})$  for each iteration, with corresponding vertices  $\{X_j\}_{j=1}^{n+1}$ . Let  $X_1$  refer to the best vertex, and  $X_{n+1}$  refer to the worst vertex.

Four possible operations are determined in the algorithm: reflection, expansion, contraction, and shrink, each being associated with a scalar parameter:  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ , respectively (see Figure 2.9). The values of these parameters satisfy  $\alpha > 0, \beta > 1, 0 < \gamma < 1$  and  $0 < \delta < 1$ . In the standard implementation of the NM method, the parameters are chosen to be  $\{\alpha, \beta, \gamma, \delta\} = \{1, 2, 1/2, 1/2\}$ . Let  $\bar{X}$  be the centroid of the  $n$  best vertices computed by  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ .



### 2.7.1 An Iteration of the NM Algorithm

- 1) Sort. Evaluate  $f$  at the  $n+1$  vertices of  $\Delta$  and sort the vertices so that  $f(X_1) \leq f(X_2) \leq \dots \leq f(X_{n+1})$  holds.
- 2) Reflection. Compute the reflection point with  $X_r$  by  $X_r = \bar{X} + \alpha(\bar{X} - X_{n+1})$ . Evaluate  $f_r = f(X_r)$ . If  $f_1 \leq f_r < f_n$ , replace  $X_{n+1}$  with  $X_r$  so that  $X_r = \bar{X} + \alpha(\bar{X} - X_{n+1})$ .
- 3) Expansion. If  $f_r < f_1$ , then the expansion point is computed. Let  $X_e = \bar{X} + \beta(X_r - \bar{X})$  and evaluate  $f_e = f(X_e)$ . If  $f_e < f_r$ , replace  $X_{n+1}$  with  $X_e$ , otherwise replace  $X_{n+1}$  with  $X_r$ .
- 4) Outside Contraction. If  $f_n \leq f_r < f_{n+1}$ , complete the outside contraction point by  $X_{oc} = \bar{X} + \gamma(X_r - \bar{X})$  and evaluate  $f_{oc} = f(X_{oc})$ . If  $f_{oc} \leq f_r$ , replace  $X_{n+1}$  with  $X_{oc}$ , otherwise go to step 6.
- 5) Inside Contraction. If  $f_r \geq f_{n+1}$ , the inside contraction expansion point is computed by  $X_{ic} = \bar{X} - \gamma(X_r - \bar{X})$  and evaluate  $f_{ic} = f(X_{ic})$ . If  $f_{ic} < f_{n+1}$ , replace  $X_{n+1}$  with  $X_{ic}$ , otherwise go to step 6.
- 6) Shrink. For  $2 \leq i \leq n+1$ , define  $X_i = X_1 + \delta(X_i - X_1)$
- 7) Stopping Criterion. This is based on a comparison of function values originally considered with  $|f(x_{n+1}) - f(x_1)| < \varepsilon$ , where  $\varepsilon = 1 \times 10^{-6}$ .



**Figure 2.9** Possible Operations Performed on a Simplex in  $\mathcal{R}^2$

## 2.8 Coverage Probability

A confidence interval (CI) as a region constructed under the assumption that the model contains the “true” value (the parameter of the model) with a specified probability. This region is constructed using the particular properties of an estimator and takes into account both the accuracy and precision of this estimate. There are two quantities associated with confidence intervals:

Coverage probability: this term can be explained with the probability that a procedure for constructing random regions produces an interval containing, or covering, the true value. This quantity is the chance that the parameter of interest is

covered by the interval constructed. It is independent of the particular sample to which such a procedure is applied and is a property of the interval producing procedure.

Confidence level: any particular sample produces the interval with a procedure with coverage probability  $p$  can be said to have a confidence level of  $p$ , hence the term ‘confidence interval’. Both of these definitions, the confidence level and coverage probability, are equivalent before obtaining a sample. After considering whether a parameter is either in or not in the interval, the interval containing the parameter is either 0 or 1, respectively. Thus, if constructed with 95% of 95% CI, this will cover the parameter under repeated sampling. Of course, for any particular sample, we do not know if the CI produced contains the true value.

## 2.9 Mean Absolute Percentage Error (MAPE)

MAPE is a useful statistic to compare fitted values obtained using different methods. A method with a lower value is usually indicative of a better fitting model over another;

$$\text{MAPE} = \frac{\sum_{i=1}^n (|y_i - \hat{y}_i| / y_i)}{n} .$$

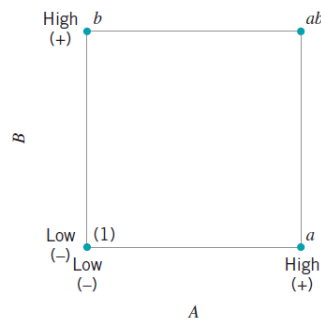
## CHAPTER 3

### THE PROPOSED METHOD TO FIT A SECOND-ORDER MODEL

In this Chapter, the proposed method for finding the optimum points for a second-order model in order to obtain a best fit for it is described. There are two steps, the details of which are covered in Sections 3.1 and 3.2.

#### 3.1 The Seeding Step for the Proposed Method

A  $2^k$  factorial design is used to begin the process and, for the case of  $k=2$ , a  $2^2$  factorial design, is shown in Figure 3.1. Both the CCD classical method and the proposed method have the same starting points set by using a  $2^k$  factorial design.



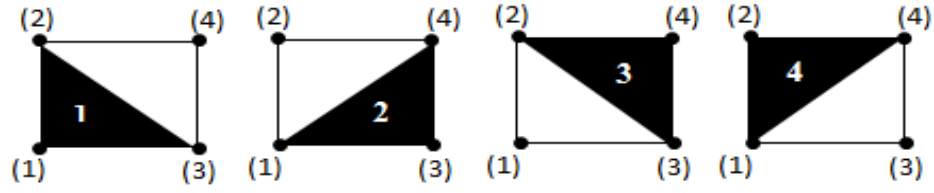
a) Design Geometry

A	B
-	-
+	-
-	+
+	+

b) Design Matrix

**Figure 3.1** The  $2^2$  Factorial Design as a Starting Point

An illustration of all possible combination starting points in four cases for  $k=2$  is shown in Figure 3.2.



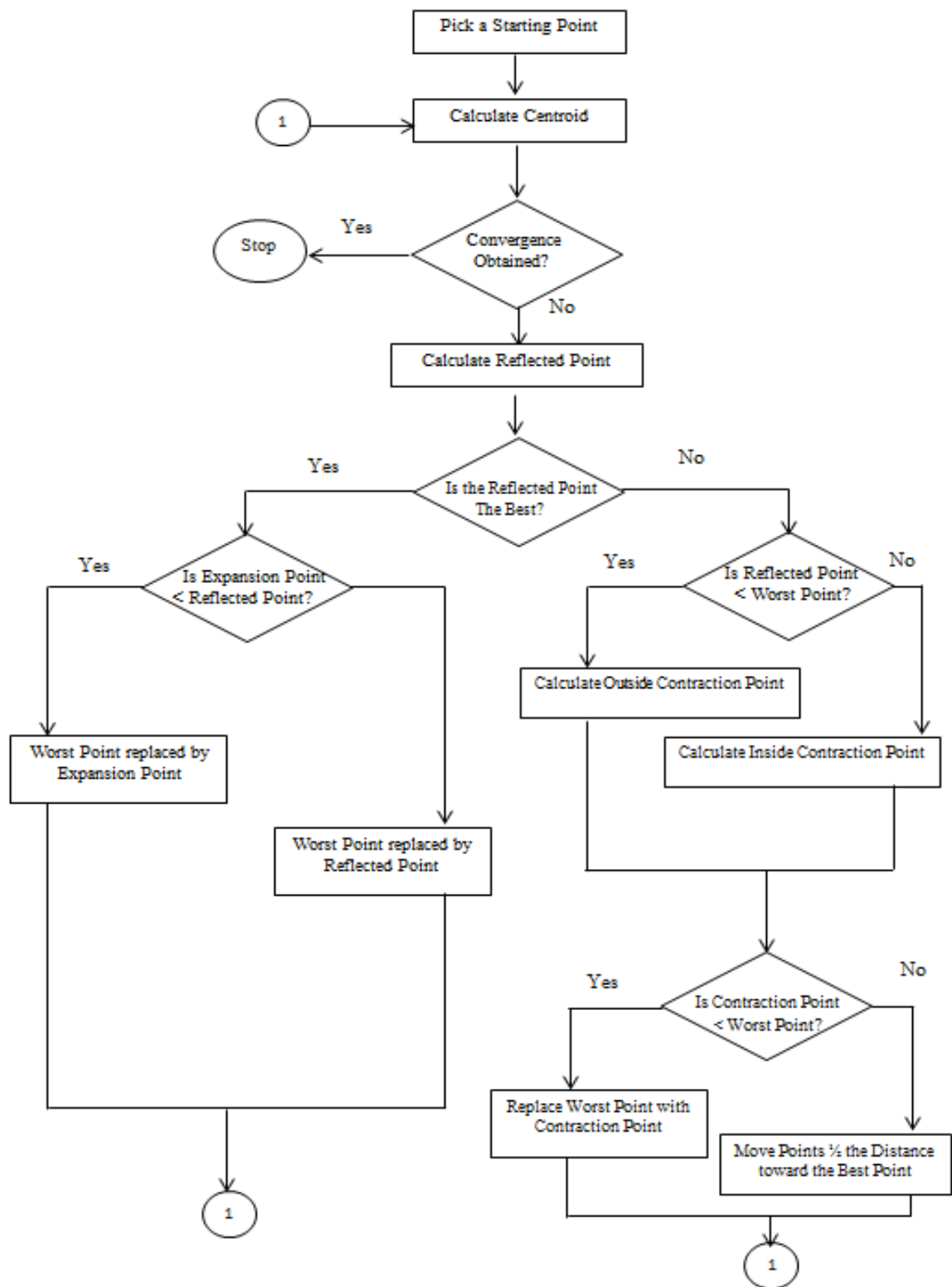
**Figure 3.2** Four Possible Starting Points in the First Phase of the Proposed Method

The possible four cases of starting points are as follows: case 1 (points (1), (2), and (3)), case 2 (points (1), (3), and (4)), case 3 (points (2), (3), and (4)), and case 4 (points (1), (2), and (4)). The initial simplex is set up using the four cases above, and so possibly consists of the four types.

### 3.2 The Second Step in the Proposed Method

The steps in the proposed method are shown in Figure 3.1. After using the  $2^k$  factorial design for setting the starting points for random without replacement  $k+1$  points, all possible starting points are noted. The NM algorithm (see Figure 3.3) is applied to each case in order to move to the optimum region with a stopping criterion based on a comparison of  $|f(x_{n+1}) - f(x_1)| < \varepsilon$ , where  $\varepsilon = 1 \times 10^{-1}$ .

After this, the experimental points are obtained and used to fit the second-order model. After the second-order model has been obtained, the stationary points are computed.



**Figure 3.3** Flow Chart of the NM Algorithm

## CHAPTER 4

### SIMULATION STUDY

Five mathematical test functions were used to compare the performance of the classical RSM (CCD) to the proposed method; they are expressed as follows:

$$f_1(x_1, x_2) = 2 + 0.01(x_2 - x_1^2)^2 + (1 - x_1) + 2(2 - x_2)^2 + 7 \sin\left(\frac{x_1}{2}\right) \sin\left(\frac{7x_1x_2}{10}\right); x_1, x_2 \in [1, 4]$$

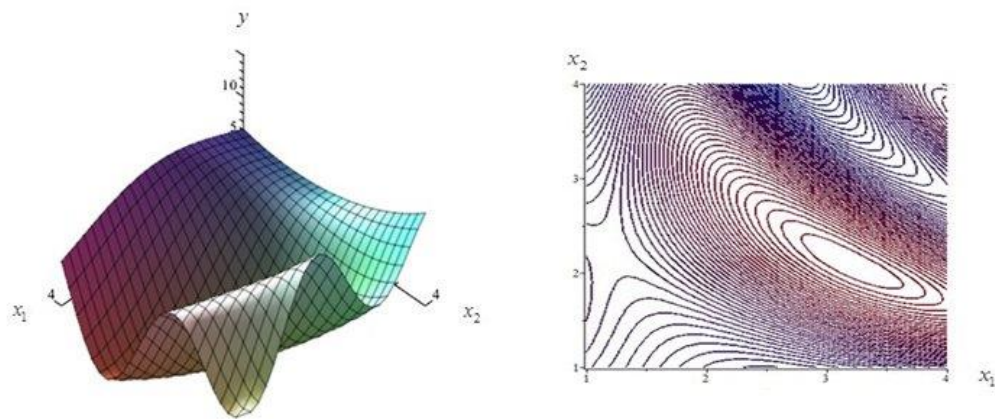
$$f_2(x_1, x_2) = -(x_1^2 + x_2 - 11)^2 - (x_1 + x_2^2 - 7)^2; x_1, x_2 \in [-2, 2]$$

$$f_3(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4; x_1 \in [-1, 0.5], x_2 \in [0, 1]$$

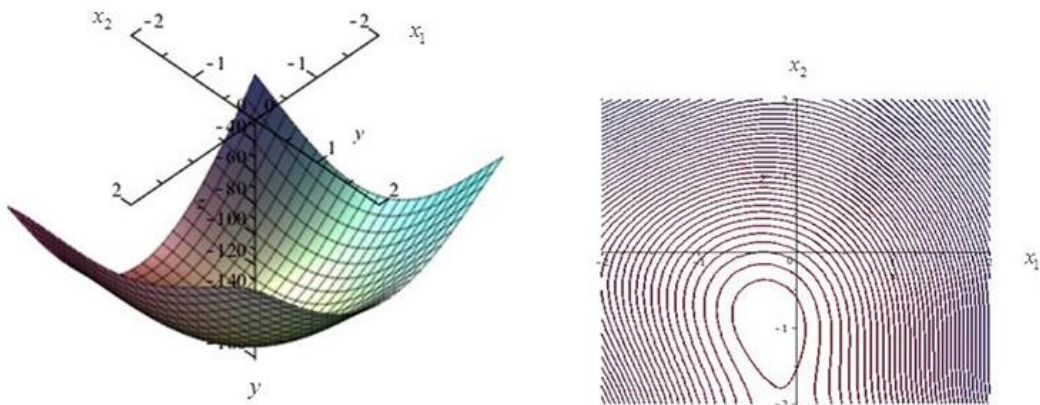
$$f_4(x_1, x_2) = x_1 \sin(4x_1) + 1.1x_2 \sin(2x_2); x_1, x_2 \in [1.5, 3.5]$$

$$f_5(x_1, x_2) = 1431 - 7.81x_1 - 13.3x_2 + 0.0551x_1^2 + 0.0401x_2^2 - 0.01x_1x_2; x_1 \in [50, 120], x_2 \in [150, 200]$$

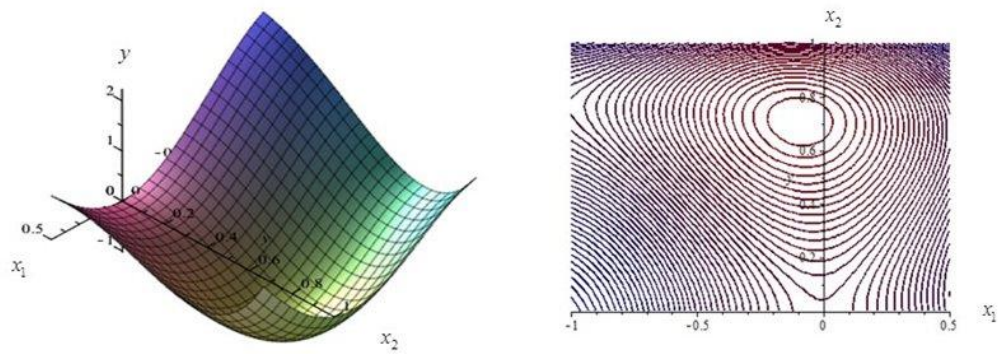
Figures 4.1-4.5 illustrate the response surface and contour of each mathematical test function. Table 4.1 shows the minima and the minimum response of the test functions.



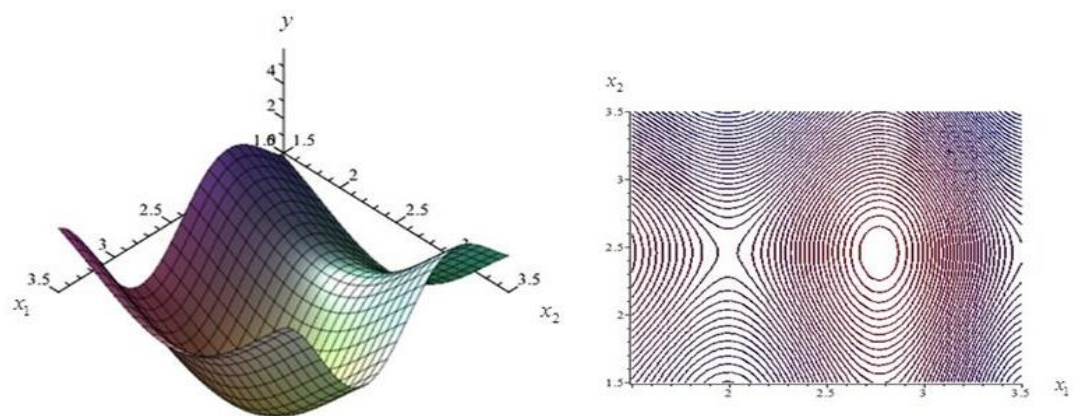
**Figure 4.1** The Response Surface and Contour of  $f_1$



**Figure 4.2** The Response Surface and Contour of  $f_2$

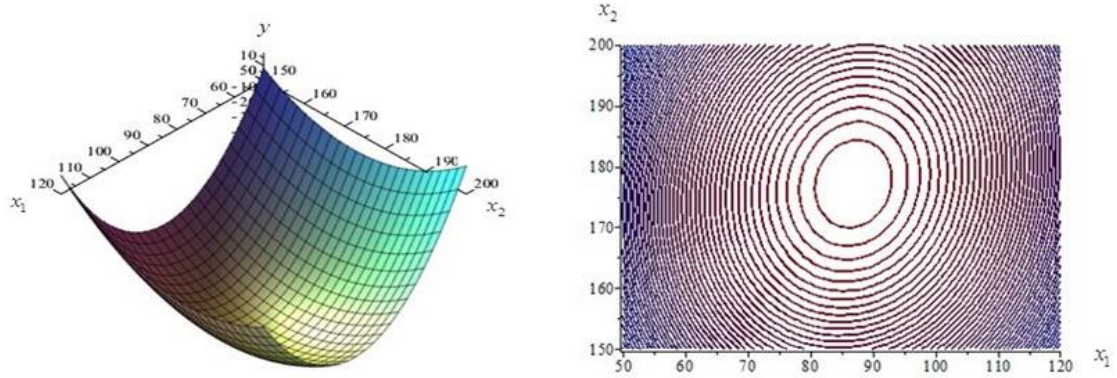


**Figure 4.3** The Response Surface and Contour of  $f_3$



**Figure 4.4** The Response Surface and Contour of  $f_4$





**Figure 4.5** The Response Surface and Contour of  $f_5$

**Table 4.1** The Minima and the Minimum Response of the Test Functions.

Test function	Minimum Points		Minimum
	$x_{\min_1}$	$x_{\min_2}$	$y_{\min}$
$f_1$	3.20	2.10	-6.51
$f_2$	-0.27	-0.92	-181.6
$f_3$	-0.092	0.713	-1.032
$f_4$	2.77	2.46	-5.41
$f_5$	86.9	176.67	-83.22

#### 4.1 Steps in the Simulation Study

**Step 1** Independent variables for each mathematical test function were generated with a uniform distribution. Values of the independent variables were randomly selected from the ranges. The functions  $f_1$  to  $f_5$  have  $x_1, x_2 \in [1, 4]$ ,  $x_1, x_2 \in [-2, 2]$ ,  $x_1, x_2 \in [1.5, 3.5]$  and  $x_1, x_2 \in [-2, 2]$ ,  $x_1 \in [50, 120]$ ,  $x_2 \in [150, 200]$  respectively.

**Step 2** Classical RSM (CCD) was constructed starting with a factorial design. For functions  $f_1$  to  $f_5$ ,  $x_1, x_2$  are both called factors in the experiment. They were generated using  $X_i \sim U(a_i, b_i); i=1,2$ , supported by  $X_i \in [a_i, b_i]$ , mean  $= \frac{1}{2}(a_i + b_i)$ , and variance  $= \frac{1}{12}(b_i - a_i)^2$ , where  $a_i$  and  $b_i$  are the lower and upper bound of  $x_i$ , respectively.

**Step 3** 50 replications were carried out for each.

- 1) Set up  $L_{ij} \in \left[ a_{ij}, \frac{1}{2}(a_{ij} + b_{ij}) \right]; i=1,2; j=1,2,\dots,50$ .
- 2) Set up  $\delta_{ij} \in [0, b_{ij} - \frac{1}{2}(a_{ij} + b_{ij})]; i=1,2; j=1,2,\dots,50$ .
- 3) Set up  $H_{ij} \in [L_{ij} + \delta_{ij}]; i=1,2; j=1,2,\dots,50$ .
- 4) Set up  $\varepsilon_i \sim N(0, \sigma^2)$ .

For  $f_1$   $\varepsilon_i \sim N(0, 0.01^2)$

For  $f_2$   $\varepsilon_i \sim N(0, 1^2)$

For  $f_3$   $\varepsilon_i \sim N(0, 0.01^2)$

For  $f_4$   $\varepsilon_i \sim N(0, 0.01^2)$

For  $f_5$   $\varepsilon_i \sim N(0, 1^2)$

**Step 4** From step 3, where  $i = 1$  to 50, 50 factorial designs were obtained to begin with. Five center points were added for each factorial design. The center point in a  $2^2$  factorial design  $C_{ij}$  is given by  $(L_{ij} + H_{ij})/2$ , for  $i=1,2, j=1,2,\dots,50$ .

Curvature was tested for and, if found, it was deemed necessary to add an axial point. However, if it was not found, the steepest ascent was used to move to a new region until curvature was found.

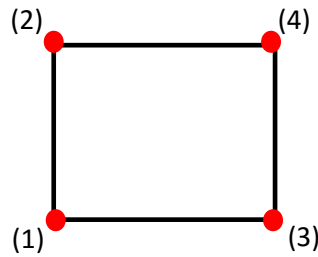
Where the axial point is computed by  $\alpha = [\text{number of factorial run}]^{1/4}$  for a factorial design with  $k=2$ ,  $\alpha = [2^2]^{1/4} = 2^{2/4} = 2^{1/2} = 1.414$ .

After that, the second-order procedure was performed and the parameters of the model estimated by the ordinary least squares (OLS), then the stationary points

$x_1, x_2$ , and  $y$  were obtained. Following this, the 95% confidence interval was constructed for the means of the stationary points  $x_1, x_2$ , and  $y$ .

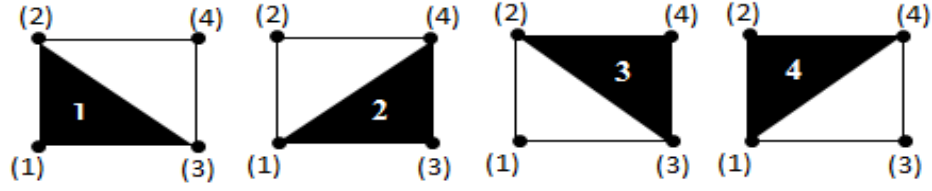
**Step 5** From steps 3 and 4, only one 95% confidence interval for the means of stationary points  $x_1, x_2$  and  $y$  were selected. Steps 1-4 were iterated 100 times, producing a 95% confidence interval for each mean of the stationary points  $x_1, x_2$  and  $y$ , resulting in one hundred CIs of average the stationary points  $x_1, x_2$  and  $y$  confidence intervals.

**Step 6** First, the classical RSM (CCD) was performed to create a  $2^2$  factorial design (see Figure 4.6).



**Figure 4.6**  $2^2$  Factorial Design

The performance of the proposed method was tested with the same starting points as the classical RSM (CCD), and used a  $k+1$  simplex for moving to the optimum region, where  $k$  is the number of factors. The functions  $f_1$  to  $f_5$  consist of the stationary points  $x_1, x_2$  i.e. two factors, and so the possible starting point for the proposed method could be one of four possible cases, as is shown in Figure 4.7. The resultant algorithms for the four cases are referred to as NM(1), NM(2), NM(3), and NM(4).



**Figure 4.7** Four Possible Starting Points in the First Phase of the Proposed Method

**Step 7** The NM algorithm was applied in the proposed method (either NM(1), NM(2), NM(3), or NM(4)) and utilized to optimize the response of interest instead of the steepest ascent/descent method. The functions  $f_1$  to  $f_5$  were applied using the proposed method. Following this, each function obtained one hundred CIs with 95% for the mean of the stationary points  $x_1, x_2$  and  $y$  were obtained for each function for each of the four cases dependent on NM(1), NM(2), NM(3) and NM(4), respectively.

**Step 8** To compare the efficiency of the classical RSM(CCD) with the proposed method, the coverage probability was computed as counts of the identification of the true value of each function contained in the 100 95% confidence intervals for the mean of the stationary points  $x_1, x_2$  and  $y$ .

**Step 9** MAPE was computed for both methods and used to compare their efficiencies:

$$\text{MAPE}(\hat{x}_{\min_1}) = \frac{1}{50} \sum_{j=1}^{50} \frac{\left| \hat{x}_{\min_1 j} - x_{\min_1} \right|}{x_{\min_1}} \times 100$$

$$\text{MAPE}(\hat{x}_{\min_2}) = \frac{1}{50} \sum_{j=1}^{50} \frac{\left| \hat{x}_{\min_2 j} - x_{\min_2} \right|}{x_{\min_2}} \times 100;$$

$$\text{MAPE}(\hat{y}_{\min}) = \frac{1}{50} \sum_{j=1}^{50} \frac{\left| \hat{y}_{\min j} - y_{\min} \right|}{y_{\min}} \times 100.$$

**Step 10** The average number of points from 100 replications was computed for each method.

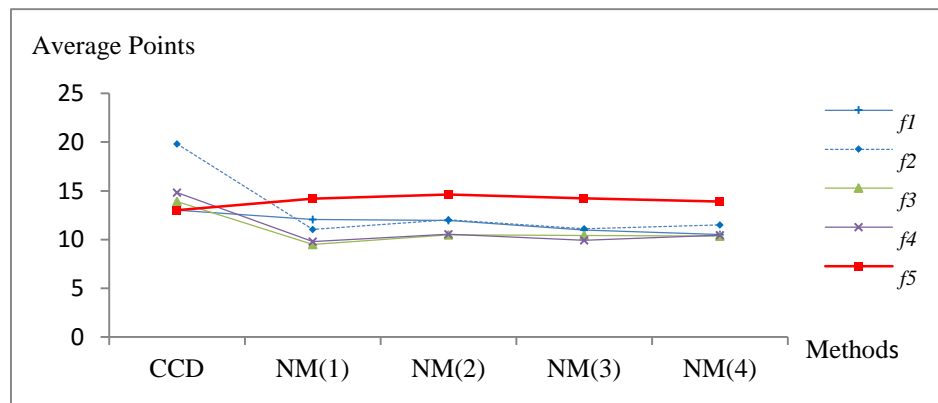
## 4.2 Results of the Simulation Study

Simulation study results were computed from 100 sets of 50 replications. The efficiencies of the RSM (CCD) and the proposed method forms NM(1), NM(2), NM(3) and NM(4) were compared in terms of the average number of points, coverage probability, and MAPE.

**Table 4.2** The Average Number of Required Experiments for RSM (CCD) and the Proposed Method

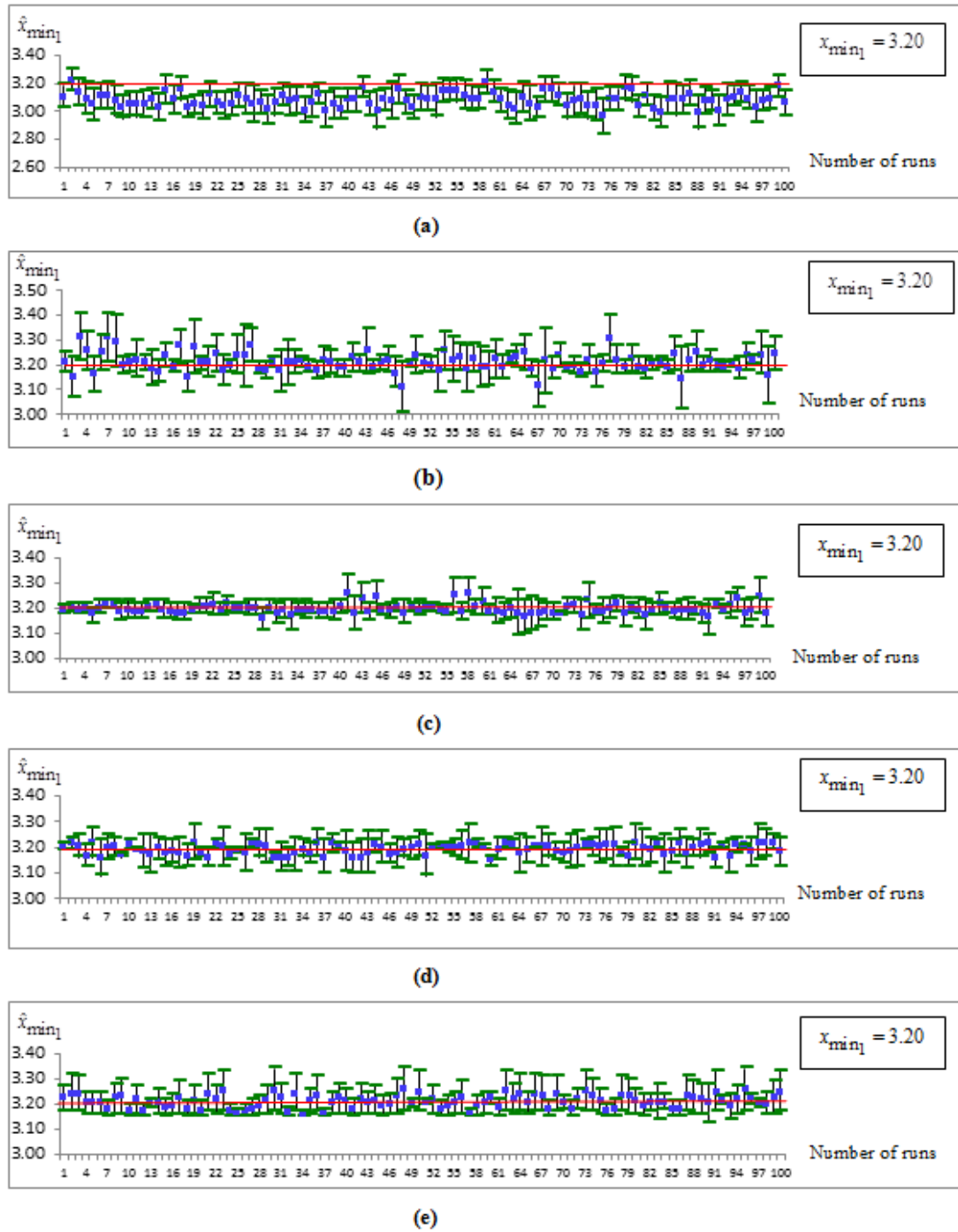
Test Function	Average Number of Required Experiments				
	CCD	NM(1)	NM(2)	NM(3)	NM(4)
$f_1$	13.00	12.06(92.77)	11.95(91.92)	10.97(84.38)	10.51(80.85)
$f_2$	19.83	11.04(55.67)	12.02(60.62)	11.12(56.08)	11.51(58.04)
$f_3$	13.89	9.48(68.25)	10.49(75.52)	10.41(74.95)	10.36(74.59)
$f_4$	14.83	9.79(66.01)	10.54(71.07)	9.91(66.82)	10.46(70.53)
$f_5$	13.00	14.19(109.15)	14.61(112.38)	14.23(109.46)	13.89(106.85)

**Note:** The Numbers in Parentheses are the Percentage Relative to CCD

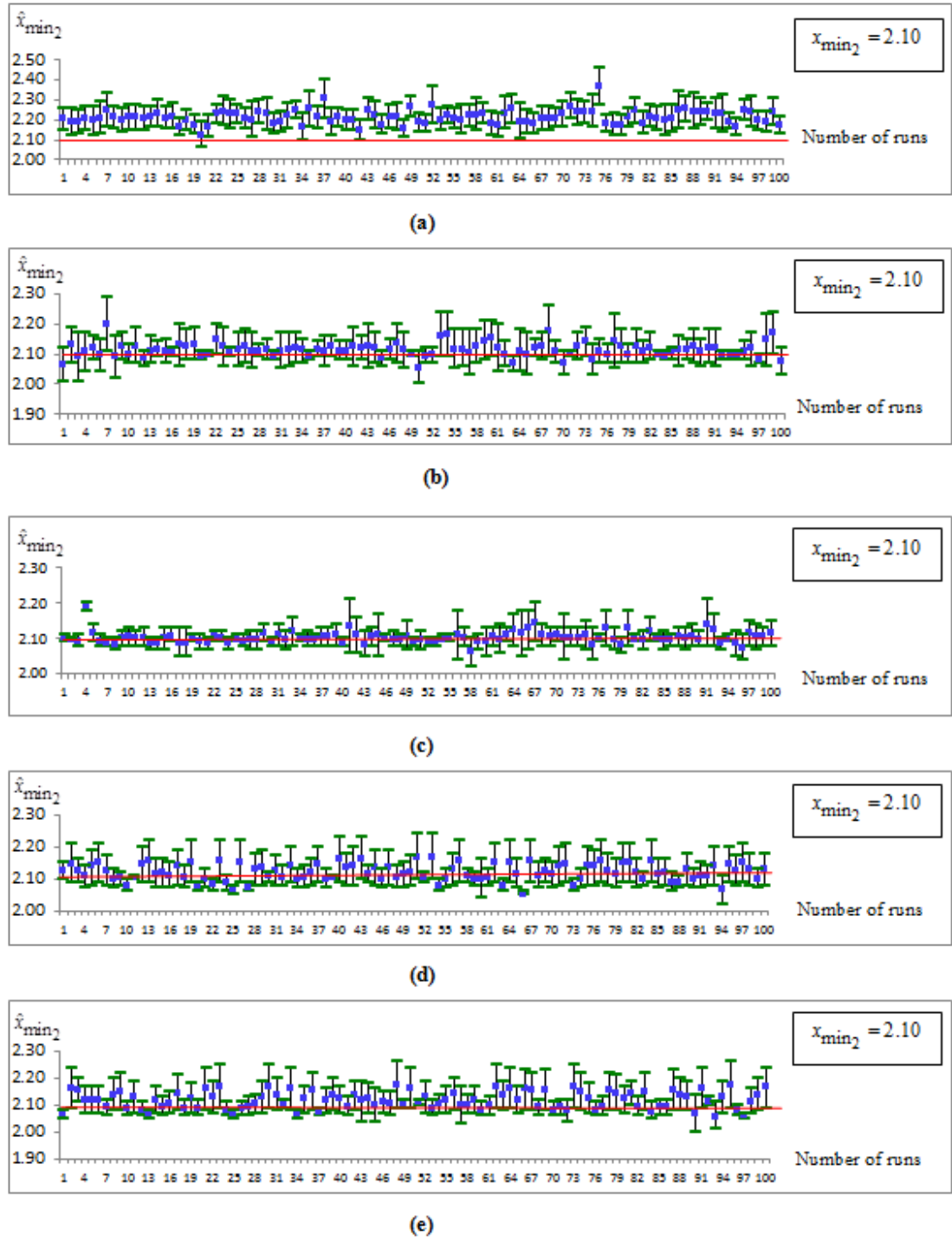


**Figure 4.8** Comparison of Average Points for CCD and the Proposed Method for Functions  $f_1$  to  $f_5$

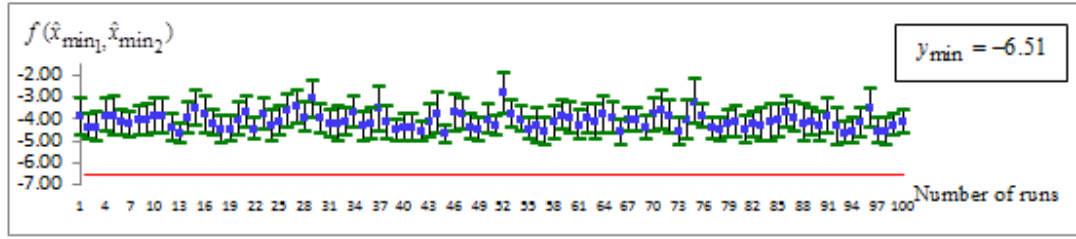
In Table 4.2 and Figure 4.8, we can see that the average points for classical RSM (CCD) are more than the proposed method for almost all of the functions. Except function  $f_5$ , both methods are similar in terms of average number of points.



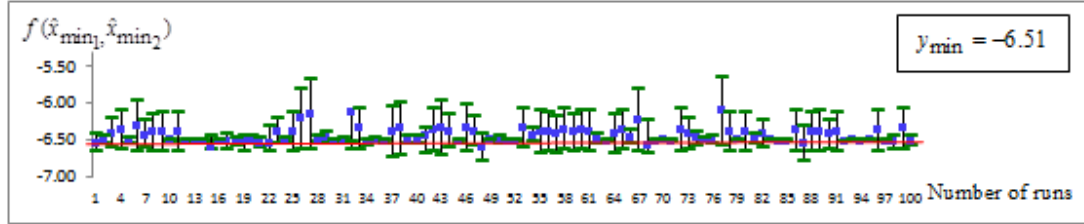
**Figure 4.9** One Hundred CIs of Average  $\hat{x}_{min1}$  of  $f_1$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



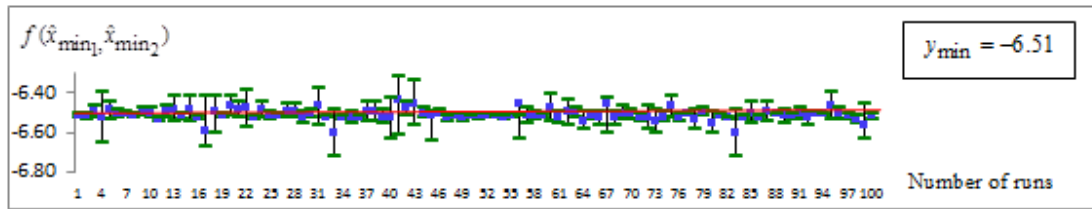
**Figure 4.10** One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_1$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



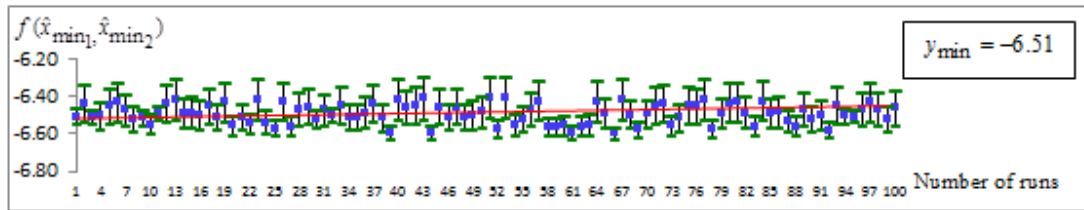
(a)



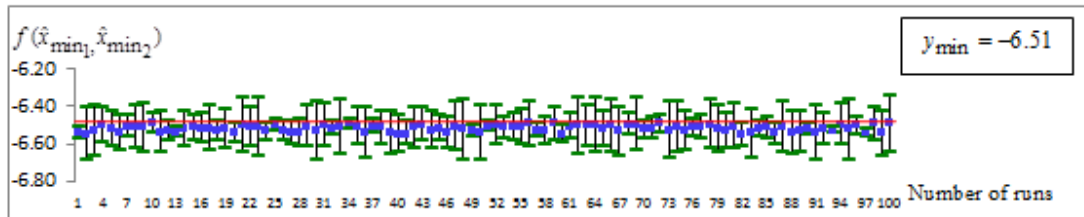
(b)



(c)



(d)



(e)

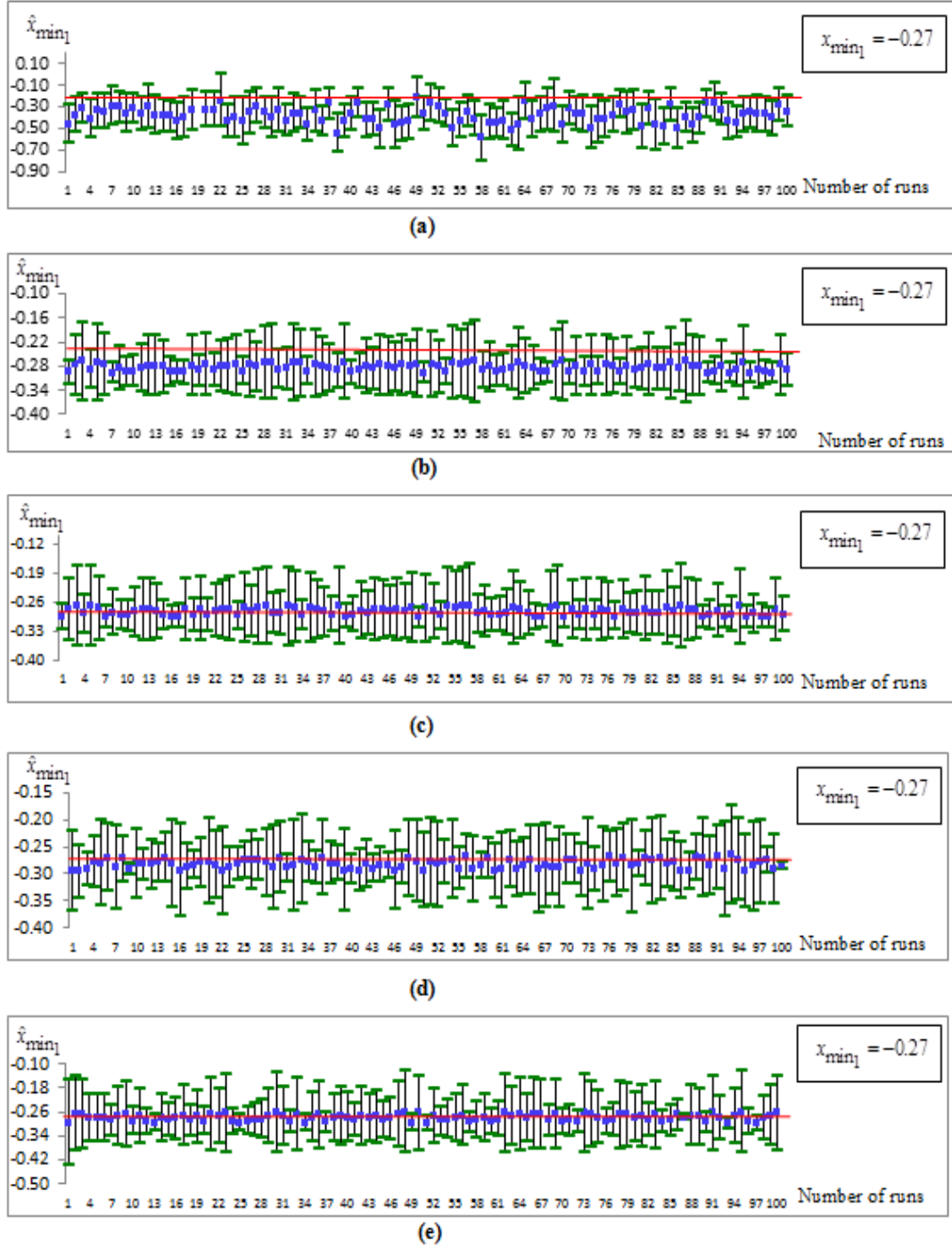
**Figure 4.11** One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_1$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



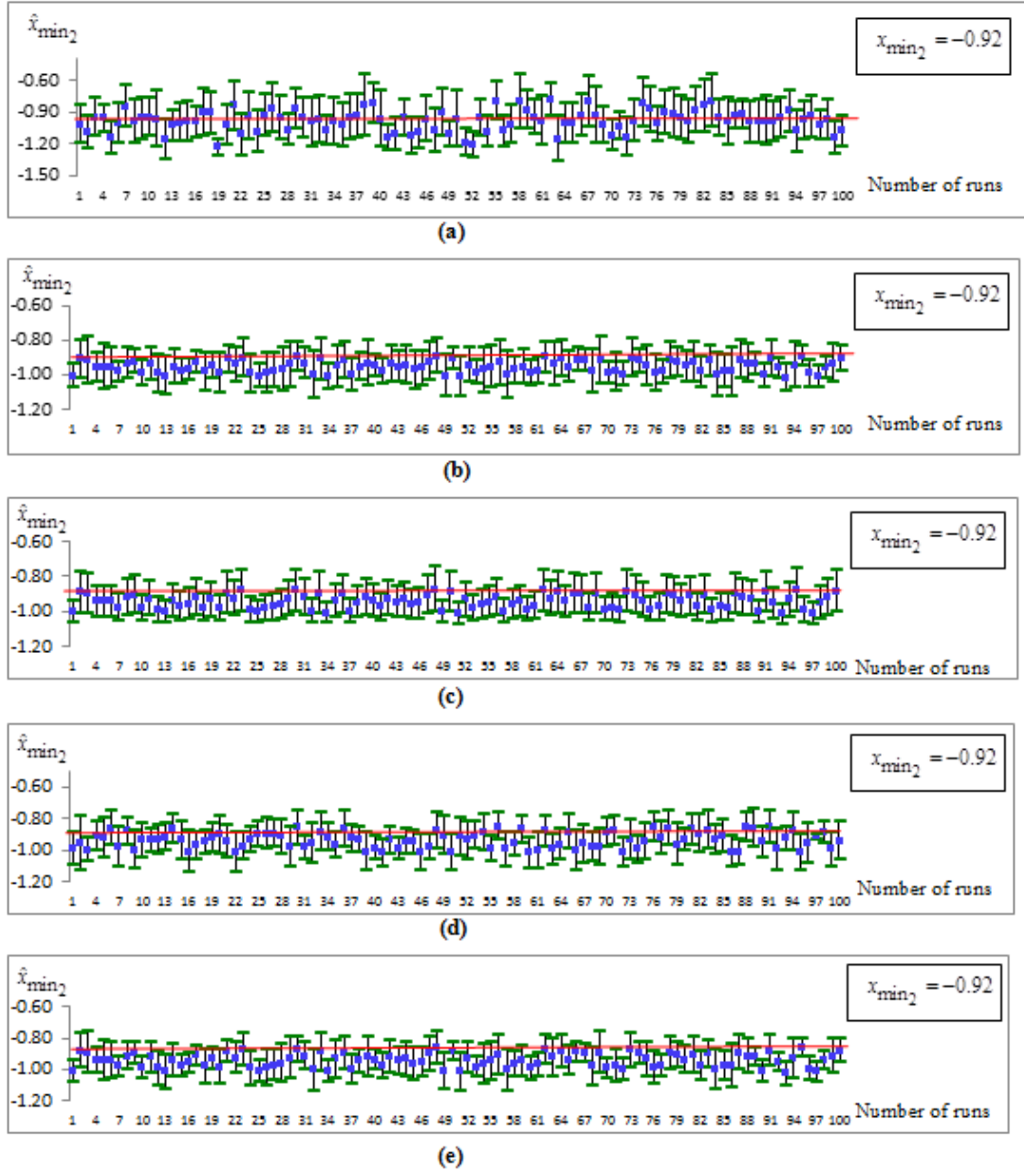
**Table 4.3** Comparison of the Coverage Probability of RSM(CCD) to the Proposed Method (NM(1),NM(2),NM(3),NM(4)) for  $f_1$

Test function	Method	Coverage probability of			
		$x_1$	$x_2$	Both $x_1, x_2$	Response( $y$ )
$f_1$	CCD	0.27	0.01	0.00	0.00
	NM(1)	0.96	0.98	0.95	0.93
	NM(2)	0.99	0.99	0.98	0.85
	NM(3)	0.98	0.96	0.94	0.84
	NM(4)	0.88	0.93	0.87	0.98

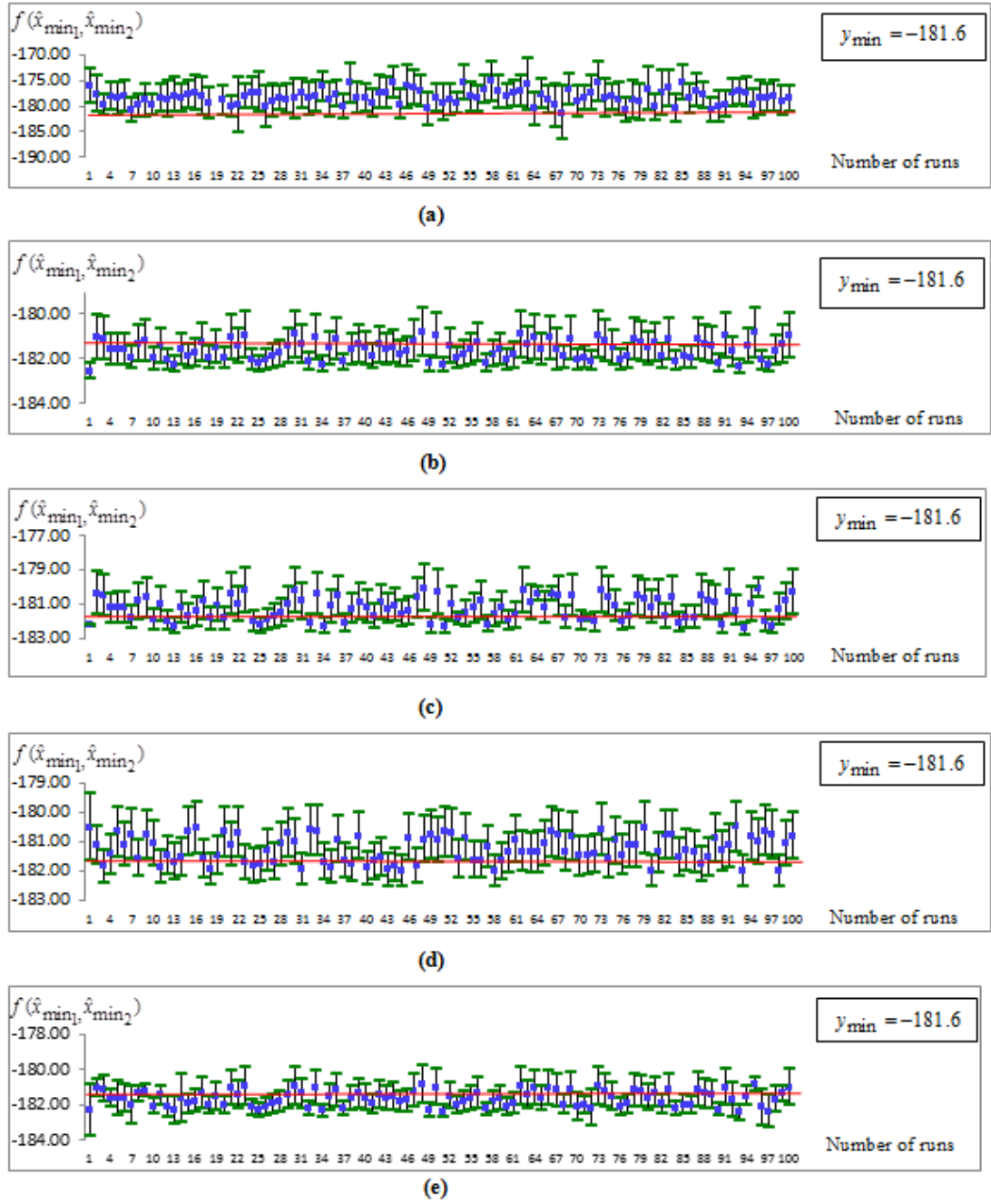
A summary of the results of coverage probability for function  $f_1$  from Figures 4.9 - 4.11 are contained in Table 4.3. We can see that the proposed method consisting of forms NM(1), NM(2), NM(3), and NM(4) performed better than the classical RSM(CCD), especially in the case of both true values ( $x_1, x_2$ ) being contained in the 95% CI of their means ( $\hat{x}_1, \hat{x}_2$ ). The worst case of coverage probability of 0.87 was shown by proposed method form NM(4). Even though this was the worst case for the proposed method, it still outperformed the classical RSM (CCD). Note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4)) are not much different from each other.



**Figure 4.12** One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_2$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



**Figure 4.13** One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_2$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)

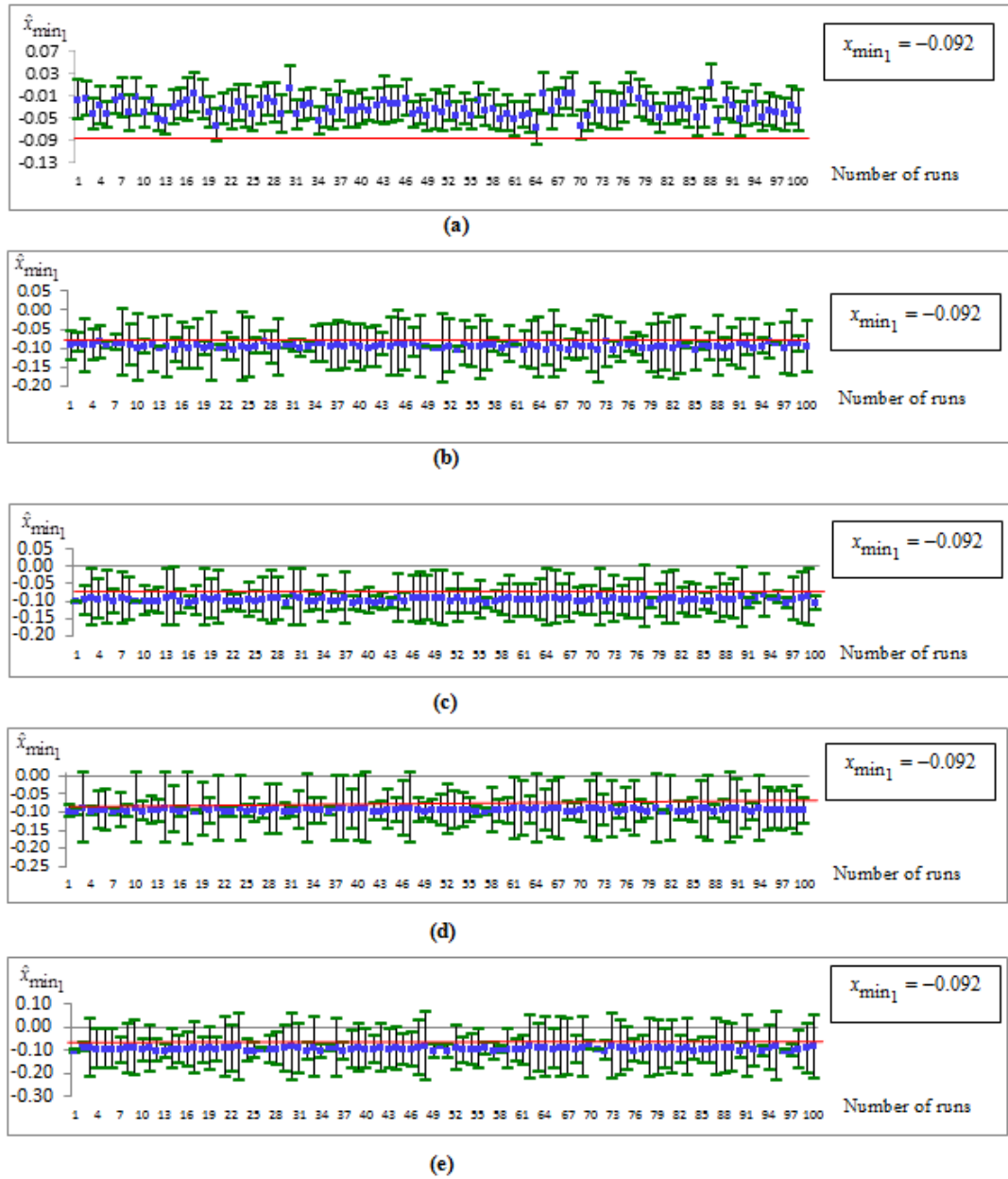


**Figure 4.14** One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_2$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)

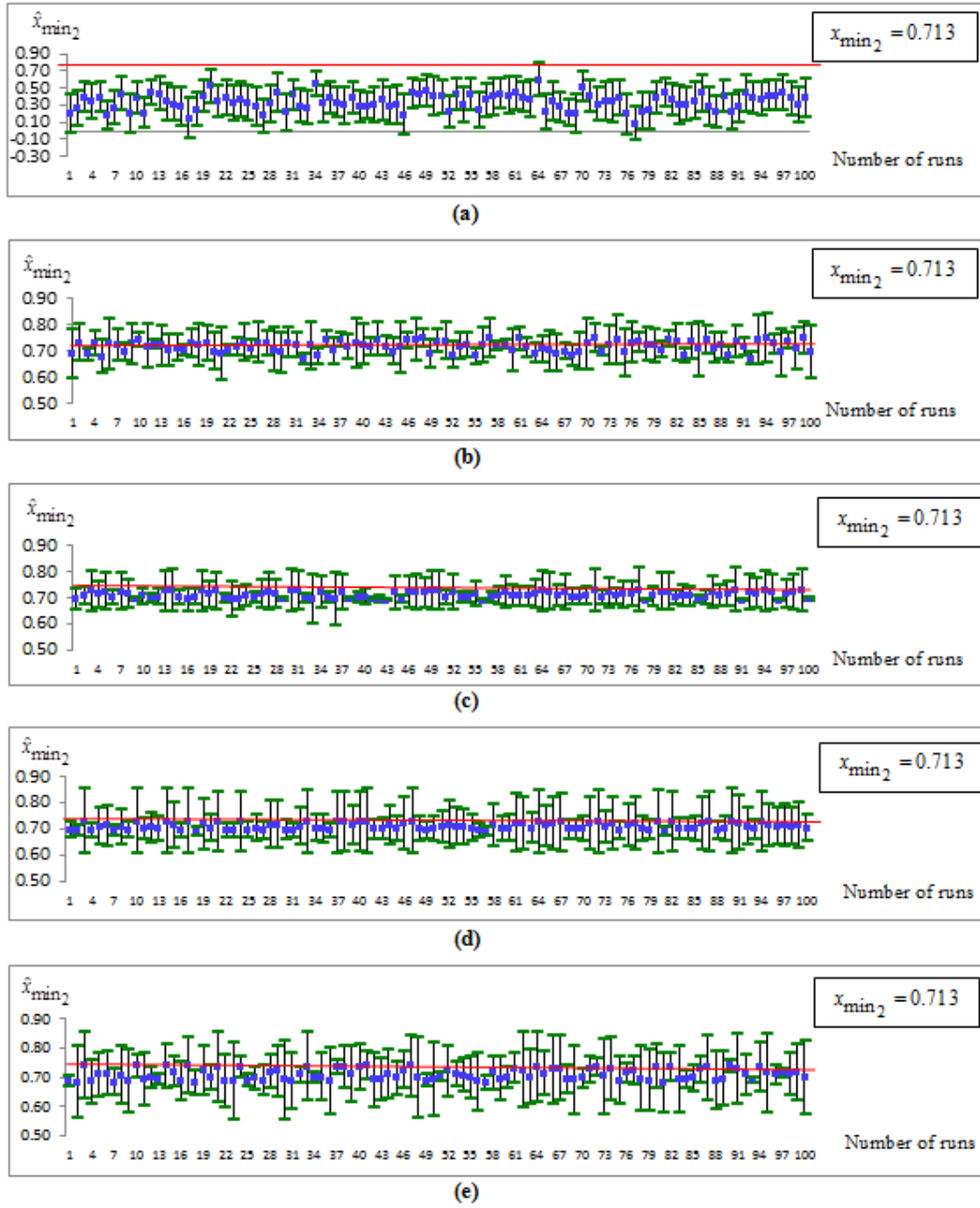
**Table 4.4** Comparison of the Coverage Probability of the Classical RSM(CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_2$

Test function	Method	Coverage probability of			
		$x_1$	$x_2$	Both $x_1, x_2$	Response( $y$ )
$f_2$	CCD	0.60	0.65	0.28	0.16
	NM(1)	0.96	0.91	0.89	0.84
	NM(2)	0.89	0.86	0.84	0.78
	NM(3)	0.98	0.96	0.94	0.80
	NM(4)	0.95	0.91	0.89	0.82

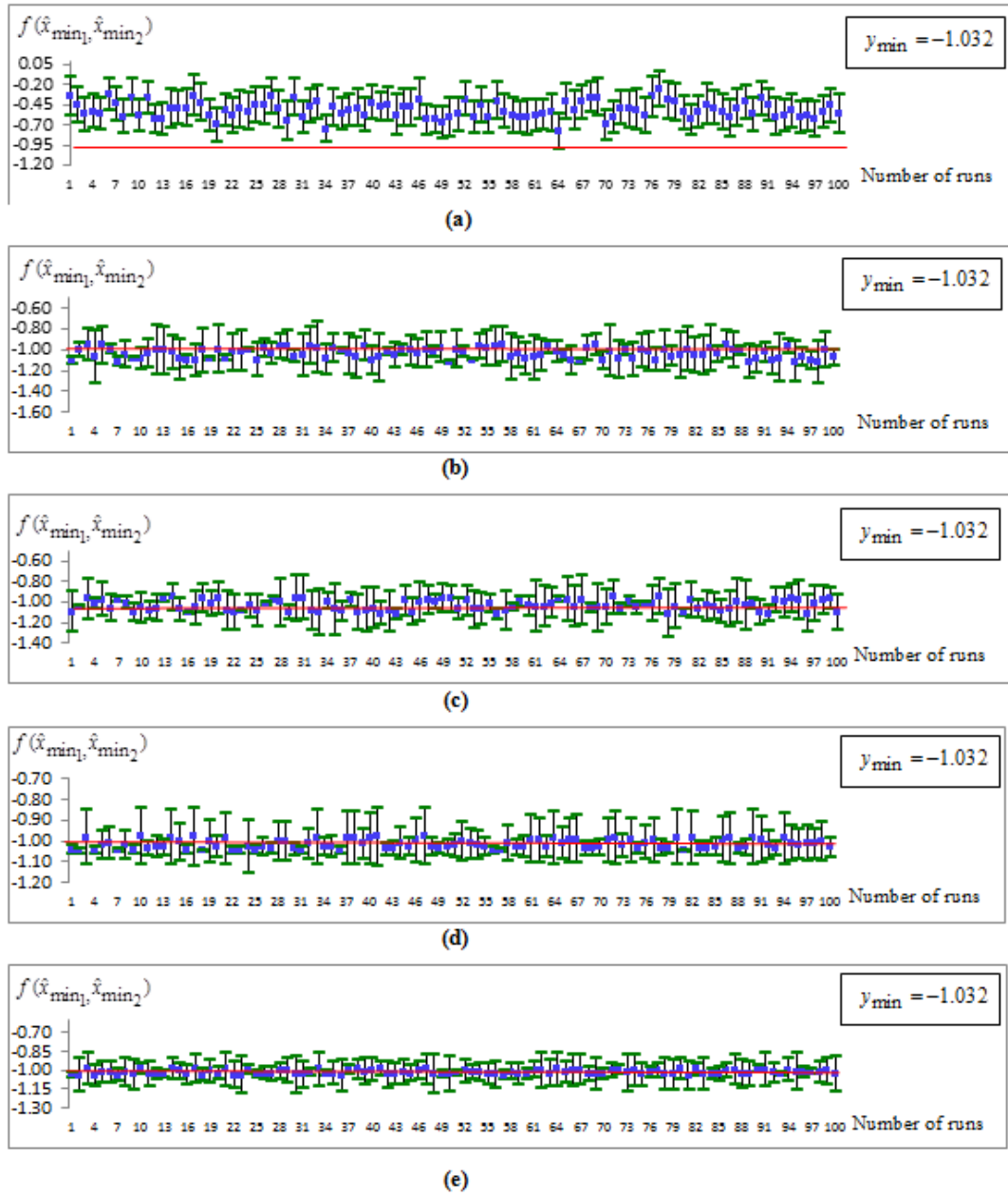
The results of coverage probability for function  $f_2$  from Figures 4.12 - 4.14 are summarized in Table 4.4. Once again, it is evident that the proposed method forms NM(1), NM(2), NM(3), and NM(4) showed superior performance to the classical RSM(CCD), especially when considering the true values  $(x_1, x_2)$  together at the 95% CIs of their means  $(\hat{x}_1, \hat{x}_2)$ . The worst case of coverage probability (0.84) of the proposed method was shown by form NM(2). Even though this was the worst case scenario for the proposed method, it still had far superior performance than shown by the classical RSM(CCD). Once more, note that there are no major differences among the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4).



**Figure 4.15** One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_3$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



**Figure 4.16** One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_3$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



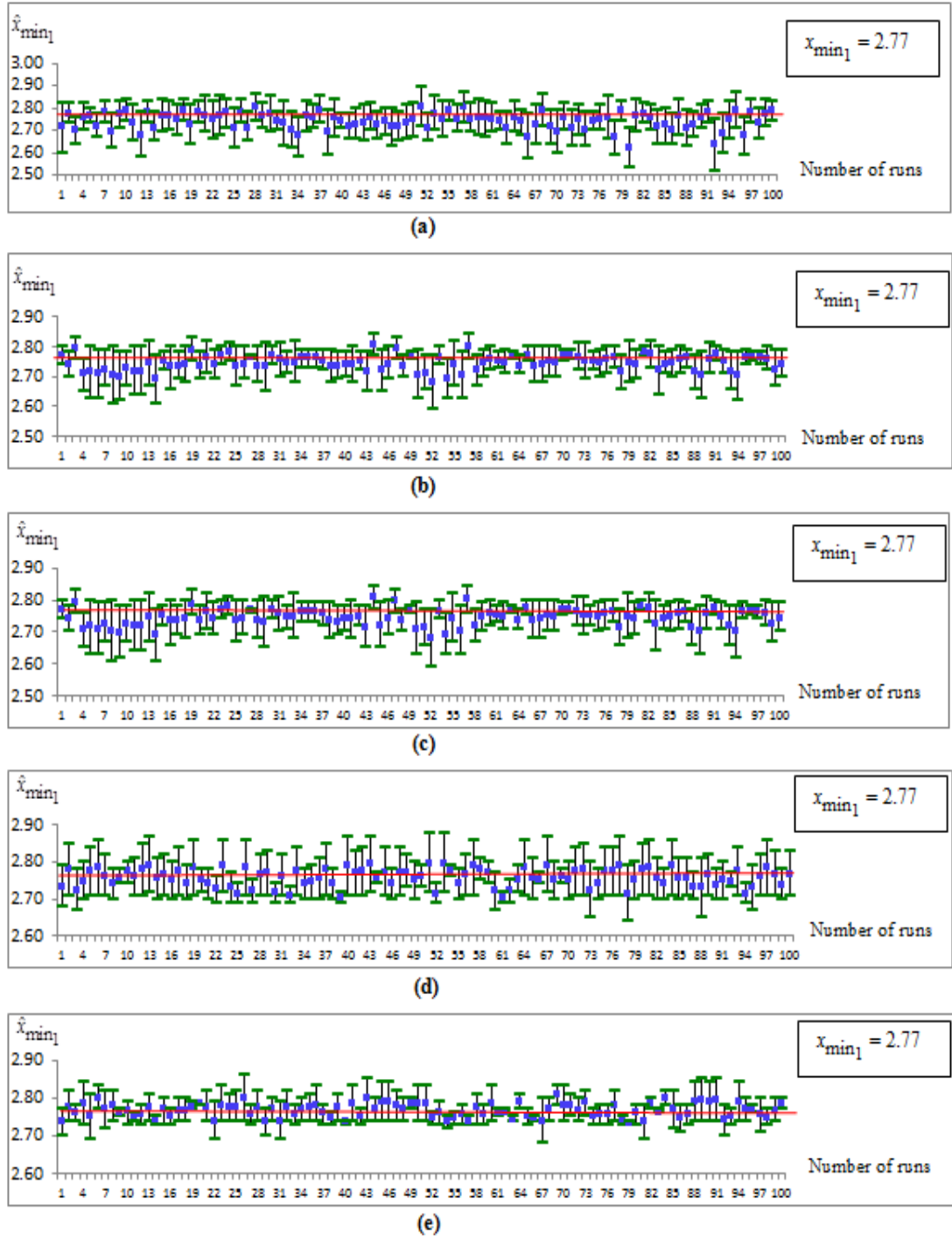
**Figure 4.17** One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_3$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



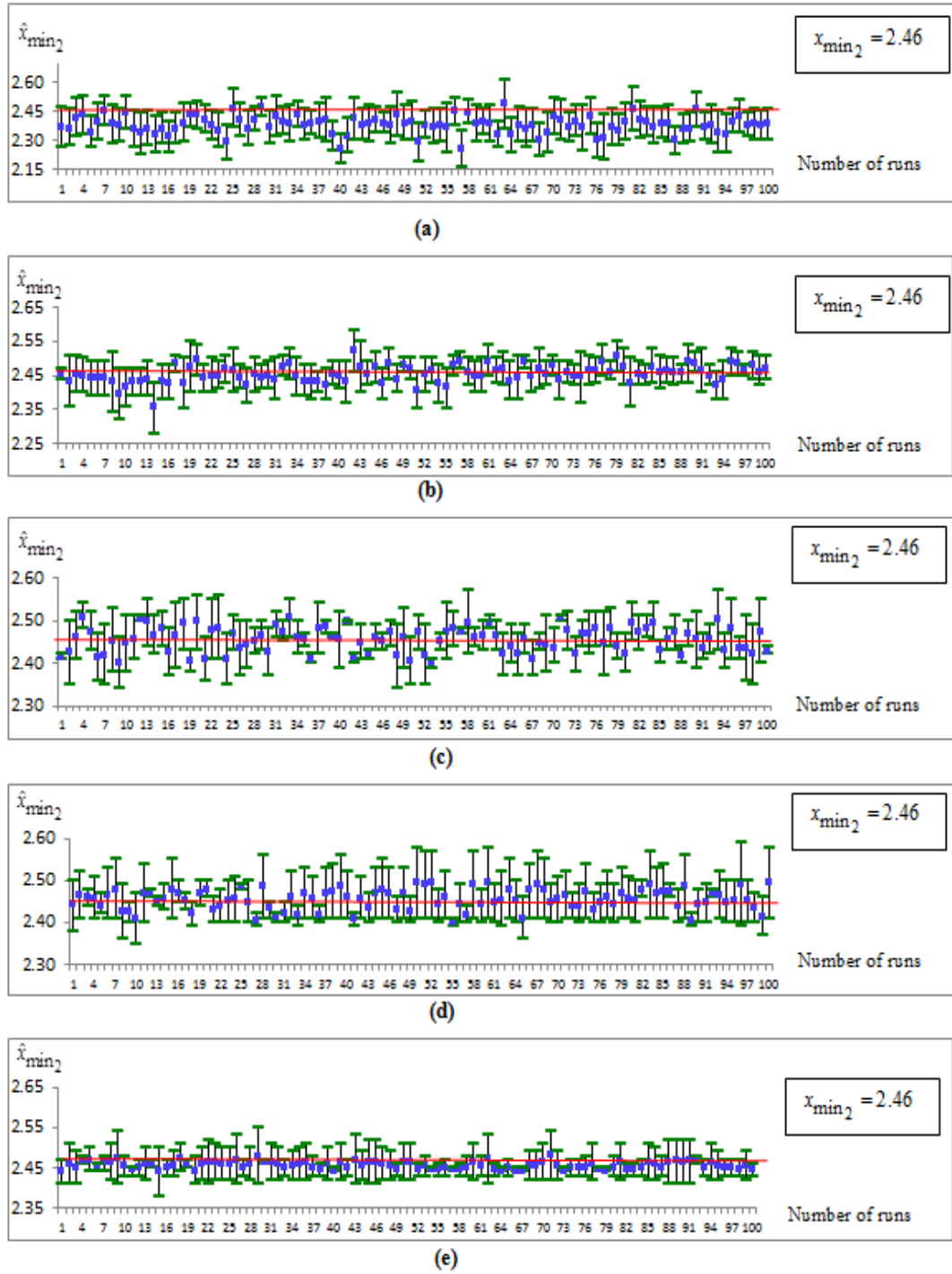
**Table 4.5** Comparison of the Coverage Probability of the Classical RSM(CCD) with the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_3$

Test function	Method	Coverage probability of			
		$x_1$	$x_2$	Both $x_1, x_2$	Response( $y$ )
$f_3$	CCD	0.08	0.08	0.04	0.01
	NM(1)	0.94	0.89	0.83	0.89
	NM(2)	0.97	0.87	0.85	0.88
	NM(3)	0.97	0.95	0.95	0.95
	NM(4)	0.96	0.94	0.91	0.97

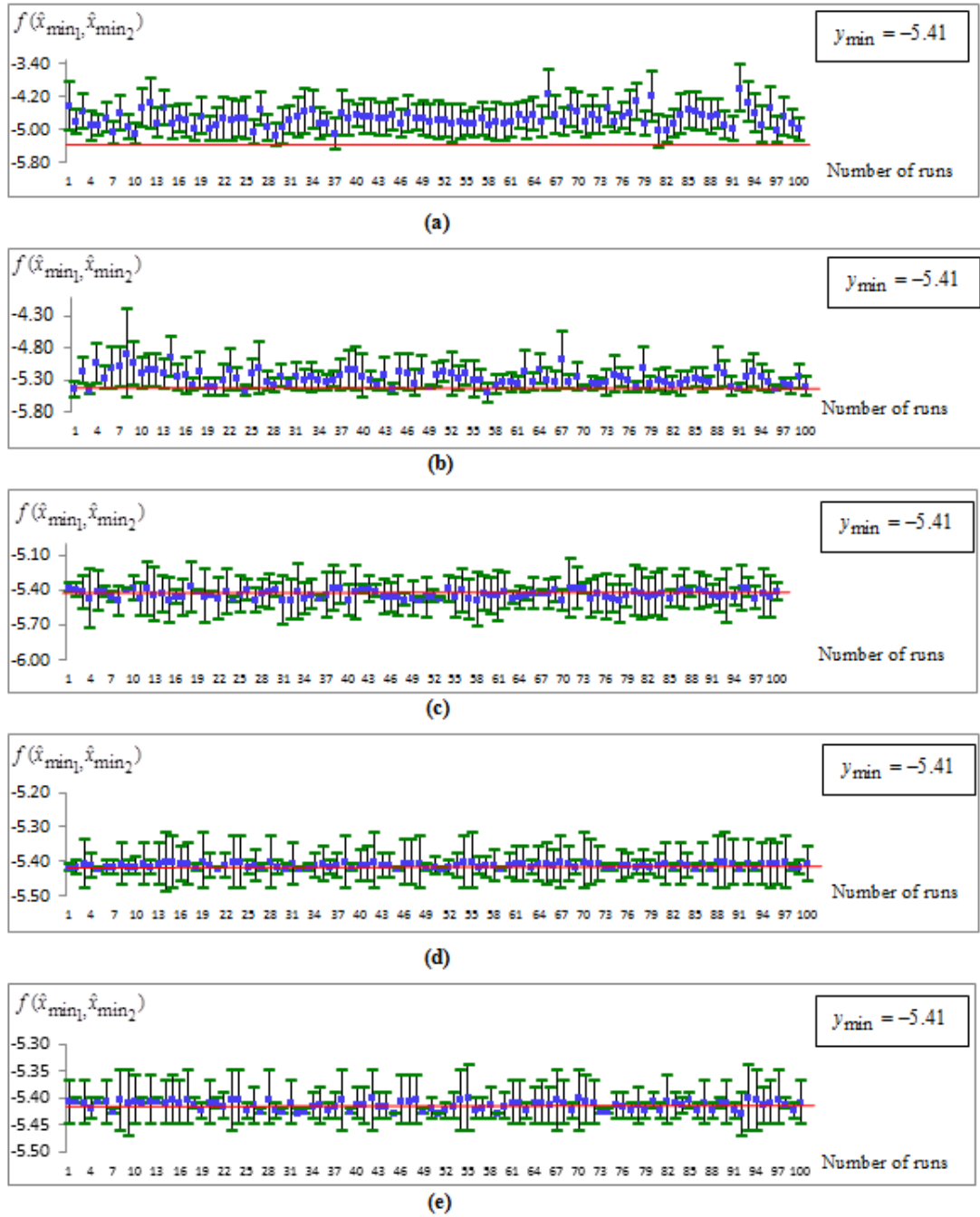
The results in Table 4.5 comprise a summary of the results of coverage probability from Figures 4.15 - 4.17 for function  $f_3$ , which is not quadratic. Once again, we can see that the proposed method forms NM(1), NM(2), NM(3), and NM(4) performed much better than the classical RSM(CCD), especially in the case of both true values ( $x_1, x_2$ ) being contained in the 95% CIs of their means ( $\hat{x}_1, \hat{x}_2$ ). The worst case of coverage probability (0.83) was shown by the proposed method form NM(1). However, all cases of the proposed method showed performances far superior to the classical RSM (CCD), although, once again note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) showed no major differences.



**Figure 4.18** One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_4$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



**Figure 4.19** One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_4$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)

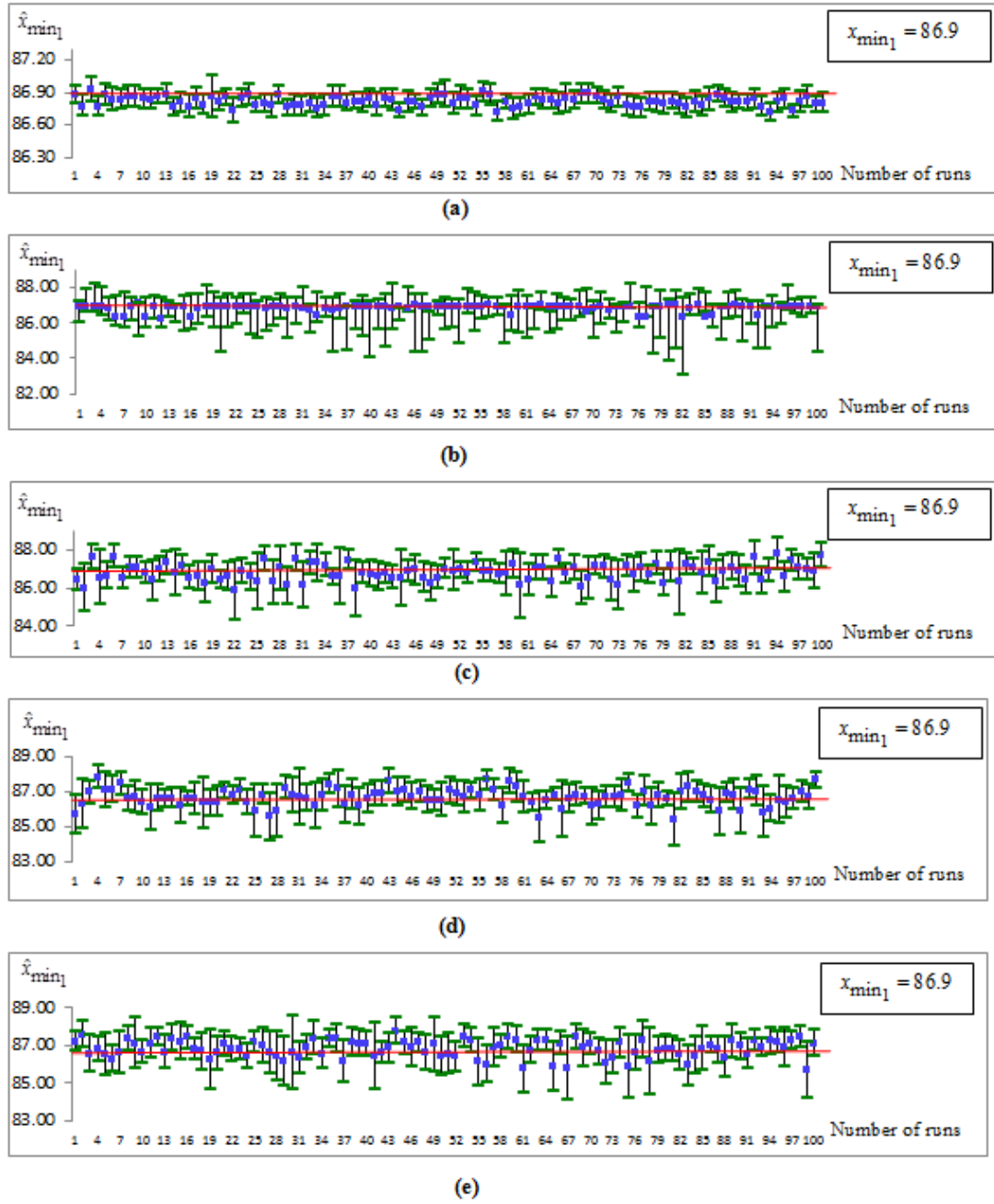


**Figure 4.20** One Hundred CIs of Average  $\hat{y}_{\min}$  of  $f_4$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)

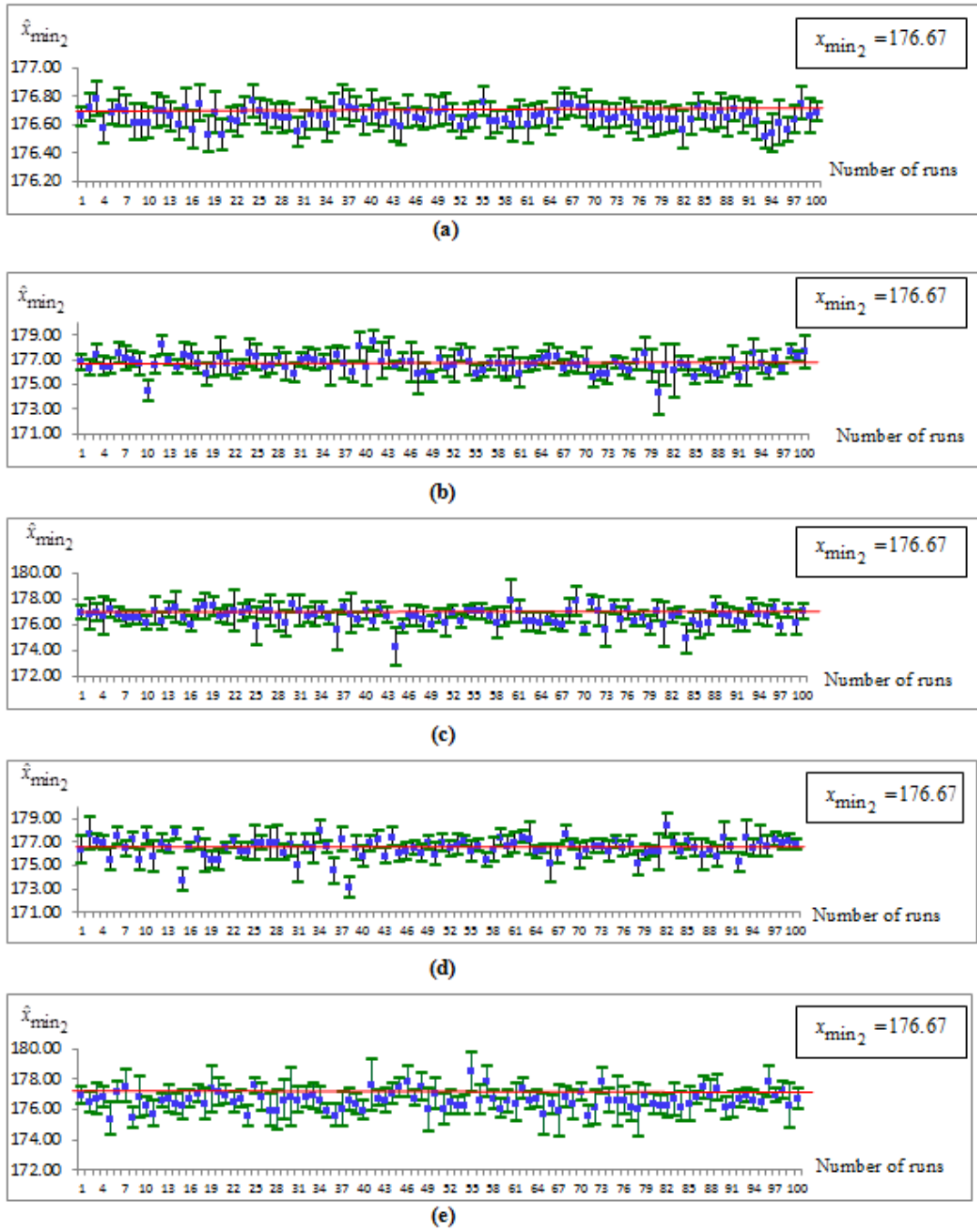
**Table 4.6** Comparison of the Coverage Probability of the Classical RSM(CCD) and the Proposed Method forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_4$

Test function	Method	Coverage probability of			
		$x_1$	$x_2$	Both $x_1, x_2$	Response( $y$ )
$f_4$	CCD	0.88	0.46	0.37	0.04
	NM(1)	0.99	0.84	0.83	0.93
	NM(2)	0.82	0.79	0.73	0.90
	NM(3)	0.88	0.83	0.74	0.88
	NM(4)	0.76	0.84	0.72	0.82

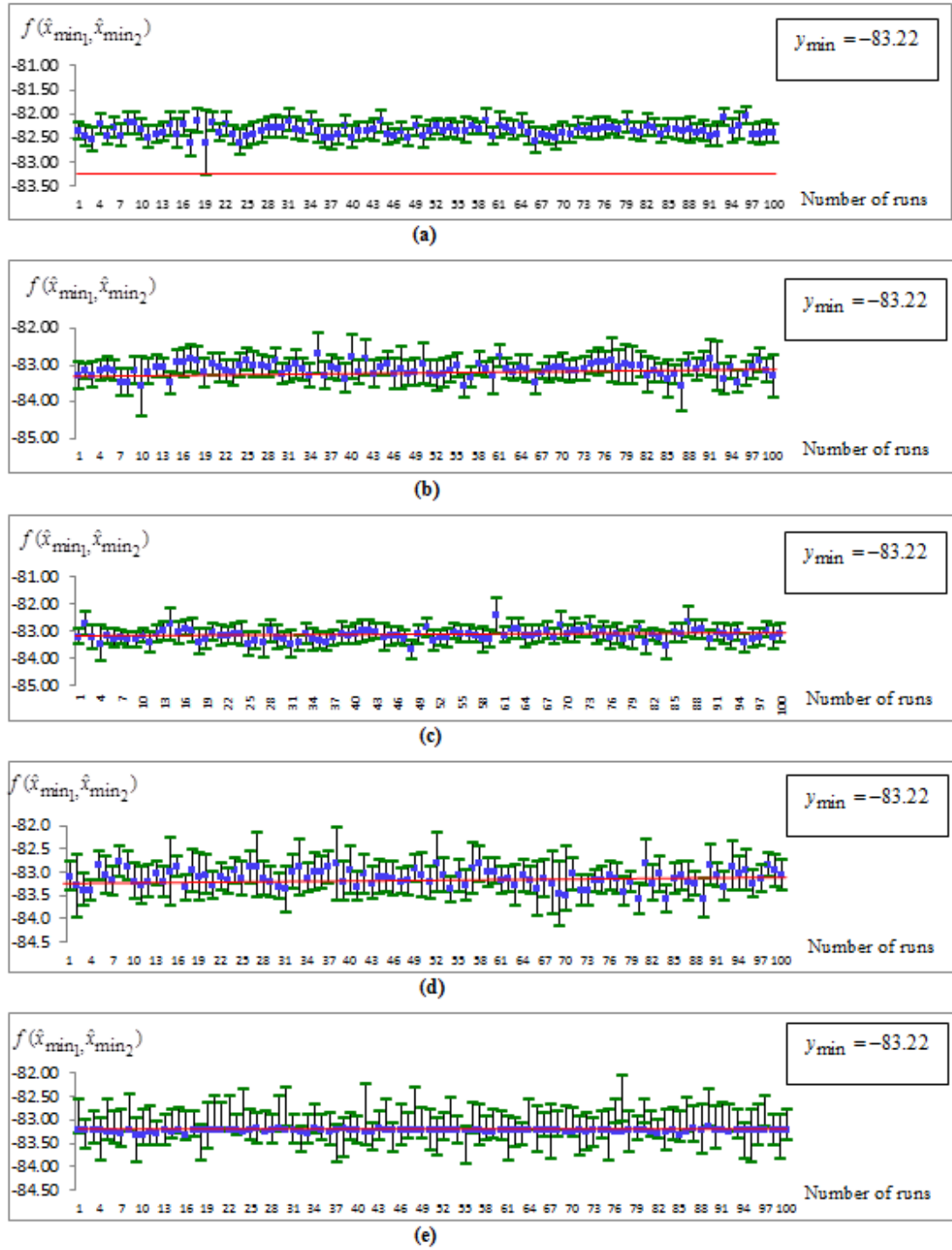
Results of coverage probability from Figures 4.18 - 4.20 for function  $f_4$  are summarized in Table 4.6. This time, the coverage probability of the classical RSM(CCD) and the proposed method showed similar results. However, in the case of both true values ( $x_1, x_2$ ) being contained in the 95% CI of their means ( $\hat{x}_1, \hat{x}_2$ ), the proposed method was markedly better for all forms. The coverage probability worst case of 0.72 for the proposed method was shown by form NM(4), although this was better than the classical RSM(CCD). Once again, note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) are not much different from each other.



**Figure 4.21** One Hundred CIs of Average  $\hat{x}_{\min_1}$  of  $f_5$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



**Figure 4.22** One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_5$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



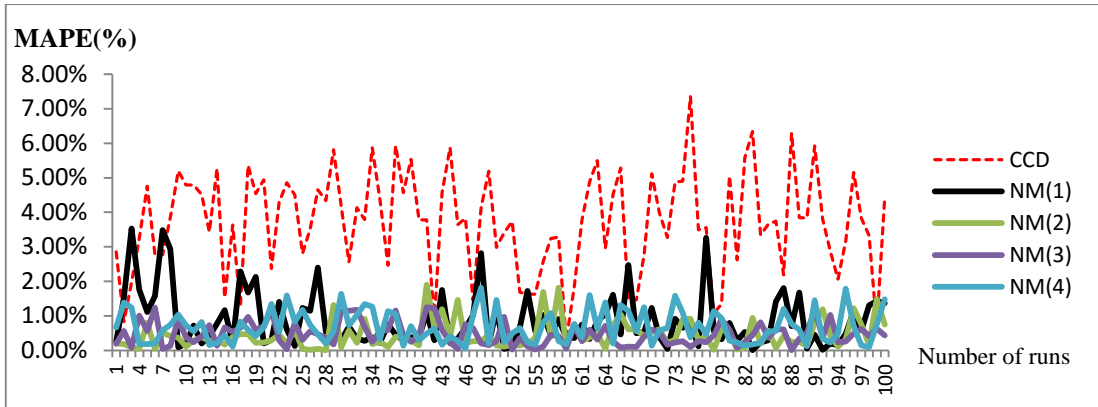
**Figure 4.23** One Hundred CIs of Average  $\hat{x}_{\min_2}$  of  $f_5$  Using (CCD) and Propose Method where (a) CCD, (b) NM(1), (c) NM(2), (d) NM(3) and (e) NM(4)



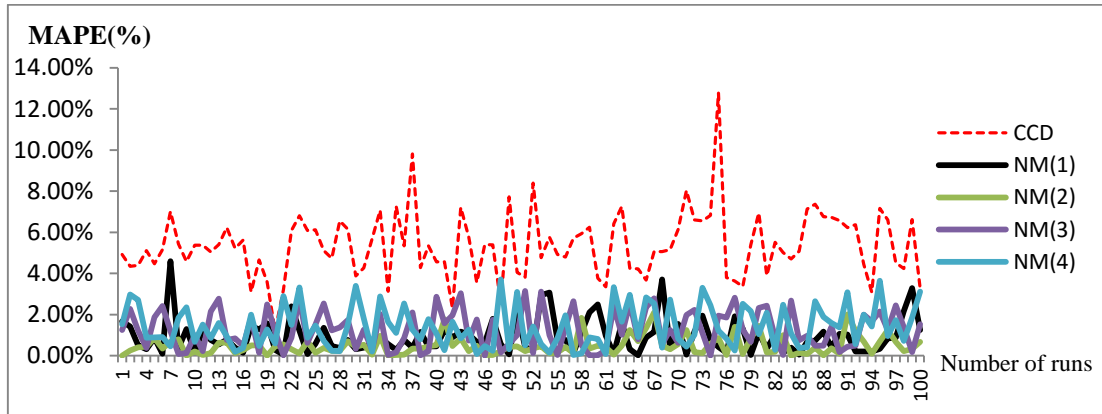
**Table 4.7** Comparison of the Coverage Probabilities of the Classical RSM(CCD) and the Proposed Method Forms NM(1), NM(2), NM(3),and NM(4) for Function  $f_5$

Test function	Method	Coverage probability of			
		$x_1$	$x_2$	Both $x_1, x_2$	Response( $y$ )
$f_5$	CCD	0.55	0.94	0.53	0.01
	NM(1)	0.95	0.83	0.78	0.94
	NM(2)	0.93	0.84	0.78	0.95
	NM(3)	0.92	0.77	0.69	0.91
	NM(4)	0.97	0.81	0.79	0.89

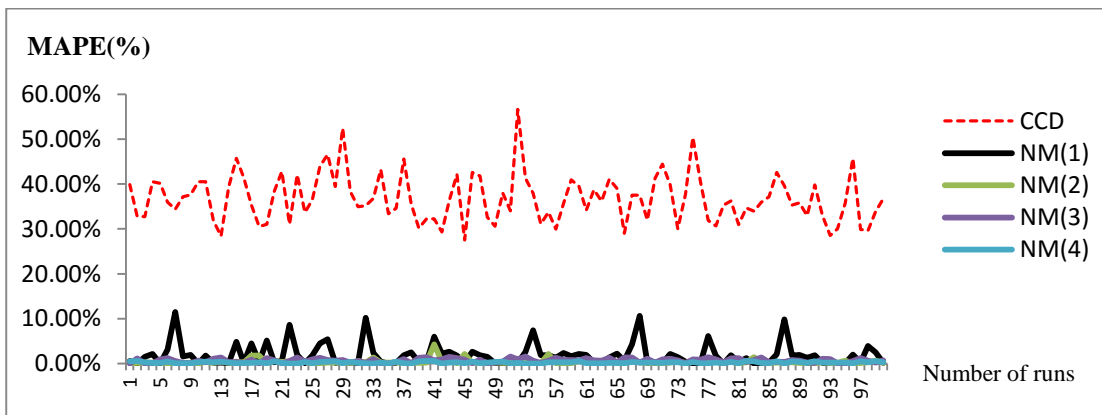
Table 4.7 is a summary of coverage probability results from Figures 4.20 - 4.23 for function  $f_5$ , which is quadratic. This time, the classical RSM(CCD) appeared to be useful for fitting quadratic models, and comparison of its coverage probability with the proposed method showed that they both performed similarly. However, when considering the case of true values  $(x_1, x_2)$ , the proposed method was superior. The worst case shown by the proposed method was NM(3) at 0.69, which was still better than the classical RSM(CCD). As with the other scenarios, note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) do not vary by much.



**Figure 4.24** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_1$  of  $f_1$



**Figure 4.25** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_2$  of  $f_1$

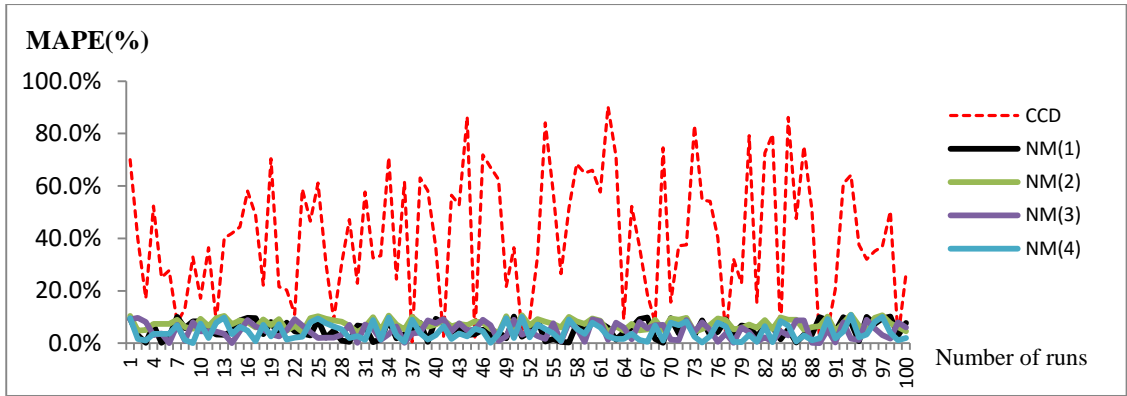


**Figure 4.26** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $y$  of  $f_1$

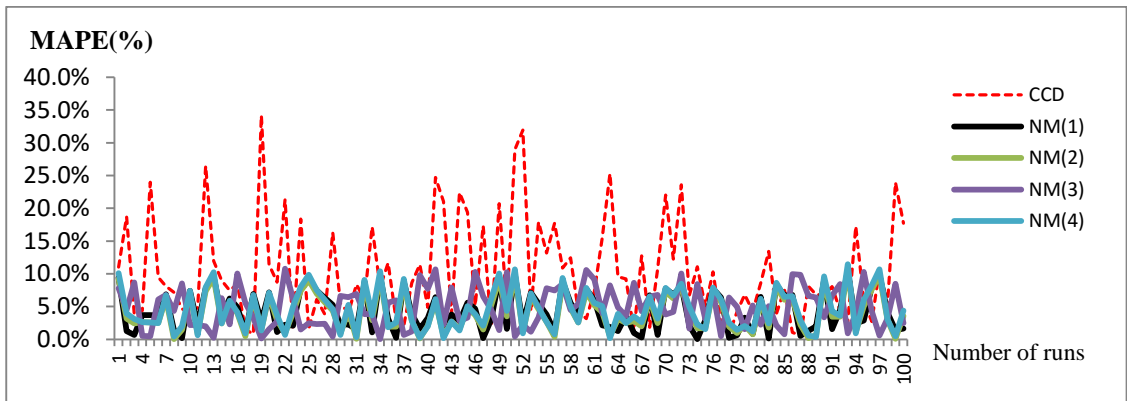
**Table 4.8** Comparison of the MAPE of the classical RSM(CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_1$

Compare with	Average MAPE (%)				
	Method				
	RSM	NM(1)	NM(2)	NM(3)	NM(4)
True $x_1$	3.6965	0.8659	0.4514	0.4805	0.6866
True $x_2$	5.3610	0.9359	0.4876	1.2673	1.4230
True $y$	36.8224	1.8790	0.3824	0.6934	0.2392

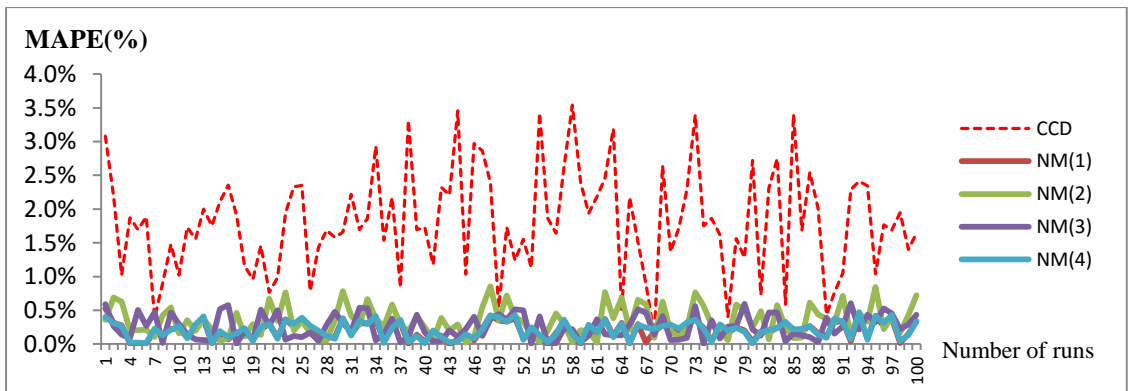
Results for MAPE for function  $f_1$  from Figures 4.24 - 4.26 are summarized in Table 4.8. It is evident that the proposed method forms NM(1), NM(2), NM(3), and NM(4) performed better than the classical RSM(CCD). Even the worst case scenario of the proposed method for form NM(1) showed a better performance than the classical RSM(CCD). Note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) were not much different from each other.



**Figure 4.27** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_1$  of  $f_2$



**Figure 4.28** Comparison of the MAPE of RSM(CCD) to the Proposed Method for the True  $x_2$  of  $f_2$

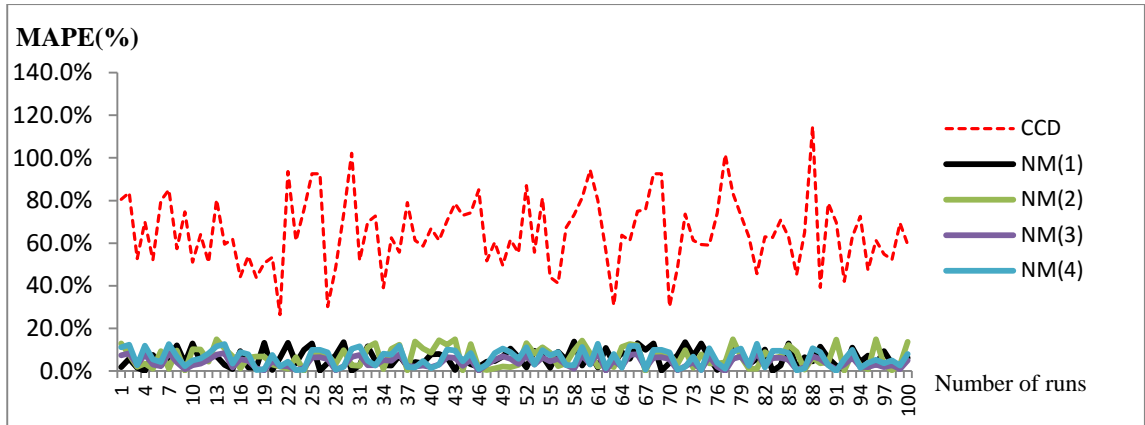


**Figure 4.29** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $y$  of  $f_2$

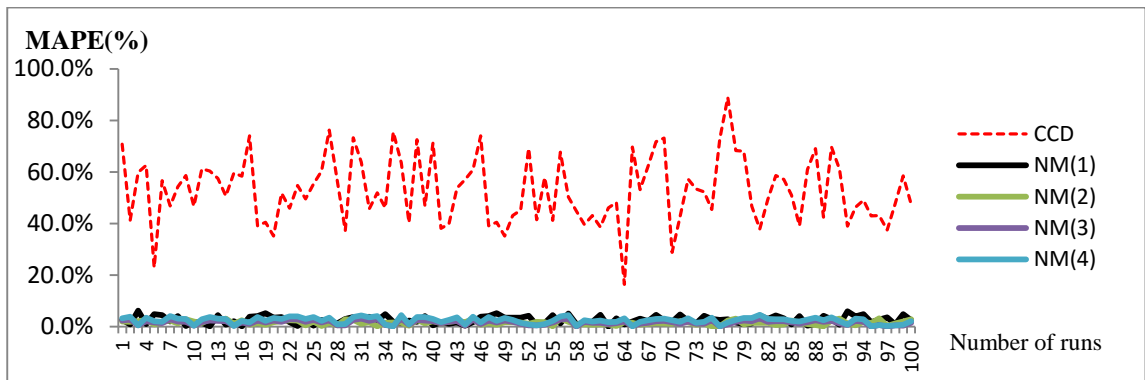
**Table 4.9** Comparison of the MAPE of the Classical RSM (CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for function  $f_2$

Compare with	Average MAPE (%)				
	Method				
	RSM	NM(1)	NM(2)	NM(3)	NM(4)
True $x_1$	40.7789	5.1883	7.3857	4.6431	4.3247
True $x_2$	10.2370	4.3123	4.1548	4.8997	4.5812
True $y$	1.8054	0.1975	0.3449	0.2556	0.2108

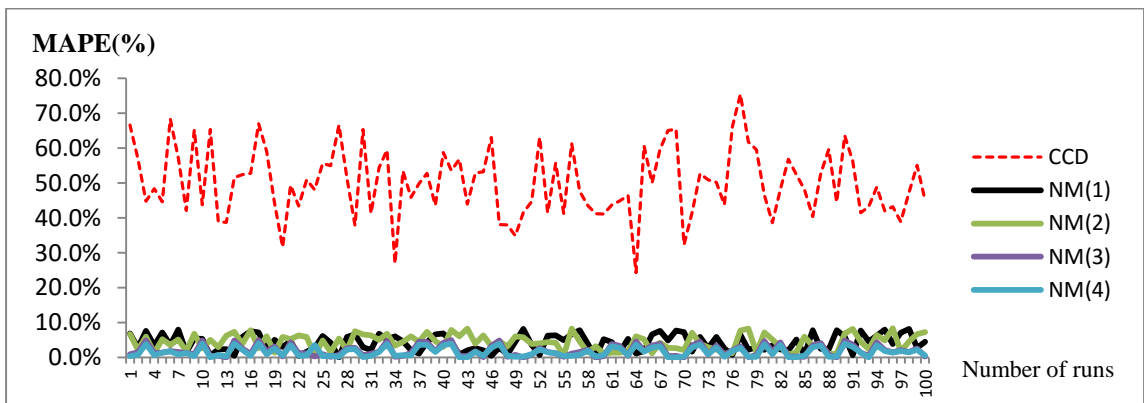
Table 4.9 shows the results of MAPE for function  $f_2$  from Figures 4.27 - 4.29. Furthermore, we can see that the proposed method forms NM(1), NM(2), NM(3), and NM(4) were superior to the classical RSM(CCD) by an order of magnitude. Note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) were pretty similar.



**Figure 4.30** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_1$  of  $f_3$



**Figure 4.31** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_2$  of  $f_3$



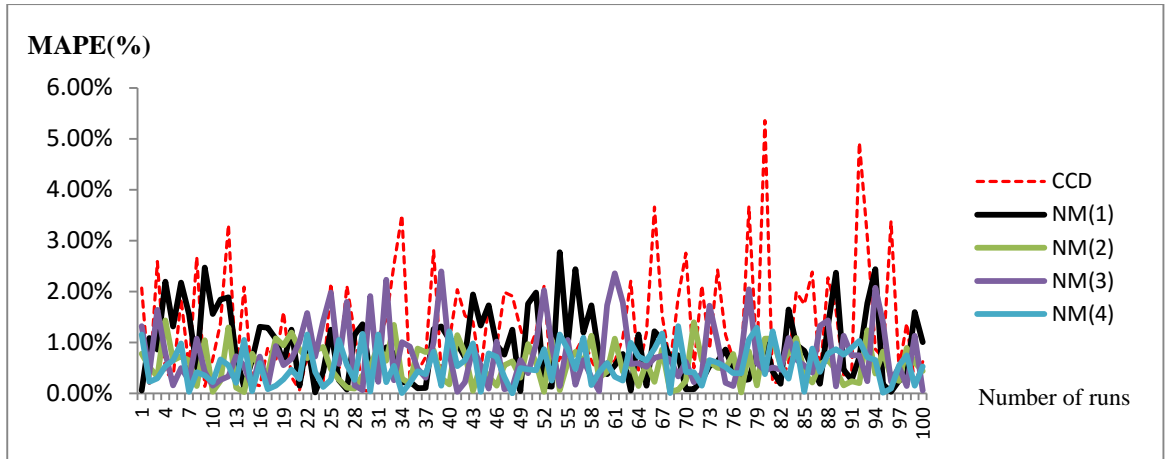
**Figure 4.32** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_3$  of  $f_3$

the True  $y$  of  $f_3$

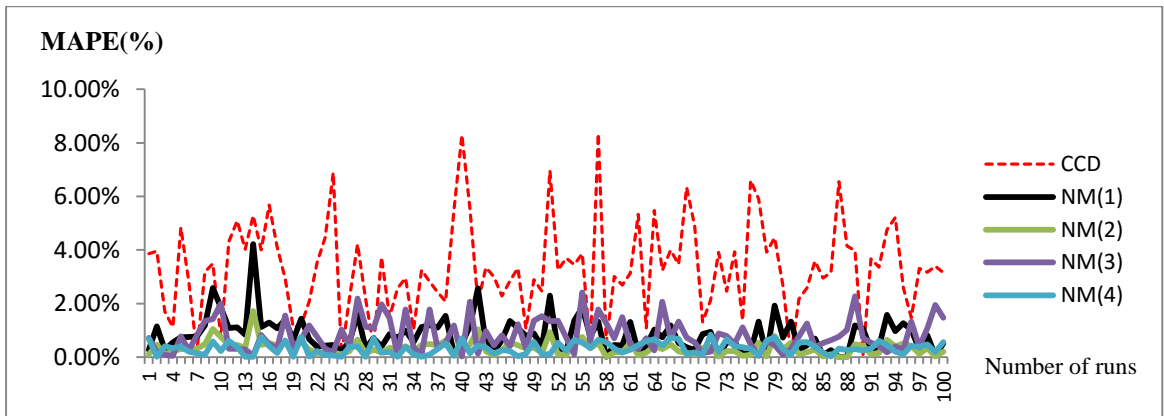
**Table 4.10** Comparison of the MAPE of the Classical RSM (CCD) and the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_3$

Compare with	Average MAPE (%)				
	Method				
	RSM	NM(1)	NM(2)	NM(3)	NM(4)
True $x_1$	64.9701	5.9455	6.5408	4.1388	6.2366
True $x_2$	53.1661	2.5116	1.5221	1.6426	2.4132
True $y$	50.1187	4.2093	4.4140	1.9570	1.6237

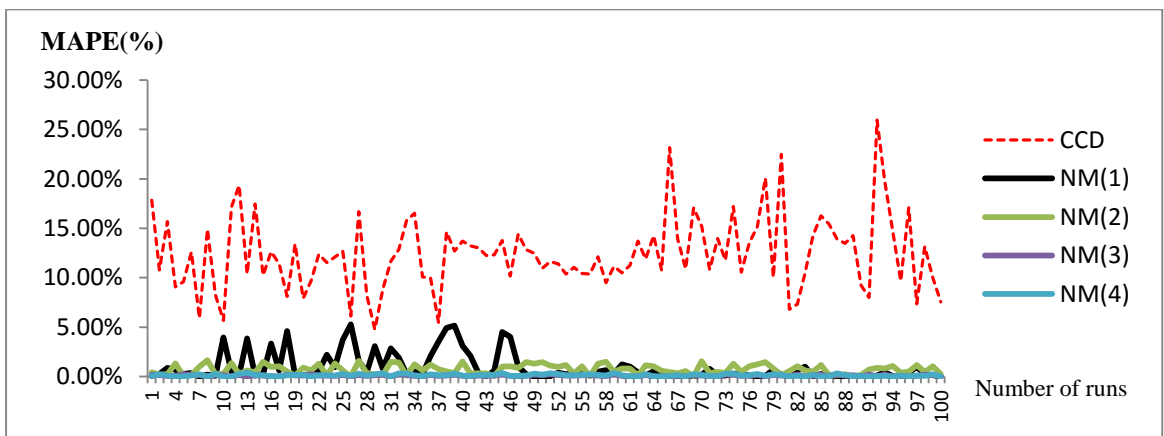
MAPE results for function  $f_3$  from Figures 4.30 - 4.32 are summarized in Table 4.10. The results show that the proposed method forms NM(1), NM(2), NM(3), and NM(4) showed superior performance over the classical RSM(CCD) by at least an order of magnitude. Even the worst results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) when used to identify the true  $x_1, x_2$  and  $y$ , respectively were far better than those of the classical RSM(CCD). Once again, take note of the fact that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) were quite similar.



**Figure 4.33** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_1$  of  $f_4$



**Figure 4.34** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_2$  of  $f_4$



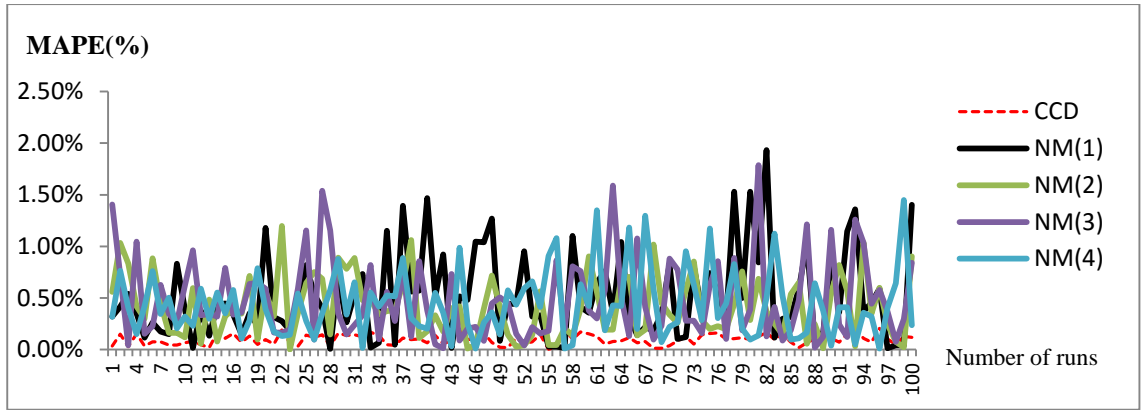
**Figure 4.35** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $y$  of  $f_4$



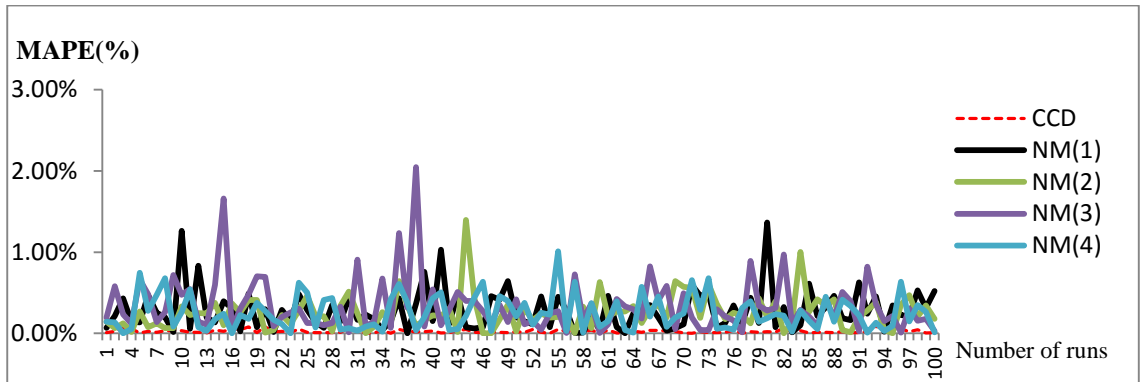
**Table 4.11** Comparison of the MAPE of the classical RSM(CCD) and the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_4$

Compare with	Average MAPE (%)				
	Method				
	RSM	NM(1)	NM(2)	NM(3)	NM(4)
True $x_1$	1.2809	0.9033	0.5789	0.7628	0.5569
True $x_2$	3.2962	0.8316	0.3380	0.8080	0.3309
True $y$	12.4090	0.9092	0.7167	0.1041	0.1338

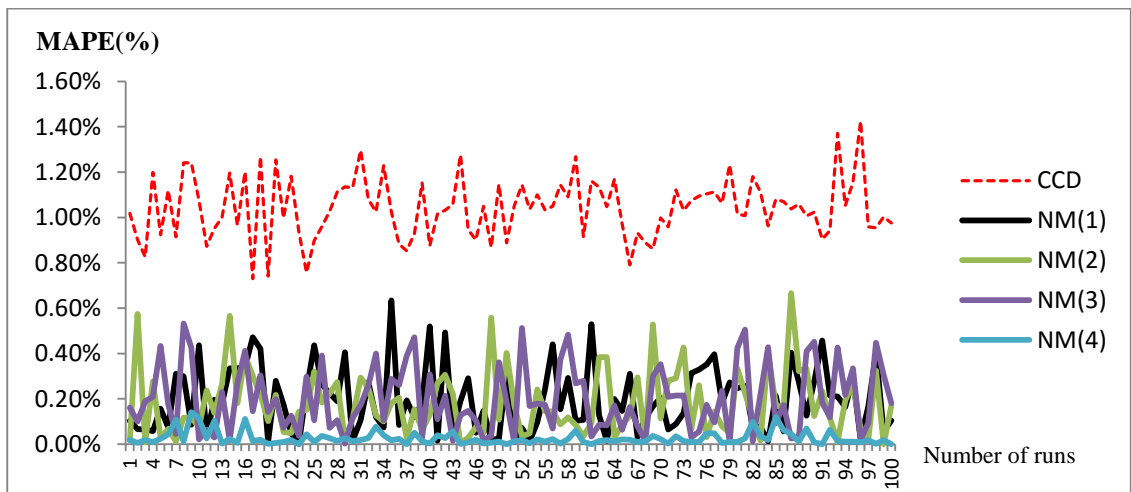
From the MAPE results for function  $f_4$  in Figures 4.33 - 4.35 summarized in Table 4.11, we can see that the proposed method forms NM(1), NM(2), NM(3), and NM(4) obtained a much higher performance than the classical RSM(CCD). Even the worst performance of the proposed method NM(1) was better than the classical RSM(CCD). As before, note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) are not much different.



**Figure 4.36** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $x_1$  and  $f_5$



**Figure 4.37** Comparison of the MAPE of RSM (CCD) to the proposed method for the True  $x_2$  and  $f_5$



**Figure 4.38** Comparison of the MAPE of RSM (CCD) to the Proposed Method for the True  $y$  and  $f_5$

**Table 4.12** Comparison of the MAPE of the Classical RSM(CCD) to the Proposed Method Forms NM(1), NM(2), NM(3), and NM(4) for Function  $f_5$

Compare with	Average MAPE (%)				
	Method				
	RSM	NM(1)	NM(2)	NM(3)	NM(4)
True $x_1$	0.0970	0.5141	0.4148	0.4848	0.4395
True $x_2$	0.0254	0.2814	0.2574	0.3401	0.2525
True $y$	1.0416	0.1979	0.1894	0.2107	0.0262

From the MAPE results from Figures 4.36 - 4.38 summarized in Table 4.12 for function  $f_5$ , we can see that the proposed method forms NM(1), NM(2), NM(3), and NM(4) and the classical RSM(CCD) performed similarly; this is because function  $f_5$  is quadratic. Note that the results of the proposed method forms NM(1), NM(2), NM(3), and NM(4) were once again quite similar.

**Table 4.13** The True Value, Mean and Standard Deviation of the Estimated Values of the Minimum Design Points and the Minimum Responses in the Simulation by the Classical RSM(CCD) and the Proposed Method

Test Function	Variable	True Value	Method				
			CCD	NM(1)	NM(2)	NM(3)	NM(4)
$f_1$	$x_1$	3.2	3.085 (0.049)	3.210 (0.036)	3.198 (0.019)	3.194 (0.018)	3.195 (0.029)
	$x_2$	2.1	2.208 (0.034)	2.113 (0.023)	2.103 (0.016)	2.12 (0.026)	2.119 (0.031)
	$y$	-6.51	-4.089 (0.352)	-6.396 (0.168)	-6.504 (0.044)	-6.493 (0.052)	-6.519 (0.017)
$f_2$	$x_1$	-0.27	-0.383 (0.070)	-0.284 (0.009)	-0.29 (0.005)	-0.282 (0.008)	-0.28 (0.010)
	$x_2$	-0.92	-0.993 (0.099)	-0.955 (0.034)	-0.945 (0.039)	-0.94 (0.050)	-0.945 (0.045)
	$y$	-181.6	-178.537 (1.14)	-181.603 (0.424)	-181.243 (0.657)	-181.255 (0.453)	-181.645 (0.447)
$f_3$	$x_1$	-0.092	-0.032 (0.015)	-0.096 (0.005)	-0.097 (0.006)	-0.095 (0.003)	-0.096 (0.006)
	$x_2$	0.713	0.334 (0.094)	0.716 (0.021)	0.71 (0.012)	0.709 (0.013)	0.708 (0.019)
	$y$	-1.032	-0.514 (0.102)	-1.038 (0.05)	-1.038 (0.051)	-1.017 (0.022)	-1.017 (0.016)
$f_4$	$x_1$	2.77	2.747 (0.03)	2.746 (0.025)	2.776 (0.018)	2.758 (0.024)	2.767 (0.018)
	$x_2$	2.46	2.376 (0.045)	2.454 (0.025)	2.455 (0.029)	2.455 (0.024)	2.456 (0.009)
	$y$	-5.41	-4.828 (0.171)	-5.272 (0.117)	-5.442 (0.034)	-5.413 (0.006)	-5.414 (0.008)
$f_5$	$x_1$	86.9	86.816 (0.044)	86.572 (0.477)	86.818 (0.436)	86.65 (0.483)	86.801 (0.463)
	$x_2$	176.67	176.655 (0.055)	176.628 (0.666)	176.624 (0.588)	176.47 (0.815)	176.613 (0.582)
	$y$	-83.22	-82.353 (0.112)	-83.131 (0.184)	-83.152 (0.196)	-83.119 (0.218)	-83.234 (0.03)

**Note:** The Numbers in Parentheses are the Standard Deviations of the Estimated Means.

From Table 4.13 we can see that the true value of  $x_1$  in  $f_1 - f_4$ , the mean of estimated  $x_1$  in the proposed method : NM(1), NM(2), NM(3) and NM(4) very close to the true value of  $x_1$  than the mean of estimated  $x_1$  in the classical RSM(CCD). Only the mean of estimated  $x_1$  in  $f_5$  close to the true value of  $x_1$  than the mean of estimated  $x_1$  in NM(1) , NM(3) and NM(4). This is because function  $f_5$  is quadratic. The true value of  $x_2$  in  $f_1 - f_4$ , the mean of estimated  $x_2$  in the proposed method : NM(1), NM(2), NM(3) and NM(4) very close to the true value of  $x_2$  than the mean of estimated  $x_2$  in the classical RSM(CCD). Only the mean of estimated  $x_2$  in  $f_5$  close to the true value of  $x_2$  than the mean of estimated  $x_1$  in NM(1), NM(2), NM(3) and NM(4). This is because function  $f_5$  is quadratic. Consider the true value of  $y$  in  $f_1 - f_5$ , the mean of estimated  $y$  in the proposed method : NM(1), NM(2), NM(3) and NM(4) very close to the true value of  $y$  than the mean of estimated  $y$  in the classical RSM(CCD).

## CHAPTER 5

### CONCLUSIONS AND FUTURE WORK

#### 5.1 Conclusions

In this dissertation, an improved method for finding the optimum region using the response surface methodological approach was proposed. This classical method is an attempt to find a suitable approximation of the true relationship by using the path of steepest ascent/descent to move points towards the optimum region. After this, a second-order model is fitted, then stationary points are obtained. In real-life situations, the true relationship between independent variables and the response is not a second-order model, then the optimum obtained by the RSM may be far from the true optimum. Therefore, an improvement on this classical method was to use the NM algorithm to move towards the optimum region. This algorithm does not require a gradient function, and instead moves points towards the optimum region by adapting itself heuristically. Therefore, this algorithm can be used to solve any pattern of function.

In the simulation study, five mathematical test functions with different patterns were used to compare the efficiency of the classical RSM(CCD) to the proposed method forms NM(1), NM(2), NM(3), and NM(4) in terms of the average of the number of points, coverage probability, and MAPE. The results of the simulation study allude to the following conclusions. In almost every situation, the average of the number of points for the proposed method was less than the classical RSM(CCD). The results of coverage probability showed that the proposed method performed better than the classical RSM(CCD), especially for the coverage probability of both true values ( $X_1$ ,  $X_2$ ) being contained in the 95% CIs of the means of stationary  $X_1, X_2$ . These results point towards the fact that the proposed method is far superior to the classical RSM(CCD). When considering MAPE, the results obtained using the proposed method showed a much higher performance than the classical RSM(CCD).

In other words, the fitted model by the proposed method has a value nearby to the true value.

The exception was testing with function  $f_5$  where the results between both methods are quite similar. This is because this test function is quadratic, which enabled the classical RSM(CCD) to obtain good results. However the other functions are not quadratic, and so results for them using classical RSM(CCD) showed less efficiency than the proposed method.

From the simulation study results, we can say that if the true relationship between response and independent variables is quadratic, the efficiency in terms of average of points, coverage probability, and MAPE is comparable for both methods. However, if the true relationship between the response and independent variables is not quadratic, the efficiency in terms of average of points, coverage probability, and MAPE of the proposed method is much higher than that of the classical RSM(CCD).

## **5.2 Future Work**

In this dissertation, classical RSM(CCD) is used as a comparison. A further study could consider other methods such as the Box-Behnken design and Plackett–Burman designs. Another point of interest would be to extend the independent variables to perhaps more than three factors.

## BIBLIOGRAPHY

- Anderson, C.; Borror, C. M. and Montgomery, D. C. 2009. Response Surface Design Evaluation and Comparison. **Journal of Statistical Planning and Inference**. 139 (2): 629-674.
- Box, G. E. P. 1966. A Note on Augmented Designs. **Technometrics**. 8(1):184-188.
- Box, G. and Behnken, D. 1960. Some New Three Level Designs for the Study of Quantitative Variables. **Technometrics**. 2: 455-475.
- Box, G. E. P. and Draper, N. R. 1959. A Basis for the Selection of a Response Surface Design. **Journal of the American Statistical Association**. 54 (287): 622-654.
- Box, G. E. P. and Draper, N. R. 1963. The Choice of a Second Order Rotatable Design. **Biometrika**. 50 (3): 335-352.
- Box, G. E. P. and Draper, N. R. 1987. **Empirical Model-Building and Response Surfaces**. New York: Wiley.
- Box, G. E. P. and Draper, N. R. 2007. **Response Surfaces, Mixtures, and Ridge Analyses**. 2<sup>nd</sup> ed. [of Empirical Model-Building and Response Surfaces, 1987]. New York: Wiley.
- Box, G. E. P. and Wilson, K. B. 1951. On the Experimental Attainment of Optimum Conditions. **Journal of the Royal Statistical Society**. 13 (1):1-45.
- Brooks, S. P. and Morgan, B. J. T. 1995. Optimization Using Simulated Annealing. **The Statistician**. 44 (2): 241-257.
- Byatt, D. 2000. **Convergent Variants of the Nelder-Mead Algorithm**. Master's thesis, University of Canterbury.
- Copleland, K. A. F. and Nelson, P. R. 1996. Dual Response Optimization Via Direct Function Minimization. **Journal of Quality Technology**. 28 (3): 331-336.
- Dennis, J. E. and Torczon, V. (1991). Direct Search Methods on Parallel Machines. **SIAM Journal on Optimization**. 1 (4): 448-474.
- Draper, N. R. and Lawrence, W. E. 1965. Designs Which Minimize Model Inadequacies: Cuboidal Regions of Interest. **Biometrika**. 52 (1/2): 111-118.



- Draper, N. R. and John, J. A. 1988. Response-surface Designs for Quantitative and Qualitative Variables. **Technometrics**. 30 (4): 423-428.
- Fan, S. K. S. and Zahara, E. 2004. Simulation Optimization Using an Enhanced Nelder Mead Simplex Search Algorithm. **Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference**.
- Gavin, H. P. 2013. The Nelder-Mead Algorithm in Two Dimensions. **CEE 201L**. Duke U.
- Glover, F. and Laguna, M. 1997. **Tabu Search**. Kluwer Academic Publishers, Boston.
- Grabitech Solutions, A. B. 2001. **MultiSimplex Release 2.1 for Windows**, Grabitech Solutions AB, Sundsvall, Sweden.
- Han, L. and Neumann M. 2006. Effect of Dimensionality on the Nelder-Mead Simplex Method. **Optim. Methods Software**. 21(1): 1-16.
- Holland, J. H. 1992. Genetic Algorithms. **Scientific American**. 267 (1): 66-72.
- Kelley, C. T. 2000. Detection and Remediation of Stagnation in the Nelder-Mead Algorithm Using a Sufficient Decrease Condition. **SIAM Journal on Optimization**. 10 (1): 43-55.
- Khuri, A. I. 1988. A Measure of Rotatability for Response Surface Designs. **Technometrics**. 30 (1): 95-104.
- Khuri, A. I. and Cornell, J. A. 1987. **Response Surfaces: Designs and Analyses**. New York: Marcel Dekker.
- Khuri, A. I. and Cornell, J. A. 1996. **Response Surfaces: Design and Analysis**. 2<sup>nd</sup> ed. New York: Marcel Dekker, Monticello.
- Kolda, T. G.; Lewis, R. M. and Torczon, V. 2003. Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods. **SIAM Journal Review**. 45 (3): 385-482.
- Lagarias, J. C.; Reeds, J. A.; Wright, M. H. and Wright, P. E. 1998. Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. **SIAM Journal on Optimization**. 9 (1): 112-147.
- Mckinnon, K. I. M. 1998. Convergence of the Nelder-Mead Simplex Method to a Nonstationary Point. **SIAM Journal Optimization**. 9 (1): 148-158.

- Montgomery, D. C. 2005. **Introduction to Statistical Quality Control.**  
5<sup>th</sup> ed. Hoboken, NJ: John Wiley
- Myers, R. H. and Montgomery, D. C. 2002. **Response Surface Methodology: Process and Product Optimization Using Designed Experiments.**  
2<sup>nd</sup> ed. New York: Wiley.
- Myers, R. H.; Montgomery, D. C. and Anderson-Cook, C. M. 2009. **Response Surface Methodology: Product and Process Optimization Using Designed Experiments.** 3<sup>rd</sup> ed. New York: Wiley.
- Myers, R. H.; Montgomery, D. C.; Vining, G. G.; Borror, C. M. and Kowalski, S. M. 2004. Response Surface Methodology: A Retrospective and Literature Survey. **Journal of Quality Technology.** 36(1): 53–77.
- Nelder, J. A. and Mead, R. 1965. A Simplex Method for Function Minimization. **The Computer Journal.** 7 (4): 308-313.
- Olsson, D. M. 1974. A Sequential Simplex Program For Solving Minimization Problems. **Journal of Quality Technology.** 6 (1): 53-57.
- Olsson, D. M. and Nelson, L. S. 1975. The Nelder-Mead Simplex Procedure for Function Minimization. **Technometrics.** 17 (1): 45-51.
- Parkinson, J. M. and Hutchinson, D. 1972. An Investigation into the Efficiency of the Variants on the Simplex Method. In F. A. Lootsma (Ed.), **Numerical Methods for Non-Linear Optimization.** London: Academic Press. Pp. 115-135.
- Price, C.J.; Coope, I. D. and Byatt, D. 2002. A Convergent Variant of the Nelder-Mead Algorithm. **Journal Optimization Theory Applications.** 113 (1): 5-19.
- Russell, V. 2009. Response-Surface Methods in R, Using rsm. **Journal of Statistical Software.** 32 : 7.
- SAS Institute. 1998. **SAS/IML Software Release 6.12 for Windows,** SAS Institute Inc., Cary, N. C.
- Sexton, R. S., Alidaee, B., Dorsey, R. E. and Johnson, J. 1998. Global Optimization for Artificial Neural Networks. A Tabu Search Application. **European Journal Operational Research.** 106 (2-3): 570-584.

- Spendley, W.; Hext, G.R. and Himsworth, F.R. 1962. Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation. **Technometrics**. 4: 441-461.
- Torczon, V. 1989. **Multi-Directional Search: A Direct Search Algorithm for Parallel Machines**. PhD thesis, Rice University, TX.
- Torczon, V. 1997. On the Convergence of Pattern Search Methods. **SIAM Journal on Optimization**. 7 (1): 1-25.
- Tseng, P. 2000. Fortified-Descent Simplicial Search Method: A General Approach. **SIAM Journal on Optimization**. 10 (1): 269-288.
- Wolfram Research. 2002. **Mathematica/Optimization and Statistics Release**. 4.2. Wolfram Research Inc.
- Wright, M. H. 1996. Direct Search Methods: Once Scorned, Now Respectable. In D.F. Griffiths and G.A. Watson(Eds.) Numerical Analysis 1995: **Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis**(Harlow, UK: Addison Wesley Longman. Pp. 191-208.

## **BIOGRAPHY**

### **NAME**

Miss Chantha Wongoutong

### **ACADEMIC BACKGROUND**

B.Sc. (Computer science),  
Silpakorn University, 1999  
M.Sc. (Statistics),  
Kasetsart University, 2005

### **PRESENT POSITION**

Lecturer, Program of Statistics,  
Department of Statistics  
Faculty of Science, Kasetsart University.

### **EXPERIENCES**

#### **Teaching:**

Introduction to Statistics,  
Time Series Analysis,  
Statistical Analysis with Program  
Quantitative Analysis