Thesis Proposal:

THE RESPONSE SURFACE METHODOLOGY

Nuran Bradley
Department of Mathematical Sciences
Indiana University of South Bend
E-mail Address: nbradley@iusb.edu

Submitted to the faculty of the
Indiana University South Bend
in partial fulfillment of requirements for the degree of

MASTER OF SCIENCE

in

APPLIED MATHEMATICS & COMPUTER SCIENCE

Advisor

Dr. Yi Cheng
Department of Mathematical Science

Committee:
Dr. Zhong Guan
Dr. Dana Vrajitoru

Abstract

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

There are many types of experiments used in real-world situations and problems. When treatments are from a continuous range of values then the true relationship between y and x's might not be known. The approximation of the response function $y = f(x_1, x_2, \ldots, x_q) + \varepsilon$ is called *Response Surface Methodology*. This thesis puts emphasis on designing, modeling, and analyzing the *Response Surface Methodology*. The three types of *Response Surface Methodology*, the first-order, the second-order, and the mixture models, will be explained and analyzed in depth. The thesis will also provide examples of application of each model by numerically and graphically using computer software.

TABLE OF CONTENTS

List of Figures

1. Introduction

As an important subject in the statistical design of experiments, the *Response Surface Methodology* (*RSM*) is a collection of mathematical and statistical techniques useful for the modeling and analysis of problems in which a response of interest is influenced by several variables and the objective is to optimize this response (Montgomery 2005). For example, the growth of a plant is affected by a certain amount of water x_1 and sunshine x_2. The plant can grow under any combination of treatment x_1 and x_2. Therefore, water and sunshine can vary continuously. When treatments are from a continuous range of values, then a Response Surface Methodology is useful for developing, improving, and optimizing the response variable. In this case, the plant growth y is the response variable, and it is a function of water and sunshine. It can be expressed as

$$y = f(x_1, x_2) + \varepsilon$$

The variables x_1 and x_2 are independent variables where the response y depends on them. The dependent variable y is a function of x_1, x_2, and the experimental error term, denoted as ε. The error term ε represents any measurement error on the response, as well as other type of variations not counted in f. It is a statistical error that is assumed to distribute normally with zero mean and variance σ^2. In most *RMS* problems, the true response function f is unknown. In order to develop a proper approximation for f, the experimenter usually starts with a low-order polynomial in some small region. If the response can be defined by a linear function of independent variables, then the approximating function is a **first-order model**. A first-order model with 2 independent variables can be expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

If there is a curvature in the response surface, then a higher degree polynomial should be used. The approximating function with 2 variables is called a **second-order model**:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_{11}^2 + \beta_{22} x_{22}^2 + \beta_{12} x_1 x_2 + \varepsilon$$

In general all *RSM* problems use either one or the mixture of the both of these models. In each model, the levels of each factor are independent of the levels of other factors. When the levels of each factor are not independent then a **mixture model** is appropriate for designing an *RMS* model.

In order to get the most efficient result in the approximation of polynomials the proper experimental design must be used to collect data. Once the data are collected, the *Method of Least Square* is used to estimate the parameters in the polynomials. The response surface analysis is performed by using the fitted surface. The **response surface designs** are types of designs for fitting response surface. Therefore, the objective of studying *RSM* can be accomplish by

(1) understanding the topography of the response surface (local maximum, local minimum, ridge lines), and

(2) finding the region where the optimal response occurs. The goal is to move rapidly and efficiently along a path to get to a maximum or a minimum response so that the response is optimized.

2. Literature Reviews

The *RSM* is important in designing, formulating, developing, and analyzing new scientific studying and products. It is also efficient in the improvement of existing studies and products. The most common applications of *RSM* are in Industrial, Biological and Clinical Science, Social Science, Food Science, and Physical and Engineering Sciences. Since *RMS* has an extensive application in the real-world, it is also important to know how and where *Response Surface Methodology* started in the history. According to Hill and Hunter, *RSM* method was introduced by G.E.P. Box and K.B. Wilson in 1951 (Wikipedia 2006). Box and Wilson suggested to use a first-degree polynomial model to approximate the response variable. They acknowledged that this model is only an approximation, not accurate, but such a model is easy to estimate and apply, even when little is known about the process (Wikipedia 2006). Moreover, Mead and Pike stated origin of RMS starts 1930s with use of *Response Curves* (Myers, Khuri, and Carter 1989).

According to research conducted (Myers, Khuri, and Carter 1989), the *orthogonal design* was motivated by Box and Wilson (1951) in the case of the first-order model. For the second-order models, many subject-matter scientists and engineers have a working knowledge of the *central composite designs* (CCDs) and *three-level designs* by Box and Behnken (1960). Also, the same research states that another important contribution came from Hartley (1959), who made an effort to create a more economical or *small composite design*. The techniques used in mixture models became important to *RSM* users by the work of Scheffé in 1950s. According to (Myers, Khuri, and Carter 1989), the important development of optimal design theory in the field of experimental design emerged following Word World II. Elfving (1952, 1955, 1959), Chernoff (1053), Kiefer (1958, 1959, 1960, 1962), and Kiefer and Wolfowitz were some of the various authors who published their work on optimality.

One of the important facts is whether the system contains a maximum or a minimum or a saddle point, which has a wide interest in industry. Therefore, *RMS* is being increasingly used in the industry. Also, in recent years more emphasis has been placed by the chemical and processing field for finding regions where there is an improvement in response instead of finding the optimum response (Myers, Khuri, and Carter 1989). In result, application and development of *RMS* will continue to be used in many areas in the future.

3. Response Surface Methods and Designs

Response Surface Methods are designs and models for working with continuous treatments when finding the optima or describing the response is the goal (Oehlert 2000). The first goal for Response Surface Method is to find the optimum response. When there is more than one response then it is important to find the compromise optimum that does not optimize only one response (Oehlert 2000). When there are constraints on the design data, then the experimental design has to meet requirements of the constraints. The second goal is to understand how the response changes in a given direction by adjusting the design variables. In general, the response surface can be visualized graphically. The graph is helpful to see the shape of a response surface; hills, valleys, and ridge lines. Hence, the function $f(x_1, x_2)$ can be plotted versus the levels of x_1 and x_2 as shown as Figure 3-1 (Montgomery 2005).

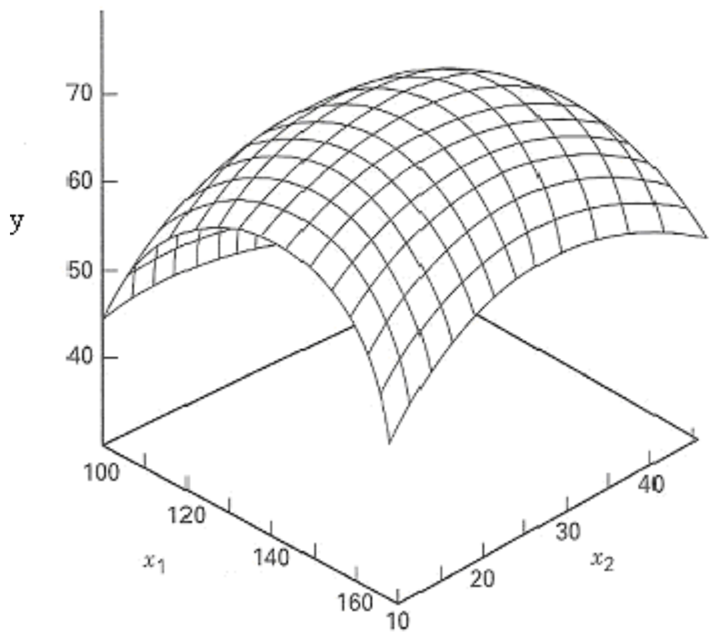

Figure 3-1 Response surface plot

$$y = f(x_1, x_2) + \varepsilon$$

In this graph, each value of x_1 and x_2 generates a y-value. This three-dimensional graph shows the response surface from the side and it is called a **response surface plot**. Sometimes, it is less complicated to view the response surface in two-dimensional graphs. The contour plots can show contour lines of x_1 and x_2 pairs that have the same response value y. An example of contour plot as shown in Figure 3-2.

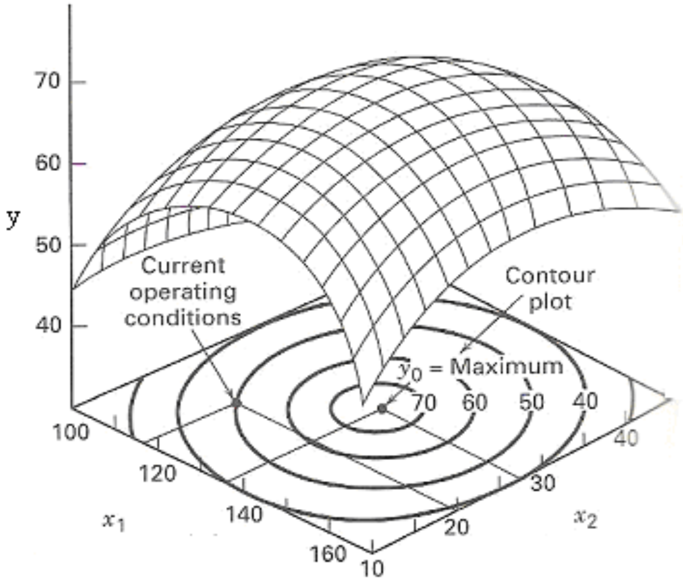

Figure 3-2 Contour plot

In order to understand the surface of a response, graphs are helpful tools. But, when there are more than two independent variables, graphs are difficult or almost impossible to use to illustrate the response surface, since it is beyond 3-dimension. For this reason, response surface models are essential for analyzing the unknown function *f*.

4. First-Order Model

4.1 Analysis of a First-Order Response Surface

The relationship between the response variable *y* and independent variables is unknown. In general, the low-order polynomial model is used to describe the response surface *f*. The polynomial models are usually a sufficient approximation in a small region of response surface. Therefore, depending on the approximation of unknown function *f*, either first-order or second-order models are employed.

Furthermore, the approximating function *f* is a first-order model when the response is a linear function of independent variables. The first-order model with *N* experimental runs is carrying out on *q* design variables and a single response *y* as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots\ldots\ldots + \beta_q x_{iq} + \varepsilon_i \qquad (i = 1, 2, \ldots\ldots, N)$$

The response *y* is a function, *f*, of the design variables $x_1, x_{2,\ldots,} x_q$, plus the experimental error. The first-order model is a *multiple-regression* model and β_i 's are regression coefficients.

First-order model is used to describe flat surfaces with or without tilted surfaces. This model is not suitable for analyzing maximum, minimum, and ridge lines. Using first-order model approximation of f is reasonable when f is not too curved in that region and the region is not too big. First-order model is assumed to be an adequate approximation of the true surface in a small region of the x's (Montgomery 2005). Moreover, first-order model indicates which way is up and down in the response. The *method of steepest ascent* is a procedure in which the algorithm follows the direction to move to increase response the most, which is used to identify a maximum. The *method of steepest descent* consists in taking the direction of the most quickly decrease in the response, which is used to identify the minimum.

4.2 Designs for Fitting the First-Order Model

The design of response surface models starts with the estimation of parameters, *pure error*, and *lack of fit*. Also, the experimenter needs to design a model that is efficient. Therefore, estimation of variances has to be taken into consideration. The **orthogonal first-order designs** minimize the variance of the regression coefficients β_k. A first-order design is orthogonal if the off-diagonal elements of the $(\mathbf{X}'\mathbf{X})$ matrix are all zero (Montgomery 2005). The orthogonal first-order designs includes 2^q *factorial with center points* and 2^{q-k} *fraction with resolution III or greater*.

4.3 My Objective of First-Order Model

In order to conduct a first-order model, I intent to study the operating conditions that maximize the yield of a process. There are two independent variables which influence the process yield: reaction time and reaction temperature. I will use low-order polynomial terms to describe some part of the response surface. Once the estimated equation is obtained, I will be able to use statistical techniques to check for the model adequacy.

My objective is to determine if the current levels or settings of the reaction time and the temperature result in a value of a response that is close to the optimum. If the response is not near the optimum, then the application of the method of steepest ascent will be needed. When the region of the optimum is located, I will begin to study a more structured response surface model such as second-order model.

5. Second-Order Model

5.1 Analysis of a Second-Order Response Surface

When there is a curvature in the response surface the first-order model is insufficient. Therefore, second-order model is useful in approximating a portion of the true response surface with curvature. The second-order model includes all the terms in the first-order model, and quadratic and cross product terms. It is usually represented as

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

The second-order models illustrate quadratic surfaces such as minimum, maximum, ridge, and saddle. If there exits an optimum then this point is called *stationary point*. The *stationary point* is the combination of design variables where the surface is at either a maximum or a minimum in all directions. If the *stationary point* is a maximum in some direction and minimum in another direction, then the *stationary point* is a *saddle point* (Oehlert 2000). The graphical visualization is very helpful in understanding the second-order response surface as it shown in Figure 3.2. Specifically, contour plots can help characterize the shape of the surface and locate the optimum response roughly.

5.2 Designs for Fitting the Second-Order Model

The most popular design for fitting the second-order model is **Central Composite Design** (**CCD**). It consists of factorial point (from a 2^q design and 2^{q-k} *fraction with resolution V or greater*), central point, and axial points. CCD often develops through a *sequential experimentation*. When the first-order model shows an evidence of *lack of fit*, then axial points can be added to quadratic terms and with more center points to develop CCD. The number of center points *m* at the origin and the distance *α* of the axial runs from the design center are two parameters in the CCD design.

There are a couple of ways of choosing *α* and *m*. First, CCD can run in incomplete blocks. A block is a set of relatively homogeneous experimental conditions so that an experimenter divides the observations into groups that are run in each block. An incomplete block design can be conducted when all treatment combinations cannot be run in each block. In order to protect the shape of the response surface, block effects need to be orthogonal to treatment effects. This can be done by choosing the correct *α* and *m* in factorial and axial blocks.

Also, *α* and *m* can be chosen so that the CCD is not blocked. If the precision of the estimated response surface at some point x depends only on the distance from x to the origin, not on the direction, then the design is said to be *rotatable* (Oehlert 2000). The rotatable design provides equal precision of estimation of the surface in all directions. The choice of *α* will make the CCD design rotatable by using either $\alpha = 2^{q/4}$ for the full factorial or $\alpha = 2^{(q-k)/4}$ for a fractional factorial.

In addition to CCD, **Box-Behnken design** can also be used for designing response surfaces. This model is a combination of 3^q factorials with incomplete block designs.

5.3 My Objective of Second-Order Model

With the purpose of exploring the second-order model, I intent to use a statistical modeling to develop an appropriate approximating relationship between the yield and the process variables: the temperature and the time. I will fit a second-order model using Central Composite Design. Since the second-order model is very flexible, consequently I can experiment with a wide variety of functional forms. The method of least squares can be used to estimate the parameters. Once an appropriate approximating model is obtained, I can analyze to determine the optimum conditions.

Moreover, my other objective is to include in my study another data set which combines with more independent variables. Therefore, the complexity and the design of the data will illustrate a different perspective on the second-order model.

6. Mixture Model

6.1 Analysis of a Mixture Experiment

In the mixture model the levels of each factor are independent of the levels of other factors. The mixtures depend on the proportion of a variety of components. For example, if design variables $x_{1,} x_{2,}, ..., x_q$ denote the proportion of q components of a mixture, then

$$x_k \geq 0 \qquad k = 1,2,\ldots\ldots,q$$

and

$$x_1, + x_2 + \ldots + x_q = 1$$

This kind of design space is called a **simplex** in q dimension. For example, in two dimensions, the design space is the segment from (0,1) to (1,0); in three dimensions it is bounded by the equilateral triangle (0,0,1), (0,1,0), and (1,0,0) (Oehlert 2000). The graphical representation of components for q = 2 and q = 3 is illustrated in Figure 6.1 (Montgomery 2005).

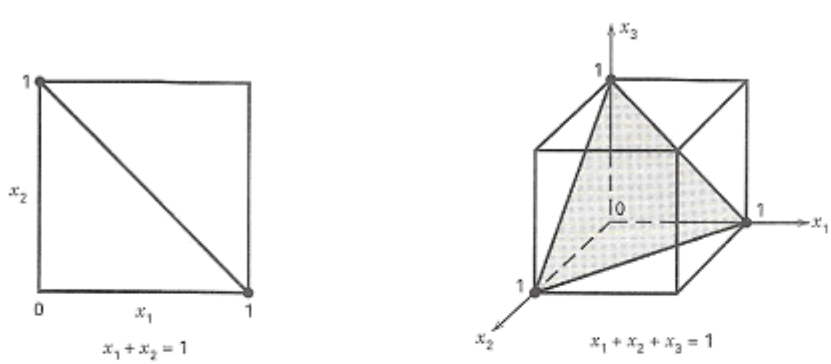

Figure 6-1 Constrained factor space for mixtures with q = 2 and q = 3

6.2 Designs for Fitting the Mixture Model

Simplex designs study the effects of mixture components on the response variable. A $\{q,m\}$ **simplex lattice design** for q components consists of points defined by the following coordinate settings (Montgomery 2005):

$$x_k = 0, \frac{1}{m}, \frac{2}{m}, ..., 1 \qquad k = 1.2....., q$$

where each component takes the m + 1 equally spaced values from 0 to 1. For example, let q = 3 and m = 2.

$$x_k = 0, \frac{1}{2}, 1 \qquad k = 1,2,3$$

$$(x_1, x_2, x_3) = (1,0,0), (0,1,0), (0,0,1), (\frac{1}{2}, \frac{1}{2}, 0), (\frac{1}{2}, 0, \frac{1}{2}), (0, \frac{1}{2}, \frac{1}{2})$$

Here, the first 3 vertices are pure blends; other 3 points are binary blends. In general simplex lattice designs provide N number of points.

$$N = \frac{(q + m - 1)!}{m!(q - 1)!}$$

Another simplex lattice design is the **simplex centroid design**. In this design, there are 2^q-1 points that correspond to the q permutations of $(1,0,0,.....,0)$, the $\binom{q}{2}$ permutations of $(\frac{1}{2}, \frac{1}{2}, 0, ...,0)$, the $\binom{q}{3}$ permutations of $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, ...,0)$ and the overall centroid $(\frac{1}{q}, \frac{1}{q}, ..., \frac{1}{q})$ (Montgomery 2005).

6.3 My Objective of Mixture Model

In this section, my goal is to analyze the use of synthetic mixtures of sand, silt, and clay which affect the growth of potatoes. The form of mixture polynomial is slightly different from the standard polynomials used in response surface methodology. Therefore, I will use simplex designs that allow response surface model to be fit over the entire mixture space. In the basis of analysis of variance for this case study, I will fit a model for the data. The contour plots will also help analyzing the area of highest growth potatoes. A second objective is to visualize how each component has an effect on the response relative to the reference blend. As a result, I will conclude my study with an analysis of trace plots.

7. Conclusion

The success of *RMS* depends on an estimation of y at different locations in the response surface. Therefore, the experimenter can draw a conclusion about whether the system contains optimum or improvement in response. Before any experimenter starts the analysis of the response surface, the application of *RMS* first begins with investigation of factors or variables. In order to obtain an efficient experiment, unimportant independent variables need to be separated from important ones. One should never start an analysis of the surface until significant factors are identified. After that, the response surface study can start. Hill and Hunter outline four steps for response surface analysis (Myers, Khuri, and Carter 1989). They are:

(1) perform a statistically designed experiment,
(2) estimate the coefficients in the response surface equation,
(3) check on the adequacy of the equation (via a lack-of-fit test), and
(4) study the response surface in the region of interest.

These steps are very important because they help the experimenter answer certain questions regarding the response surface such as (1) how much replication necessary, (2) the location of the region of the optimum, (3) the type of approximating function required, (4) the proper choice of experimental designs, and (5) whether or not transformations on the responses or any of the process variables is required (Myers, Montgomery 1995). Beside statistical and mathematical techniques the graphical representation of the response surface is also helpful in finding answers to problems. Due to broader applications in real-word problems, *RMS* will continue to attract statisticians, engineers, and scientists in order to develop, improve, and optimize new or existing products and processes.

Furthermore, my goal is to bring some perspective on how to design and analyze response surfaces in this thesis. Some case studies will help to illustrate applications of response surface models. Since many response surface designs are available, I will only use the ones which will be appropriate for the data. Moreover, the thesis will include a statistical analysis of the surface numerically and graphically using Minitab. The computations in matrix form will be accomplished by Excel. When the response surface requires a further study, the simulated response variables will be obtained by R-project. Therefore, each *RSM* model will be concluded by detailed analysis of the case studies.

Bibliography

Anderson, Mark J. and Whitcomb, Patrick J. 2004. Design solutions from concept through manufacture: Response surface methods for process optimization. *Desktop Engineering.* http://www.deskeng.com/ (accessed May 16, 2006).

Minto, Charles. 2006. Response Surface Modeling of Drug Interactions. http://eurosiva.org/Archive/Vienna/abstracts (accessed May 16, 2006).

Montgomery, Douglas C. 2005. *Design and Analysis of Experiments: Response surface method and designs.* New Jersey: John Wiley and Sons, Inc.

Myers, Raymond H., Khuri, Andre I. and Carter, Walter H., Jr. 1989. Response surface methodology: *1966-1988*. *Technometrics* 31 (2): 137-153 http://www.jstor.org/ (accessed January 29, 2007).

Myers, Raymond H. and Montgomery, Douglas C. 1995. *Response Surface Methodology: process improvement with steepest ascent, the analysis of response Surfaces, experimental designs for fitting response surfaces,* 183-351. New York: John Wiley and Sons, Inc.

Oehlert, Gary W. 2000. *Design and analysis of experiments: Response surface design.* New York: W.H. Freeman and Company

Shi, Y. and Weimer P.J. 1992. Response surface analysis of the effects of pH and Dilution rate on ruminococcus flavefaciens FD-1 in cellulose-fed continues Culture. *American Society for Microbiology* 58 (8): 2583-2591.

Verseput, Richard. 2000. Digging into DOE: Selecting the right central composite Design for response surface methodology applications. *Quality Digest*. http://www.qualitydigest.com/june01/html/doe.html (accessed June 7, 2006).

2006. Response surface methodology. http://en.wikipedia.org/wiki/response_surface_methodology (accessed January 22, 2007).