



Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations

Graduate School

2010

Multivariate Steepest Ascent Using Bayesian Reliability

Jeffrey Fuerte

Virginia Commonwealth University

Follow this and additional works at: <http://scholarscompass.vcu.edu/etd>

 Part of the [Physical Sciences and Mathematics Commons](#)

© The Author

Downloaded from

<http://scholarscompass.vcu.edu/etd/75>

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Multivariate Steepest Ascent Using the Bayesian Reliability Function

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

Jeffrey Norman Fuerte, Jr.
Bachelor's of Science, 2008

Director: Dr. David Edwards, Assistant Professor, Department of Statistical Sciences and Operations Research

Virginia Commonwealth University
Richmond, Virginia
May, 2010

Acknowledgment

The author wishes to thank several people. I would like to thank my wife, Caryn, for her support and love, as well as her patience in this “never ending” school career. Also, for instilling the values and ethics needed to survive in this world, I would like to thank my parents, grandparents, and other family members. Similarly, without Dr. Edwards, this thesis would have been without a direction; thus, I would like to express my gratitude towards him. If not for his support and willingness to help me, this thesis would have, without a doubt, failed.

TABLE OF CONTENTS

List of Tables	iv
List of Figures	v
Abstract	vii
Introduction	1
Literature Review	5
Reliability Function	24
Examples	28
4.1 Example 1.....	28
4.2 Example 2	35
4.3 Example 3	38
4.4 Example 4	46
4.5 Example 5	52
Conclusion	55
Bibliography	59
Bibliography	61

List of Tables

Table Page

1. Example 1: Experimental Runs	29
2. Example 1: Estimates of the Responses	29
3. Example 1: Paths of Steepest Ascent	29
4. Example 1: 10 Highest Reliable Points	34
5. Example 2: True β	35
6. Example 2: X Matrix	36
7. Example 3: 2 Factors, 2 Responses Experiment	39
8. Example 3: Estimates for the Responses.....	39
9. Example 3: Paths of Steepest Ascent.....	40
10. Example 3: Path of Steepest Ascent and Descent.....	40
11. Example 3: 10 Highest Reliable Points.....	45
12. Example 4: True Values of the Parameter Estimates.....	46
13. Example 4: X Matrix.....	47
14. Example 5: Experimental Runs and Responses.....	53
15. Example 5: Estimates for the Responses.....	53
16. Example 5: 10 Highest Reliable Points.....	54

List of Figures

Figure Page

1. Example 1: The paths of steepest ascent	30
2. Example 1: Confidence cone for the first path of steepest ascent	31
3. Example 1: Confidence cone for the second path of steepest ascent	31
4. Example 1: Overlap of the confidence cone of the paths of steepest ascent.....	32
5. Example 1: Convex combinations of the paths of steepest ascent.....	32
6. Example 1: Pareto Optimal plot.....	33
7. Example 1: Highest reliable direction for the paths of steepest ascent.....	34
8. Example 2: Paths of steepest ascent.....	36
9. Example 2: Simulation of highest reliable points for error variance equal to 1.....	37
10. Example 2: Simulation highest reliable points for error variance equal to 1.....	38
11. Example 3: The two paths of steepest ascent.....	41
12. Example 3: Confidence cones for the two paths of steepest ascent.....	41
13. Example 3: Convex combinations of the two paths of steepest ascent.....	42
14. Example 3: Pareto Optimal plot for the two responses.....	43
15. Example 3: Possible different paths of steepest ascent.....	44
16. Example 3: Possible different convex combinations of the paths of steepest ascent	44
17. Example 3: Highest reliable point	46
18. Example 4: Paths of steepest ascent	48
19. Example 4: Convex combinations of the paths of steepest ascent	48

20. Example 4: Simulation highest reliable points for error variance equal to 1	50
21. Example 4: Simulation highest reliable points for error variance equal to 3	51

Abstract

MULTIVARIATE STEEPEST ASCENT USING THE BAYESIAN RELIABILITY FUNCTION

By Jeffrey Norman Fuerte, Jr. M.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2010.

Major Director: Dr. David Edwards, Assistant Professor, Department of Statistical Science and Operations Research

The path of steepest ascent can be used to optimize a response in an experiment, but problems can occur with multiple responses. Past approaches to this issue such as Del Castillo's overlap of confidence cones and Mee and Xiao's Pareto Optimality, have not considered the correlations of the responses or parameter uncertainty. We propose a new method using the Bayesian reliability to calculate this direction. We utilize this method with four examples: a 2 factor, 2-response experiment where the paths of steepest ascent are similar, ensuring our results match Del Castillo's and Mee and Xiao's; a 2 factor, 2-response experiment with disparate paths of steepest ascent illustrating the importance of the Bayesian reliability; two simulation examples, showing parameter uncertainty is considered; and a 5 factor, 2-response experiment proving this method is not dimensionally limited. With a Bayesian reliable point, a direction in multivariate steepest ascent can be found.

CHAPTER 1 Introduction

In 1951, Box and Wilson proposed a revolutionary new way of thinking about industrial experiments, which would later pave the way for response surface methodology. One usually views response surface methodology in the context of design of experiments (DOE), model fitting, and process optimization. By introducing a chronological approach to experiments, including screening, region seeking (as well as steepest ascent), product optimization and other concepts, Box and Wilson seamlessly opened a new research field, which is continuing to evolve today. The attention drawn to response surface methodology has been very intense in the last 20-25 years. Progress in this interval has surpassed that of the previous 20-25 years, but much of this had to do with computing capabilities lagging behind until the 1980's. The approach described by Box and Wilson involves a set of mathematical techniques consisting of designed experiments and basic optimization. An experimenter is trying to find the optimal conditions for a set of input factors by using the designed experiment and analyzing the results. By controlling some of the factors, the experimenter is able to manipulate the other inputs of interest to determine which combination of factors will yield the preferred result. The experimenter is usually trying to either maximize or minimize this response. For instance, a cost response and a yield response could be studied for the purpose of finding optimum operating conditions that reduce the cost and increase the yield simultaneously.

Similarly, a chemist may be interested in what percentages of compounds produce the optimal solution, minimizing the acid content, while maximizing the base content.

In most response surface designs the true relationship between the response and the explanatory variables is not known. Because of this, the first step in response surface methodology is to find an adequate approximation of this relationship, called the first-order model, using estimates of the true parameters that define the true relationship. This model represents only the linear relationship between the data; thus, it contains only main effects. The least squares method is then implemented to approximate the parameters in the model and to screen the variables, eliminating the factors that are not significant. Usually, response surface methodology is a sequential approach; more than one experiment must be carried out, using the information gathered in the previous experiment. By using the first order model, one can guide the experimenter swiftly and efficiently along a path of improvement in the direction of the optimum, using a step approach. Thus, the first order model is represented by

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k \quad (1)$$

In this equation, \hat{y} is the estimated response variable, $\hat{\beta}_0$ is the estimate corresponding to the intercept, $\hat{\beta}_k$ is the estimate for the k^{th} factor, and x_i corresponds to the point of interest in the design region. Taking the partial derivatives with respect to the x_i s gives us:

$$\frac{d\hat{y}}{dx_i} = \hat{\beta}_i \quad i = 1, 2, \dots, k \quad (2)$$

This yields the direction of steepest ascent,

$$\lambda(\hat{\beta}_1, \dots, \hat{\beta}_k) \quad \lambda > 0 \quad (3)$$

Therefore, the points chosen along the path of steepest ascent are proportional to the first-order model parameter estimates. Once improvements in the response are no longer seen along the path of steepest ascent, the experimenter then performs another design, and repeats the path of steepest ascent if necessary.

This process is simple when there is only one response being optimized. The path is easy to obtain and easy to implement. However, in most situations there is more than one response that needs to be optimized. A simple approach to optimizing simultaneous responses is to build an appropriate response surface model for each response and then attempt to find a set of operating conditions that achieves the target value for each response, or at least maintains them in a desired range. One way the experimenter might determine this is to build a contour plot for each response and overlay them, deciding where the overlap of the contour plots occurs and setting the operating conditions within that overlap. The responses are guaranteed to be optimized based on the factor settings in the region where this overlap occurs. However, this approach is not appropriate for cases with more than three responses. An alternative to setting target values for each response is constrained optimization, and usually nonlinear programming techniques are utilized to decide the optima.

Several methods have been proposed to find the path of steepest ascent when there is more than one response, but most do not consider the correlation between the responses of future predictions. Therefore, instead of using past techniques, we propose a new way to examine multiple response steepest ascents using a Bayesian technique: the posterior predictive reliability function, which does take into account the correlations of future predictions, as well as parameter uncertainty. When looking at these two issues, it

is important to know why these can be harmful. Del Castillo (2007) explains further, stating that classical methods provide a single point estimate where the process should be run, but this neglects the variability of the parameter estimates. When a different experiment is run, it is possible that a different optimum could be obtained. Along the same lines, a confidence region may be created, but this region cannot be interpreted as a region that contains the optimum with some probability. This is merely interpreted in the classical sense of being a region that would result after repeated sampling and optimization. Also, one of the central problems in multivariate optimization is that the responses may be correlated. In the classical approach, the correlation can be considered when fitting the models, but not at the optimization step. The Bayesian predictive approach considers the whole multivariate distribution of the responses, so it takes into account the correlation of the responses at the optimization step. These two qualities are especially important in quality assessment. Understanding the variability of the responses has been stressed by Myers (1999) as a highly significant issue for practitioners.

Chapter 2 explores past literature concerning multiple response optimization, as well as multivariate steepest ascent. The reliability function and how it is applied is examined in Chapter 3. Chapter 4 looks at five examples using the reliability function to calculate the path of steepest ascent with multiple responses: two simulation examples and three practical examples. Lastly, Chapter 5 summarizes the conclusions from the examples and shows the reliability function's advantages and future research which can be done.

CHAPTER 2 Literature Review

In this section, we give a background to techniques that have been previously utilized for multiple response optimization. According to Myers and Montgomery (2002), the main technique to solve this problem is to overlay contour plots to locate the best region to operate is for all responses. Contour plots provide an illustration of the behavior of the multivariate system, which can provide valuable information and added insight into the problem of optimization. This technique works best when there are only two factors and two responses. However, as the number of factors and responses increases, contour plots become extremely hard to read, and misinterpretation becomes increasingly likely. Also, contour plots graphically represent point estimates of the response value. Random sampling error is not represented in the plot, and this problem increases exponentially as the number of responses increases. As a result, other methods must be considered. To combat the problem of dimensionality, researchers have employed various approaches.

When there are only two responses, Myers and Carter (1973) introduced the dual responses approach, where two responses are classified as primary and secondary. This approach attempts to maximize (or minimize) the primary response while placing a constraint on the secondary response. That is,

$$\begin{aligned} & \text{Min (or Max)} y_{primary} \\ & \text{Subject to: } y_{secondary} = \varepsilon \end{aligned} \tag{4}$$

where ε is a specified value.

Various values of the secondary response are considered and the best operating condition is the one where the primary response is maximized (or minimized) based on those selected secondary responses. Biles (1975) generalized this procedure, optimizing a primary response function while keeping secondary responses in specified ranges. His procedure employs a modified version of Box and Wilson's original steepest ascent. On the other hand, by constraining the secondary response, the experimenter might miss other possible operating conditions that are superior but were not considered; thus, a true optimum may not be found.

Response surface optimization procedures use the estimates of the gradient of the surface with respect to each of the control variables. The path of steepest ascent is a function of the measurement scales employed for each of the factors. What the experimenter selects for the ranges of each control variable vastly affects the estimates and thus the path of steepest ascent. Because of this, much discussion has been documented on what scales to use in order to predict future response. Heller and Staats (1973) dubbed their method "cheapest ascent" to factor in cost of future measurement, as well as time. They describe their method as follows: 1) running experiments at every point on a multidimensional grid, with the range determined by the cost of experimentation, estimates of the changes in operating conditions, and constraints on the variables or responses, 2) the resulting data are analyzed to locate the grid point having the best performance (maximizing or minimizing the response), and 3) if this point is on the boundary of the experimental grid, and if the estimated gradient is high enough, further experiments may be run; otherwise, the inputs are set to this point. This method

does not take into account the uncertainty of the parameter estimates, and thus could lead to incorrect future predictions.

Harrington (1965) introduced an analytic technique based on a desirability function. Each response is transformed to a desirability value, f , ranging between 0 and 1, such that as f increases, the desirability of the corresponding response increases. Taking the individual desirability values and combining them into one value, F (the geometric mean), Harrington gathered an overall measure for the quality of the system. This reduces the multivariate problem into a univariate one.

Derringer and Suich (1980) extended this approach by introducing more general transformations of the responses into desirability values. Derringer and Suich's method allows for the use of a geometric mean of desirabilities that are defined according to power functions that are based on the impact priorities of the experiments regarding which specifications on the responses are the "tightest." They propose converting each response variable into an individual desirability function that varies from zero to one where:

$$D = \begin{cases} 0 & \text{if response is outside of target} \\ 1 & \text{if response is at target} \end{cases} \quad (5)$$

Going further, depending on the response, one could have three potential desirability functions:

- Larger-the-better response:

$$d_i = \begin{cases} 0 & \text{if } \hat{y} < L \\ \left(\frac{\hat{y} - L}{H - L} \right)^s & \text{if } L \leq \hat{y} \leq H \\ 1 & \text{if } \hat{y} > H \end{cases} \quad (6)$$

- Smaller-the-better response:

$$\circ d_i = \begin{cases} 0 & \text{if } \hat{y} > H \\ \left(\frac{H - \hat{y}}{H - L}\right)^s & \text{if } L \leq \hat{y} \leq H \\ 1 & \text{if } \hat{y} < L \end{cases} \quad (7)$$

- Target is best:

$$\circ d_i = \begin{cases} 0 & \text{if } \hat{y} > H \\ \left(\frac{\hat{y} - L}{T - L}\right)^s & \text{if } L \leq \hat{y} \leq T \\ \left(\frac{H - \hat{y}}{H - T}\right)^s & \text{if } T \leq \hat{y} \leq H \\ 0 & \text{if } \hat{y} < L \\ 1 & \text{if } \hat{y} = T \end{cases} \quad (8)$$

In these equations, L is the lowest possible accepted response, H corresponds to the highest possible accepted response, and T is the target value. The desirability function becomes linear when $s=1$. When $s > 1$, more emphasis is placed on being close to the target value, whereas when $0 < s < 1$, less importance is stressed on being near the target value. The individual desirability scores are then combined into an "overall" desirability:

$$D = \left(\prod_{i=1}^m \hat{d}_i \right)^{1/m} \quad (9)$$

The design variables are then chosen to maximize this overall desirability. This technique involves a compromise between important responses, as an experimenter is usually not able to maximize the desirability function without the cost of some of the other responses. Some advantages of this approach are that optimization becomes simpler after the transformations, making the functions flexible, and theoretical optimal solutions can be stated. On the other hand, disadvantages are present, such as the statistical properties of the desirability function are unknown, and the desirability model

used in the approach is assumed to be correct. Also, Myers (1999) states another disadvantage concerning desirability functions: "It is tempting to treat the optima or constrained optima as if they were based on deterministic functions. One must always remember that optima are stochastic in nature and use of [any desirability function] should be followed by rather extensive confirmatory experiments." He warns that all response surface models involve predicted values that have considerable variance, which causes the optima to have unknown variance. Another disadvantage is that the importance of each response is imperative to know. If an experimenter makes an incorrect judgment on the values in the desirability function, this could lead to poor predictions in the future. Along these same lines, this approach does not take into account the variance and correlation structures of the responses. By ignoring these correlations, the experimenter alters the overall desirability, which may harm the determination of optimum operating conditions.

Khuri and Conlon (1981) utilize polynomial regression functions to handle the multivariate optimization problem. They assume that all response functions depend on the input variables, and these functions can be represented by polynomial regression models of the same degree within a certain region of interest. By removing linear dependencies first, and then obtaining individual optima of the estimated responses, an "ideal" optimum can be set where all individual optima are achieved. Using a distance function, the deviation from the "ideal" optimum can be calculated, and ultimately minimized. This minimized distance function, expressible in terms of the estimated responses and their covariance structure, can arrive at a set of operating conditions suitable for a "compromised" optimum. Vining (1998) later established that the Khuri

and Conlon procedure is a special case of a weighted squared error loss function, and showed several other plausible weighting schemes. However, the uncertainty of the parameter estimates is not considered, and this could lead to future predictions that are not suitable.

Another approach is to use loss functions to provide a conceptual framework for combining different criteria into a single objective function. This represents the "total loss to society" from departures of a vector of criteria from their target values. Taguchi (1986) introduced the loss function to model the concept that any departure from intended targets causes economic loss. This loss function describes the loss arising from deviations from a target as a result of random variation and systematic errors. The total loss to society due to deviation in all the product characteristics may be taken as the sum of those due to the individual characteristics. The first designed experiments using loss functions for multivariate optimization were conducted in 1984. Since then, well over 500 sets of loss functions have been undertaken for both process development and manufacturing support. For very difficult problems, loss functions have proven to be powerful and flexible, because they provide a compact unified approach to compromising even large numbers with conflicting objectives. The global quality loss function is defined as follows:

$$GQL = \sum_{r=1}^{R_g} W_r (V_r - T_r)^2 \quad (10)$$

where V_r is a random variable that measured the response at criteria r , T_r is the target value of V_r , and W_r is the weight factors which scale the importance of the different criteria. The sum is taken over all the responses included in system R_g . This formula

describes the weighted distance of the measured response from the target at each criteria, r . By taking the sum over all the responses, we get an overall loss function.

Ames et Al. (1997) recommends another loss function which includes only response polynomials and targets and does not include random errors because the errors are small compared to the errors in hitting the targets. His equation models V_r as a function of the inputs:

$$QLP = \sum_{r=1}^R W_r (Y_r(X_1, X_2 \dots) - T_r)^2. \quad (11)$$

Minimizing these functions with respect to process inputs locates the best operating conditions. However, this method requires the experimenter to compromise by putting all the multiple responses into one function. Sometimes this cannot be done without causing a detriment to the responses so other methods need to be considered.

Tang and Xu (2002) look at the approach of a dual response optimization in terms of the mean and the variance. In the past, two models have been formulated, one for the mean and one for the standard deviation or variance:

$$\hat{y}_\mu = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i=1}^k a_{ii} x_i^2 + \sum_{i < j}^k a_{ij} x_i x_j \quad (12)$$

$$\hat{y}_\sigma = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k b_{ii} x_i^2 + \sum_{i < j}^k b_{ij} x_i x_j \quad (13)$$

These equations are then optimized simultaneously in a region of interest. However,

Tang and Xu propose a unified formula:

$$\begin{aligned} & \text{Min } \delta_\mu^2 + \delta_\sigma^2 \\ & \text{subject to: } \hat{y}_\mu(x) - \omega_\mu \delta_\mu = T_\mu^* \\ & \quad \hat{y}_\sigma(x) - \omega_\sigma \delta_\sigma = T_\sigma^* \\ & \text{and either } x' x \leq r^2 \text{ or } x_l \leq x \leq x_u \end{aligned} \quad (14)$$

The objective function is the function to be minimized, where δ_μ and δ_σ are unrestricted scalar variables, so these can be positive or negative, including the mean and variance of the response variable, ω_μ , ω_σ (≥ 0) are user defined weights, and T^* is the ideal response mean and standard deviation associated with a set of response functions. Using this unified formula, Tang and Xu attempt to minimize this function. However, this approach only considers the mean and variance for the two responses. If the responses seeking to be optimized are not the mean and variance, then the formulation proposed by Tang and Xu does not optimize the responses.

The preceding techniques are for the problem of multivariate optimization when looking at the second order model i.e., the main effects, interactions, and quadratics that are significant. However, our proposed method takes the methodology a step back, merely looking at just the model with the main effects and interactions. Here we look at the research that has been involved in only steepest ascent and the applications to our method.

Del Castillo (1996) implements a method to deal with multiple response steepest ascent, proposing the use of confidence regions (if the responses are quadratic) or confidence cones (if the responses are linear) to analyze multiresponse processes. The formula for finding the confidence region is defined as:

$$\sum_{i=1}^k b_i^2 - \frac{\left(\sum_{i=1}^k b_i X_i \right)^2}{k-1 \sum_{i=1}^k X_i^2} \leq s_b^2 F_{\alpha, k-1, v_b} \quad (15)$$

where $s_b^2 = SS_{\text{error}} C_{jj} / (n-p)$, C_{jj} is the j^{th} diagonal element of the $(X'X)^{-1}$, and X is the model matrix of the experiment. The direction the formula generates lies within the 100(1- α)% confidence cone of steepest ascent if

$$\sum_{i=1}^k b_i x_i > 0 \quad (16)$$

or inside the 100(1- α) confidence cone of steepest descent if

$$\sum_{i=1}^k b_i x_i < 0 \quad (17)$$

The fraction of directions excluded by the confidence cone (see Box and Draper (1987)) is given by:

$$1 - \phi = 1 - T_{k-1} \left(\frac{\sum_{i=1}^k b_i^2}{s_b^2 F_{\alpha, k-1, n-p}} - (k-1) \right)^{1/2} \quad (18)$$

where T_{k-1} denotes the Student's t distribution with $k-1$ degrees of freedom.

The confidence region is a cone when there are two or three factors and a hyper cone if more than three variables with the apex at the design origin and all points a unit distance from the origin satisfying the inequality. By first finding a confidence region around the direction that would maximize the responses, rather than just a path, the experimenter is able to decide where the cones overlap. This overlap represents the compromise direction in which the experimenter can proceed, knowing that the experimenter is achieving the goal of optimizing the responses. The experimenter can then operate the sequential experiments in that overlap of the optimum of the responses and continue analyzing as if there was only one response. As described by Del Castillo, his "approach consists in finding operating points x that simultaneously satisfy

constraints [of the response variables]." Once the best direction is found, sequential experimentation is continued and the process may be repeated if need be. Del Castillo (1993) utilizes nonlinear programming techniques to determine operating conditions that are in the overlap of the confidence cones of the responses. Del Castillo proposes solving the nonlinear programming problem:

$$\text{Maximize } \|x - x_c\| \quad (19)$$

$$\text{Subject to } \left[\sum_{i=1}^k b_i^2 - \frac{\left(\sum_{i=1}^k b_i x_i \right)^2}{\sum_{i=1}^k x_i^2} \right] \leq (k-1) s_{b_j}^2 F_{\alpha_j, k-1, n-p} \quad j \in L_M \cup L_m \quad (20)$$

$$\left[\sum_{i=1}^k b_i x_i \right]_j \leq 0 \quad j \in L_m \quad (21)$$

$$\left[\sum_{i=1}^k b_i x_i \right]_j \geq 0 \quad j \in L_M \quad (22)$$

$$x'x \leq \rho^2 \quad (\text{if a spherical region}) \quad (23)$$

$$x_l \leq x \leq x_u \quad (\text{if a cuboidal region}) \quad (24)$$

where x_c represents the sample point and x_l and x_u denote vectors of lower and upper bounds respectively. This model attempts to find the farthest point from the current operating point x_0 such that it lies within the $100(1-\alpha)\%$ confidence cones of maximum improvement of all responses j and within the experimental region. Del Castillo recommends using $0.01 \leq \alpha \leq 0.1$ for the "primary" responses and $0.001 \leq \alpha \leq 0.01$ for the "secondary" responses, providing wider confidence cones for the secondary responses. If the equations above have no feasible solution, this indicates the experiment contains disparate paths of steepest ascent. When this happens Del Castillo recommends using a different non linear program which includes some subjective considerations and

process knowledge to define priorities among the responses. For example, if a subset of responses P are most important, Del Castillo advises using:

$$\text{Maximize } \frac{\sum_{j=1}^r (1-\phi_j)x'd_j}{k - \sum_{i=j}^r \phi_j} \quad (25)$$

$$\text{Subject to } \left[\frac{\sum_{i=1}^k b_i^2 - \frac{\left(\sum_{i=1}^k b_i x_i \right)^2}{\sum_{i=1}^k x_i^2}}{\sum_{i=1}^k b_i^2} \right] \leq (k-1) s_{b_j}^2 F_{\alpha_j, k-1, n-p} \quad j \in L_M \cup L_m \quad (26)$$

$$\left[\frac{\sum_{i=1}^k b_i x_i}{\sum_{i=1}^k x_i^2} \right]_j \leq 0 \quad (\text{if } p \in P \text{ must be a min.}) \quad (27)$$

$$\left[\frac{\sum_{i=1}^k b_i x_i}{\sum_{i=1}^k x_i^2} \right]_j \geq 0 \quad (\text{if } p \in P \text{ must be a max.}) \quad (28)$$

$$x'x \leq \rho^2 \quad (\text{if a spherical region}) \quad (29)$$

$$x_l \leq x \leq x_u \quad (\text{if a cuboidal region}) \quad (30)$$

where d_1, d_2, \dots, d_r represent the estimated direction of maximum improvement for each linear response. A weighted sum of the cosines of the angles between the desired vector x and each individual direction of maximum improvement is given in (25). The weights are given by the fraction of directions excluded by each confidence cone. In other words, the solution to the system of equations is a compromise of the paths of steepest ascent, weighted by the percentage of directions excluded by each cone. More weight is given to the primary responses as their cones are smaller and exclude more directions. Therefore, the result is inside the primary responses' confidence cones.

Del Castillo's method is adequate for finding a compromise path of steepest ascent when the directions are not disparate. Del Castillo finds the path that is the best compromise, especially when the cones overlap. However, if we consider disparate directions, the chosen path will be closer to the path for the "primary" response. What if the best compromise path for the true optimum lies outside the "primary" responses' confidence cone? This may present problems for future predictions.

Mee and Xiao (2008) (referred to from now on as MX) suggest a different method when tackling the problem of multivariate steepest ascent. MX illustrate their method on an example with three responses and five factors, focusing on identifying useful compromise directions, especially for cases where the paths of steepest ascent are widely disparate. The example the authors choose to motivate their method typifies this type of discrepancy. First, MX prove that when considering compromise directions only convex combinations of the paths of steepest ascent should be considered:

Theorem 1. Every nonnegative linear combination x_c is Pareto Optimal among the set of vectors $\|x\| \leq \|x_c\|$.

The theorem infers that MX only considered paths that compromised all responses; thus, only the compromised path of steepest ascent that lies in the angle created by the original paths of steepest ascent were considered. Considering only convex combinations of the paths allows the researchers to construct several graphs that make the choice of a suitable compromise direction easy.

To find the angle between two paths of steepest ascent, Mee and Xiao employ:

$$\cos(\theta_{b_j, b_{j'}}) = \frac{(b_j' b_{j'})}{\sqrt{(b_j' b_j)(b_{j'}' b_{j'})}}, \quad (31)$$

Where b_j represents the vector of $\hat{\beta}'$'s for the j^{th} response and b_j represents the $\hat{\beta}'$'s for the j^{th} response. Formula (31) directly relates to the correlation between the predicted values for the two responses at the design points. In fact, for any coded design D such that $1'D=0$ and $D'D=S_{xx}I$, the cosine and the correlation will be equal. When calculating the cosine of the angle between the paths of steepest ascent, any angle that is greater than 90° will be disparate enough that the compromise direction will not be within the confidence cones that Del Castillo proposed.

Mee and Xiao use Pareto Optimal points to locate the desired compromise path of steepest ascent. To calculate a Pareto Optimal point, MX considered only convex combinations of the paths. First, MX let $\hat{y}(x) = (\hat{y}_1(x), \dots, \hat{y}_j(x))'$ for any number of j responses, be the predicted responses at a given vector x and define the norm $\|x\| = (x'x)^{1/2}$. The vector $\hat{y}(x_{PO})$ is said to be Pareto Optimal if for every vector x such that $\|x\| \leq \|x_{PO}\|$ and $\hat{y}(x) \neq \hat{y}(x_{PO})$, there exists a j such that $\hat{y}_j(x_{PO})$ is preferred over $\hat{y}_j(x)$. MX also proved that every nonnegative linear combination x_c is PO among the set of vectors $\|x\| \leq \|x_c\|$. Thus, by only considering convex combinations of the paths of steepest ascent, a simplification is made and the search is lessened greatly. Once the PO points are calculated, the experimenter can create a Pareto Optimal plot, which includes all of the PO points a set radius from the design center, as well as the predicted values for the factorial design points. In the PO plot, which can only be used when looking at two responses, one response is on the x-axis and the other response is on the y-axis. In the plot, each point represents the x-values plugged into the

estimated first-order model. This gives the experimenter an idea of where the best trade off for the responses occurs. An example of this is given later in the thesis.

Similarly, a “paths of improvement” region is found, and the set of vectors x , such that x is contained in the paths of improvement region for all j responses is

$\square_c = \{x : x'T / \|x\| > (v_{\alpha_1}, \dots, v_{\alpha_j})\}$, where T denotes $k \times J$ matrix of t statistics; i.e. each element of β is divided by its standard error. From this equation, the intersection of j cones is determined as the solution to a system of linear inequalities.

However, because only convex combinations are considered, the researchers ignore all other points. This is important because in doing so the researchers ignore the variances of the predictions because these variances could be high enough that the true path is not a convex combination. By not including the convex combinations of the paths of steepest ascent in their search, valuable information could be missed and a better potential “compromise” path of steepest ascent may not be found.

Peterson (2004) makes the claim that a new method for multiple response optimization must be used because most past methods do not consider the correlations among the responses and the variability of the predictions. Also, most approaches do not factor in the uncertainty of the estimates of the model parameters. The reason these are important is for quality assessment. Peterson makes an effort to investigate past approaches and prove why each method cannot be used to accurately make future predictions. Peterson goes further to discuss why loss functions should not be used: “These quadratic loss function methods ... do not take into account the uncertainty of the variance-covariance matrix of the regression model error.” Similarly, Peterson states that quadratic loss functions are difficult for practitioners to grasp. Del Castillo’s method is

considered inadequate as a result of having to state “primary” responses and “secondary” responses, as well as not addressing the correlations among the responses. Although Peterson goes on to praise Chiao and Hamada’s approach for including the variance-covariance structure of the data and simplicity of interpretation, he is quick to state that uncertainty of the parameters is not considered. As a result, Peterson proposes a Bayesian reliability approach which takes into account the correlation structure of the data, the variability of the process distribution and the model parameter uncertainty. The reliability function utilizes the posterior predictive distribution of the multivariate response to compute the probability that a future multivariate response will satisfy specified quality conditions.

First, Peterson lets $Y = (Y_1, \dots, Y_p)'$ be the multivariate ($p \times 1$) response vector and $x = (x_1, \dots, x_k)'$ be the ($k \times 1$) vector of factor variables. Creating the standard regression model, we have:

$$Y = Bz(x) + e, \quad (32)$$

Where B is a $p \times q$ matrix of regression coefficients and $z(x)$ is a $q \times 1$ vector for x . The error is distributed as a multivariate normal with mean vector 0 and variance-covariance matrix Σ . To account for the uncertainty in the model parameters, B and Σ , the posterior predictive density can be used. Using the conventional noninformative joint prior for B and Σ and the model in (32), the Bayesian predictive density for Y given B, Σ and the data can be attained. Because B is proportional to a constant and Σ is proportional to $|\Sigma|^{-(p+1)/2}$, the joint prior for B and Σ is relative to $|\Sigma|^{-(p+1)/2}$.

From here, we can sample from the multivariate t-distribution because the Bayesian predictive density for a specified x -value has this distribution with ν degrees of

freedom, where $\nu=n-p-q+1$ and n is the sample size. Therefore, using the joint prior, as well as the joint likelihood for X , B , and Σ :

The posterior predictive distribution is:

$$f(y | x, data) = c \left\{ 1 + \frac{1}{\nu} (y - \hat{\beta}z(x))' H (y - \hat{\beta}z(x)) \right\}^{-(p+\nu)/2} \quad (33)$$

where

$$c = \frac{\Gamma(\frac{p+\nu}{2})\sqrt{|H|}}{\Gamma(\nu/2)(\pi\nu)^{p/2}}; \quad H = \frac{\nu V^{-1}}{1 + z(x)' D z(x)'} \\ D = \sum_{i=1}^n z(x_i) z(x_i)'; \quad V = (y^* - (\hat{\beta}Z)')'(y^* - (\hat{\beta}Z)')$$

Z is the $q \times n$ matrix formed by the $z(x_i)$ covariate vectors.

Peterson easily simulates Y -values from this distribution by simulating a multivariate normal random variable and an independent chi-square random variable:

$$Y_j = (\sqrt{\nu}W_j / \sqrt{U}) + \hat{\mu}_j, \text{ for } j = 1, \dots, p \\ \text{where } W \sim MVN(0, H^{-1}) \text{ and } U \sim X^2(\nu) \quad (34) \\ \text{and } \hat{\mu}_j \text{ is the } j^{\text{th}} \text{ element of } \hat{\mu} = \hat{B}z(x).$$

After sampling from this distribution, we would want to maximize the probability, $p(x)$, such that:

$$\begin{aligned} P(D(Y) \geq D^* | x) \\ P(Q(Y) \leq Q^* | x) \end{aligned} \quad (35)$$

where D^* and Q^* are chosen by experimenter, as the desirability function and the quadratic loss function, respectively. Let x_0 be such that it maximizes $p(x)$ over the experimental region; we know that if $p(x_0)$ is sufficiently large, x_0 will provide operating

conditions for future responses which have a high probability of satisfying desired conditions. Peterson suggests having a product expert or a team of experts, deciding on what D^* and Q^* should be, depending on what response levels develop a “good” product. Choosing a specific point around the design region, one can then find out what the Bayesian reliability is, or a variety of x-points can be tested to see if further improvements in reliability can be obtained. If the reliability is large, the experimenter can feel comfortable that future predictions will produce good responses with a high degree of likelihood. Peterson then stresses that validation runs should be performed to double check that the statistical model for the optimal factor conditions holds. Remedial work is necessary if the reliabilities are not large. Increasing the sample size will allow the experimenter’s reliability to sufficiently increase when reducing process variation does not do so.

Similarly, one can modify and simulate from the posterior predictive distribution to add more data points. This can be done by increasing the rows of the design matrix and changing the degrees of freedom accordingly. This will give the experimenter an idea of how much the reliability can be increased by reducing model uncertainty. These ideas are akin to the “preposterior” analysis described by Raiffa and Schlaiffer (2000).

Using simulation, one can approximate the reliability $p(x)$ for various x-values in the experimental region using:

$$p(x) \approx \frac{1}{N} \sum_{s=1}^N I(C(Y_s) \in S), \quad (36)$$

where N is the number of simulations, I represents the indicator function that meets the criteria of the predicted y value, Y_s , is a member of the region of interest, S , which contains the target response values. Therefore, our interest is in $C(y) \in S$. With a small

number of factors, it is computationally reasonable to grid over the entire experimental region to calculate the values of $p(x)$. However, when $k \geq 3$ the researcher may want to have a more efficient approach to maximizing $p(x)$ using general optimization methods. These approaches include methods discussed in Nelder-Mead (1964) or Chatterjee, Laudato and Lynch (1996).

If reliable results are not found then one should look at a finer grid over a sub-region of the experimental model, along with another logistic model and more simulations. A better fit is expected as a smaller response surface is explored. This can be repeated as necessary.

One can look at the reliabilities for only one response. In other words, because $p(x)$ is a joint probability over all the responses, marginal probabilities can be easily computed using the equation:

$$p_i(x) = P(Y_i \in A_i | x, data), \quad (37)$$

where A_i is an interval where the experimenter desires Y_i to be. The interval can be one or two sided. The researcher can then monitor both $p_i(x)$ and $p(x)$ in chorus, so that by modifying A and A_i , he can observe the economic impact on changing the criteria. This can be used for other marginal (or joint, when looking at more than one response) probabilities as well.

A Bayesian credible region can be calculated for process ruggedness assessment in addition. This can be achieved by plotting out all the x -values for which $p(x)$ is at least some probability (for instance .95).

Peterson's method assumes the correct model has been chosen already, so he is looking at the second order model, including all main effects, interactions, and quadratic

terms. Peterson is trying to optimize the response after this second-order model has been fit. In our own research, we will backtrack one step and use the posterior predictive distribution to compute steepest ascent reliabilities.

For this paper, we use the posterior predictive distribution to find the optimal compromise path of steepest ascent for multiple responses. Essentially, we use a grid search, sampling around the design region to determine the path with the highest Bayesian reliability. By determining the highest reliability we hope that the best compromise path of steepest ascent will be found. This method is most useful if there are multiple responses that have disparate paths of steepest ascent and there is no overlap between the confidence cones of the individual paths of steepest ascent. Using our method, we also hope that when the confidence cones overlap our highest reliable point is not only within this overlap but also a convex combination of the paths of steepest ascent as well.

CHAPTER 3 Reliability Function

Based on previous methods, it is clear that many issues need to be considered when considering multiple responses. Del Castillo's method may be used if the paths are close together and their confidence regions intersect. However, neither his method nor Mee and Xiao's method consider the correlation among the responses and the variability of the predictions gathered from the paths of steepest ascent. This is a critical issue for quality assessment. Similarly, neither of these methods takes into account the uncertainty of the model parameters estimates. Peterson takes a Bayesian approach by using the reliability function to assess the accuracy of the first order model.

The approach suggested in this paper takes into account the correlation of the responses, the variability of the predictions and the uncertainty of the estimates of the model parameters by implementing the method described by Peterson. The standard regression model for multiple response optimization is:

$$y = \beta_0 + \beta_1(x_1) + \beta_2(x_2) + \dots + \beta_k(x_k) + \varepsilon \quad (38)$$

where $\varepsilon \sim MN(0, \Sigma)$

In order to implement the method of steepest ascent, we first need to estimate $\hat{\beta}$, according to the formula, where X is the model matrix of interest with the main effects only:

$$\hat{\beta} = (X'X)^{-1}X'y \quad (39)$$

To account for the uncertainty in the model parameters, the posterior predictive density function $f(y|x, data)$ can be used. The Bayesian predictive density for a specified x -value has the multivariate t distribution, with $\gamma = n - p - q + 1$ degrees of freedom, where n is the sample size, p is the number of factors and q is the number of responses. By sampling from this distribution for a given point outside of our design region, we can obtain a probability and compare this probability to other sampled points. The point with the highest probability is the direction in which we move for a compromise path of steepest ascent. Therefore, we need to perform a grid search around our design region with a given radius. We perform this search by sampling from a multivariate t distribution, with mean 0 , variance H^{-1} and degrees of freedom γ , where:

$$H = \frac{\gamma(V)^{-1}}{1 + x_i' D x_i}$$

$$V = (y - X\hat{\beta})'(y - X\hat{\beta})$$

$$D = X' X$$

We can sample from this distribution using the R statistical program, with the `rmvt` package, specifying the mean, variance and degrees of freedom. Sampling from this distribution and adding to it $x_i * \hat{\beta}$ (our mean) gives our simulated y . From here we answer the question: does this y meet our criteria for the response? We can find a probability based on the answer by running a large number of trials. For each trial, we sought to determine if the sampled point produced \hat{y} 's that were within our desired ranges. We sampled from the posterior predictive distribution 1000 times to find a probability of meeting every response's goal. But, the sample size can be altered and a higher sample yields a more precise probability. Searching around a radius of $\sqrt{2}$ (which

is from the equation $x^T x = r^2$, we implemented polar coordination transformation for two variables to grid around the surface of a circle. Thus, we used:

$$\begin{aligned} x_1 &= r * \cos(\theta) \\ x_2 &= r * \sin(\theta) \end{aligned} \quad (40)$$

where $r = \sqrt{2}$ and θ was the degree around the radius. For each degree, and thus each point, we found the probability:

$$p(x) = P(Y \in A), \quad (41)$$

where A represents the desired conditions for each response, Y.

The points were sorted from increasing to decreasing. The point with the highest reliability is the new compromise direction for the path of steepest ascent. When the number of factors exceeded two, a new method had to be used to grid search around the design region. With only two factors, we simulated around a circle, sampling points every one degree (0.017 radians).

To demonstrate this better, consider a small example with two factors and two responses. Suppose we want to maximize both responses at values greater than 3. To implement our method, we set up the radius of $\sqrt{2}$ and searched around the circle at only 4 points. The points searched were at $\theta=0, 90, 180, 270$. Therefore, to find the coordinates for each r and θ , equation (40) was used. From here, we can sample from the posterior predictive distribution to determine if the yields for both responses are greater than 3. If they are, the indicator function generates a 1. Performing this process 500 times, we can create a probability for each degree. Thus, a reliability can be obtained at each sampled point. Comparing the reliabilities, we can find the highest reliable point, which gives a direction to move for the compromised path of steepest ascent.

However, for more than two factors the run time for the grid search would grow exponentially. Therefore, Peterson's (1993) method for polar coordinate transformation was implemented to provide a better way to search around the design region. We need to obtain x in the form:

$$x = r\theta, \quad (42)$$

where $r^2 = x'x$, and θ is described for more than two dimensions below. The general transformation is given by:

$$\begin{aligned} \tilde{x}_1 &= r \sin a_1, \\ \tilde{x}_i &= r \left\{ \prod_{j=1}^{i-1} \cos a_j \right\} \sin a_i, \quad i = 2, \dots, k-1, \\ \tilde{x}_k &= r \prod_{j=1}^{k-1} \cos a_j. \end{aligned} \quad (43)$$

Here, \mathbf{a} is a $(k-1) \times 1$ vector of angles contained in the rectangular set

$$A = \mathbf{a} : \left\{ -\frac{\pi}{2} < a_i \leq \frac{\pi}{2} \quad (i = 1, \dots, k-2), \quad -\pi < a_{k-1} < \pi \right\}. \quad (44)$$

By grid searching over A , we can calculate the reliability function as before and find the highest reliabilities. Whichever path maximizes this reliability is the path of steepest ascent chosen. This approach was used in the case where $k \geq 3$, such as the example in 4.3. After grid searching around the design region, the reliability as described above can be found for the number of trials. Therefore we need to solve the system of equations:

$$\begin{aligned} \max_{r^2=x'x} p(x) \\ \text{where } p(x) \approx \frac{1}{N} \sum_{s=1}^N I(C(Y_s) \in S), \\ \text{and } C(Y_s) \text{ is the desirable region for all responses} \end{aligned} \quad (45)$$

The point with the highest reliability is the path to conduct further experiments.

CHAPTER 4 Examples

The following examples utilize the Bayesian reliability approach to steepest ascent. We provide four examples, each illustrating a different advantage from previous methods.

The first second example involves only two factors and two responses, and we employ Del Castillo's and Mee and Xiao's methods for finding steepest ascent, as well as our own method. In this case, the paths of steepest ascent confidence cones will overlap.

The second example involves a simulation, showing that our method produces results that match up with Del Castillo's and Mee and Xiao's methods. For the third second example the two paths are disparate enough that they do not overlap, so we will show the highest reliability approach in that case. The fourth example involves simulating data from known $\hat{\beta}$'s and altering the error variance to point out how the error variance can affect the predictions. The final example demonstrates that our method is not limited dimensionally using a five factor, two response example.

4.1 Example 1

The following example is used in *Response Surface Methodology* by Myers and Montgomery (2002). Two variables, time and temperature, influence two responses, conversion and activity. The goal is to maximize both conversion and activity. The experimental runs are given in Table 4.1.

Table 4.1 – Example 1: Experimental Runs

Intercept	X1	X2	Y1	Y2
1	-1	-1	74	69
1	1	-1	51	50
1	-1	1	88	78
1	1	1	70	90
1	0	0	81	60
1	0	0	75	60.4
1	0	0	76	59.1
1	0	0	83	60.6
1	0	0	80	60.8
1	0	0	91	58.9

This yields the estimates shown in Table 4.2, calculated from the formula

$$\hat{\beta} = (X'X)^{-1} X'y.$$

Table 4.2 – Example 1: Estimates of the Responses

$\hat{\beta}$	Y ₁	Y ₂
$\hat{\beta}_0$	76.90	64.68
$\hat{\beta}_1$	-10.25	-1.75
$\hat{\beta}_2$	8.25	12.25

From here, we can calculate the path of steepest ascent for both responses, shown in

Table 4.3, with Figure 4.1 illustrating the paths of steepest ascent.

Table 4.3 – Example 1: Paths of Steepest Ascent

Y	Step	X ₁	X ₂
Y ₁	1	-1	.804878
	2	-2	1.609756
	3	-3	2.414634
	4	-4	3.219512
Y ₂	1	-.1428571	0.95421
	2	-.2857143	1.90842
	3	-.4285714	2.86264
	4	-.5714286	3.81685

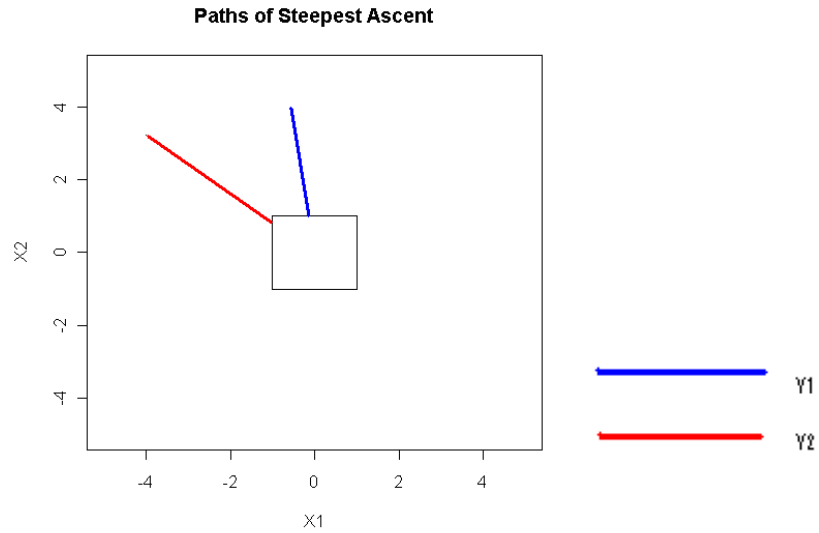


Figure 4.1 – Example 1: The paths of steepest ascent

Based on Del Castillo's method, we should observe where the two confidence regions lie to attempt to find a compromise path of steepest ascent/descent. Because we only have two factors, we can simplify the equation (15) into a simpler formula:

$$\theta = \arcsin \left(\frac{(k-1)s_b^2 F_{\alpha, k-1, v_2}}{\sum_{i=1}^k b_i^2} \right)^{1/2}, \quad (46)$$

where θ is the angle between the confidence cone and the path of steepest ascent (in other words, the margin of error). Computing this for each response, we have $\theta_{y_1} = 44.40023$ and $\theta_{y_2} = 60.13469$. Based on this result, we can now plot the confidence cones along with the paths of steepest ascent, with the first one shown in Figure 4.2.

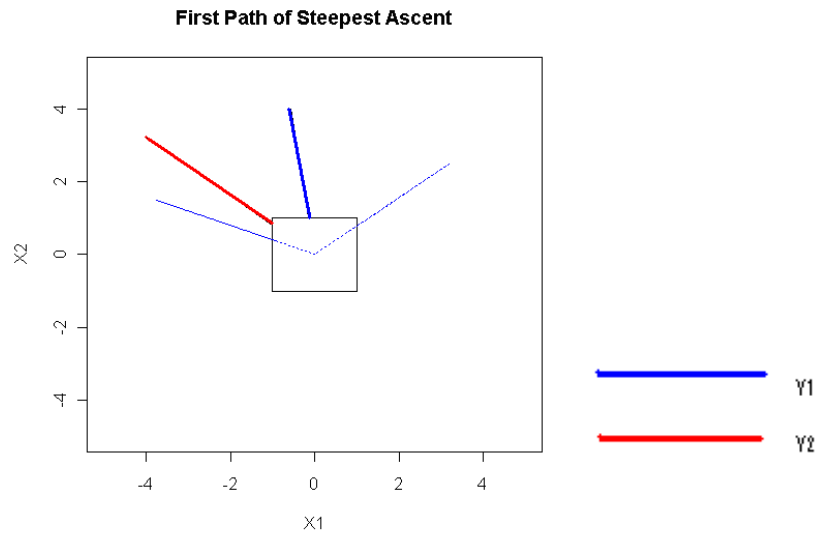


Figure 4.2 – Example 1: Confidence cone for the first path of steepest ascent

The second path and confidence cone is shown in Figure 4.3.

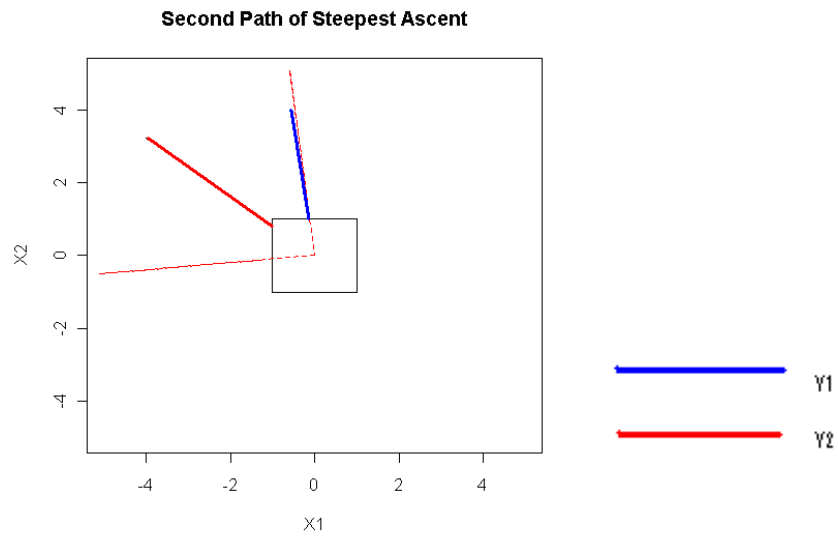


Figure 4.3 – Example 1: Confidence cone for the second path of steepest ascent

Therefore, Figure 4.4 shows where the overlap of the two confidence cones occurs (the shaded region).

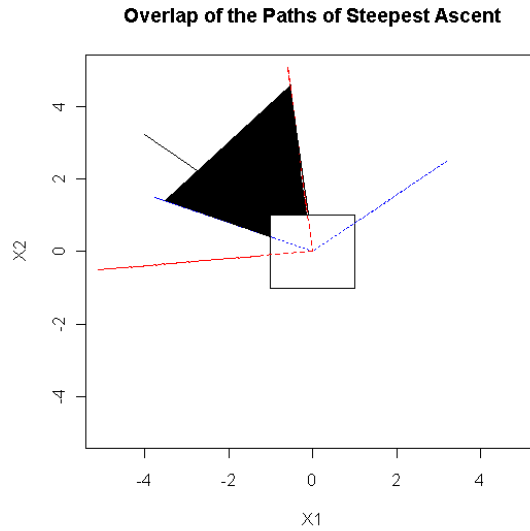


Figure 4.4 – Example 1: Overlap of the confidence cone of the paths of steepest ascent

Based on Mee and Xiao’s paper, we know that the Pareto Optimal points lie along convex combinations of the paths of steepest ascent, with this area shaded in Figure 4.5.

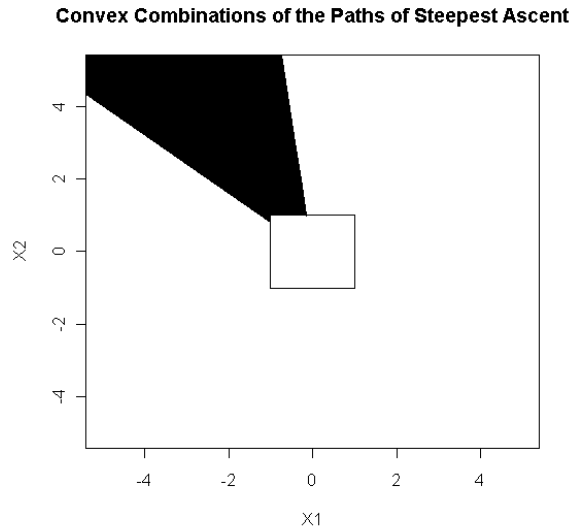


Figure 4.5 – Example 1: Convex combinations of the paths of steepest ascent

Each convex combination around the radius $\sqrt{2}$ is entered into equation (47), the first order model for each response to find the Pareto Optimal point:

$$\begin{aligned}\hat{y}_1 &= 76.9 - 10.25x_1 + 8.25x_2 \\ \hat{y}_2 &= 64.68 - 1.75x_1 + 12.255x_2\end{aligned}\tag{47}$$

We want to maximize both responses, so we should look for the Pareto Optimal point that performs this best. Looking at the Pareto Optimal plot in Figure 4.6, we can see where the points lie in relation to the design and choose the point that maximizes both y_1 and y_2 the best.

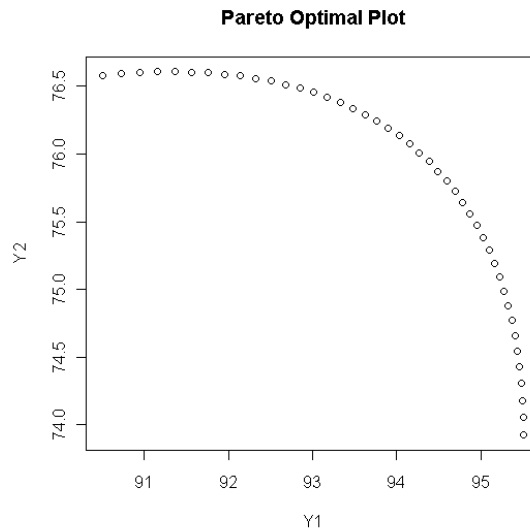


Figure 4.6 – Example 1: Pareto Optimal plot

The experimenter would calculate the best trade off at this point, and then move in that direction. However, we can calculate the Bayesian reliable point to find which direction we should move.

Utilizing our Bayesian reliability function, we consider all the points in the design region around the radius of $x'x=r^2$, or $\sqrt{2}$, to form our probability. We randomly sampled from the multivariate t distribution with mean 0, variance H^{-1} and degrees of freedom $\gamma=6$. From here we found the probability that a given point would create $y_1 \geq 86$ and $y_2 \geq 80$. These responses were arbitrarily chosen to represent the desired

minimum trade off between the two responses. The ten points with the highest probabilities are shown in Table 4.4.

Table 4.4 – Example 1: 10 Highest Reliable Points

X1	X2	Reliability
-0.770236	1.186059	0.394
-0.790818	1.172436	0.368
-0.851095	1.129441	0.368
-0.437016	1.344997	0.364
-0.749419	1.199321	0.364
-0.728374	1.212218	0.362
-0.663933	1.248676	0.36
-0.460423	1.337165	0.358
-0.318129	1.377967	0.356
-0.506809	1.320282	0.356

From here, the point with the largest reliability is the direction to move in, shown in Figure 4.7:

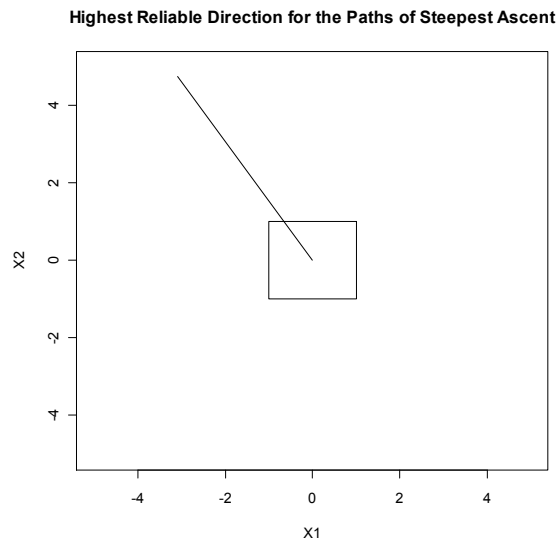


Figure 4.7 - Example 1: Highest reliable direction for the paths of steepest ascent

Plugging in the values for x_1 and x_2 into our first order model in equation (47), yielding the response values of

$$\hat{y}_1 = 76.9 + 10.25 * .77 + 8.25 * 1.186 = 94.577$$

$$\hat{y}_2 = 64.68 + 1.75 * .77 + 12.255 * 1.186 = 80.556$$

Both of the responses are above our cutoff values. What is interesting to note is that our method matches up with Del Castillo's and Mee and Xiao's methods. However, what is different from our method is that Del Castillo and Mee and Xiao propose approaches where the experimenter is forced to choose the point based on his knowledge of the system. On the other hand, our approach gives a specific point and direction for the experimenter who is not required to choose the point.

4.2 Example 2

For this example, we looked at a simulation, where we knew the true β s, as described in Table 4.5.

Table 4.5 – Example 2: True β

β	Y_1	Y_2
β_0	79.9	64.68
β_1	-10.25	-1.75
β_2	8.25	12.25

Using the X matrix described in Table 4.6, we are able to simulate data:

Table 4.6 – Example 4: X Matrix

Intercept	X ₁	X ₂
1	-1	-1
1	1	-1
1	-1	1
1	1	1
1	0	0
1	0	0
1	0	0
1	0	0

From here, we can see where the path of steepest ascent is in relation to our design region, shown in Figure 4.8:

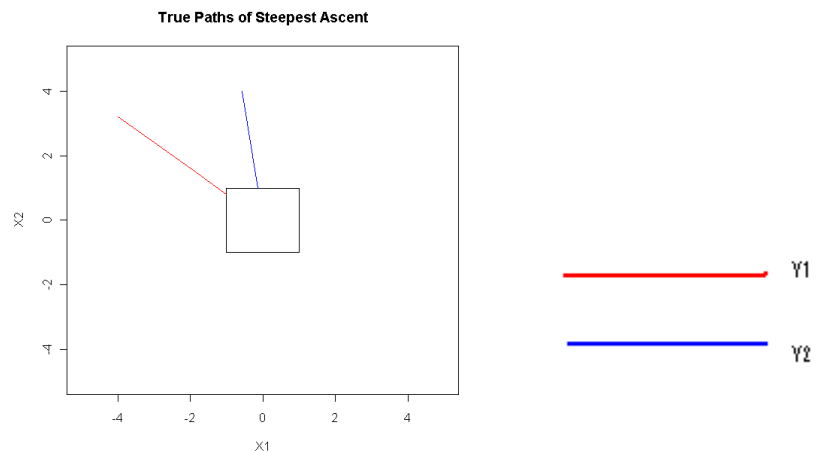


Figure 4.8 – Example 2: Paths of steepest ascent

Therefore, looking at the two paths, we know the Pareto Optimal point will be a convex combination of the two paths. From here, we can calculate our \hat{y} s from the equations:

$$\hat{y}_1 = 79.6 - 10.25x_1 + 8.25x_2$$

$$\hat{y}_2 = 64.68 - 1.75x_1 + 12.25x_2$$

To simulate data, we looked at 1000 trials for two different error variances from a random error vector. Thus, we chose to look at two different models for each response:

$$\begin{aligned} \hat{y}_1 &= 79.6 - 10.25x_1 + 8.25x_2 + \varepsilon \sim N(0,1) \\ \hat{y}_1 &= 79.6 - 10.25x_1 + 8.25x_2 + \varepsilon \sim N(0,3) \\ \hat{y}_2 &= 64.68 - 1.75x_1 + 12.25x_2 + \varepsilon \sim N(0,1) \\ \hat{y}_2 &= 64.68 - 1.75x_1 + 12.25x_2 + \varepsilon \sim N(0,3) \end{aligned}$$

We then ran through the same code as before, sampling from around the design region a radius of $\sqrt{2}$. To save computation time, we looked at every other degree. For each sampled point, we found a probability for $y_1 \geq 90$ and $y_2 \geq 75$ for $n=1000$. The values for the responses were arbitrarily chosen based on the first order models, with the responses optimized. The point with the highest reliability was selected and placed into a matrix. For all 1000 trials, there are 1000 points with each point being the highest reliability for that particular trial. Figure 4.9 shows these points graphically for when the random noise is distributed normally with a mean of 0 and a variance of 1.

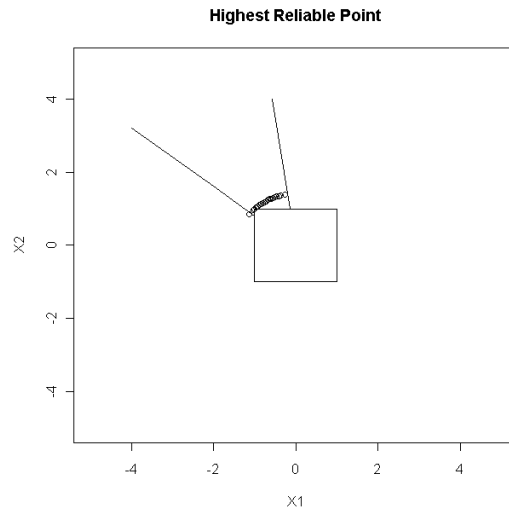


Figure 4.9 – Example 2: Simulation of highest reliable points for error variance equal to 1

The lines represent the true paths of steepest ascent. As we can see from the graph, the highest reliable points are in fact convex combinations of the true paths of steepest ascent, matching up with what Del Castillo and Mee and Xiao proposed.

Similarly, Figure 4.10 shows the highest reliable points when the error is normally distributed with a mean of 0 and a variance of 3:

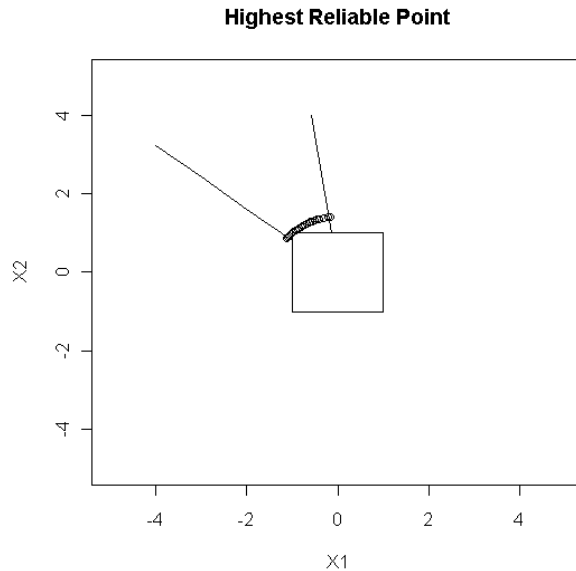


Figure 4.10 – Example 2: Simulation highest reliable points for error variance equal to 3

We can see that in both cases, almost every trial produces reliable points that are convex combinations of the true paths of steepest ascent. This matches up with past approaches and shows that our approach does produce similar results when the paths are not disparate.

4.3 Example 3

The following example was cited in Del Castillo's paper. Two factors, temperature and reaction time, influence the response, the yield of the process. Of interest is to measure both the mean yield (y_1) and the variance of the yield (y_2) based on the two factors.

Therefore, an experiment was conducted using a full factorial design for the two factors, along with five center runs. This is shown in Table 4.7.

Table 4.7 – Example 3: 2 Factors, 2 Responses Experiment

X_1	X_2	\hat{Y}_1	\hat{Y}_2
-1	-1	456.5	8.76
1	-1	595.6	21.96
-1	1	808.7	21.21
1	1	849	40.77
0	0	757.9	9.18
0	0	760.7	24.14
0	0	761.3	22.31
0	0	757.5	10.16
0	0	764.9	22.62

From this experiment, we are able to calculate $\hat{\beta}$ from the equation $\hat{\beta} = (X'X)^{-1}X'y$, shown in Table 4.8:

Table 4.8 – Example 3: Estimates For The Responses

$\hat{\beta}$	Y_1	Y_2
$\hat{\beta}_0$	723.567	20.1233
$\hat{\beta}_1$	44.85	8.19
$\hat{\beta}_2$	151.4	7.815

From there, we can use the $\hat{\beta}$'s to find our paths of steepest ascent. Calculating four steps along the path of steepest ascent yields the results shown in Table 4.9:

Table 4.9 – Example 3: Paths of Steepest Ascent

Y	Step	X ₁	X ₂
Y ₁	1	0.29624	1
	2	0.59247	2
	3	0.88871	3
	4	1.18494	4
Y ₂	1	1	0.95421
	2	2	1.90842
	3	3	2.86264
	4	4	3.81685

Suppose we wish to maximize the mean yield, \hat{y}_1 , and minimize the variation, \hat{y}_2 . Table 4.9 represents the two paths of steepest ascent. Because we want to minimize \hat{y}_2 , we need to calculate the path of steepest descent for \hat{y}_2 , we recalculate the paths as listed in Table 4.10:

Table 4.10 – Example 3: Path of Steepest Ascent and Descent

Y	Step	X ₁	X ₂
Y ₁	1	0.29624	1
	2	0.59247	2
	3	0.88871	3
	4	1.18494	4
Y ₂	1	-1	-0.95421
	2	-2	-1.90842
	3	-3	-2.86264
	4	-4	-3.81685

Figure 4.11 illustrates the two paths. We notice that the two paths are disparate.

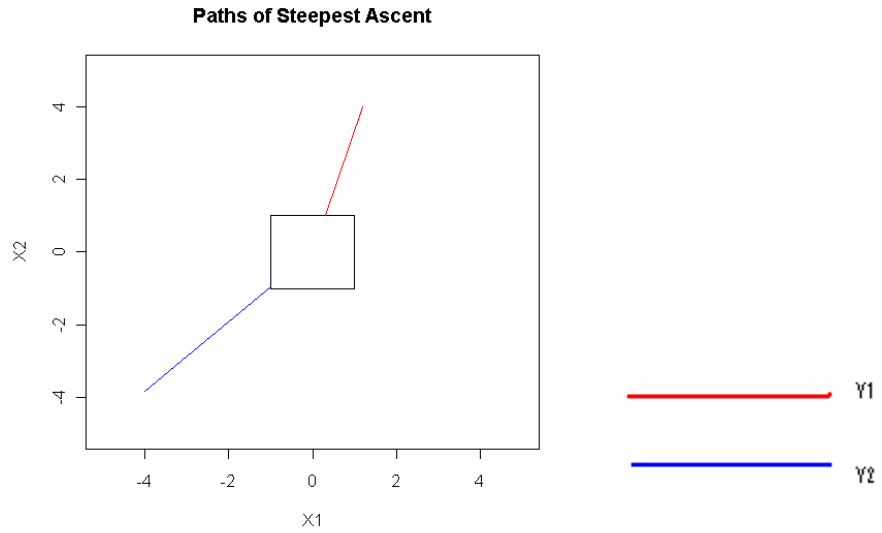


Figure 4.11 – Example 3: The two paths of steepest ascent

Using equation (46), we can calculate the angle between the confidence cone and the path of steepest ascent, yielding $\theta_{y_1} = 24.95363$ and $\theta_{y_2} = 49.10701$ for each response.

Calculating the confidence cones and adding them to the path of steepest ascent and the path of steepest descent, we can illustrate the disparity, shown in Figure 4.12.

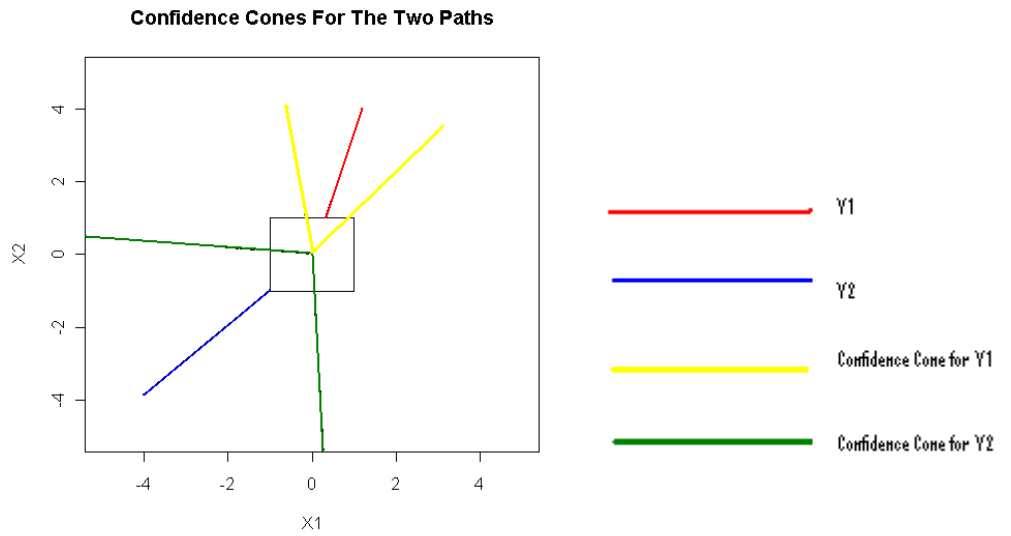


Figure 4.12 – Example 3: Confidence cones for the two paths of steepest ascent

From the figure, it is clear that the two cones do not intersect. Therefore, we need to find a compromise path of steepest ascent/descent. We can then utilize Mee and Xiao's method of finding the Pareto Optimal point. Based on this method, we should only consider points that are convex combinations of the two paths. Thus, the point should lie somewhere in the shaded region of Figure 4.13.

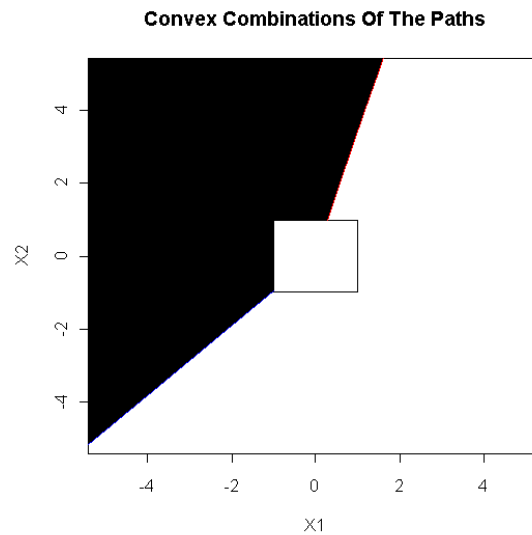


Figure 4.13 – Example 3: Convex combinations of the two paths of steepest ascent

Each convex combination around the radius $\sqrt{2}$ is entered into equation (48), the first order model for each response to find the Pareto Optimal point:

$$\begin{aligned} \hat{y}_1 &= 723.5667 + 44.85x_1 + 151.4x_2 \\ \hat{y}_2 &= 20.12333 + 8.19x_1 + 7.815x_2 \end{aligned} \quad (48)$$

Because we want to maximize the first response and minimize the second, we should look at where the optimal trade off occurs. Looking at the Pareto Optimal plot in Figure 4.14, we can see where the points lie in relation to the design and choose the point that maximizes y_1 and minimizes y_2 the best.

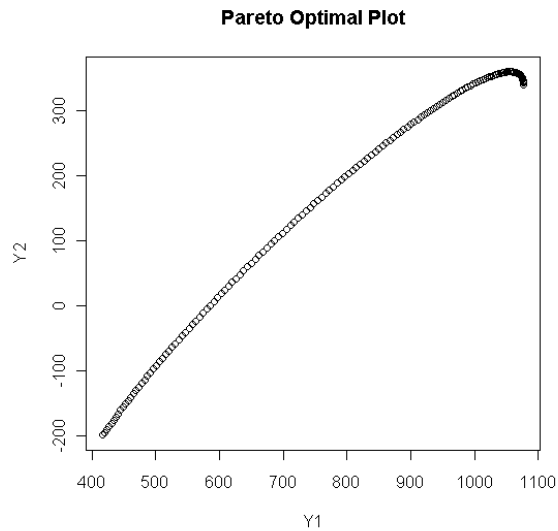


Figure 4.14 – Example 3: Pareto Optimal plot for the two responses

Based on Mee and Xiao’s approach, we should move in the direction of the point that best compromises our responses. Thus, the experimenter would have to ask what compromises should be made in order to get the desired response. Because we do not know the optimal trade off, a specific point is not chosen. However, in an actual analysis the experimenter would find the best trade off and continue in that direction with sequential experimentation.

However, this method leaves some questions unanswered. For instance, it is not clear what should be done if the true desired path is not a convex combination of the paths of steepest ascent. Although this is something we would not know, it is possible this could happen. Similarly, neither of those methods considers the uncertainty in the parameter estimates. Peering at the confidence cone from the example above, the true paths could have been slightly different. This would have made the convex combinations drastically different. For instance, the path could have been the same for y_1 but for y_2 , so we instead had the upper boundary of the confidence cone to be the actual path of

steepest ascent. Because this is a 95% confidence cone, it is possible the true path of steepest ascent could be here. Illustrating this in Figure 4.15, we could have had:

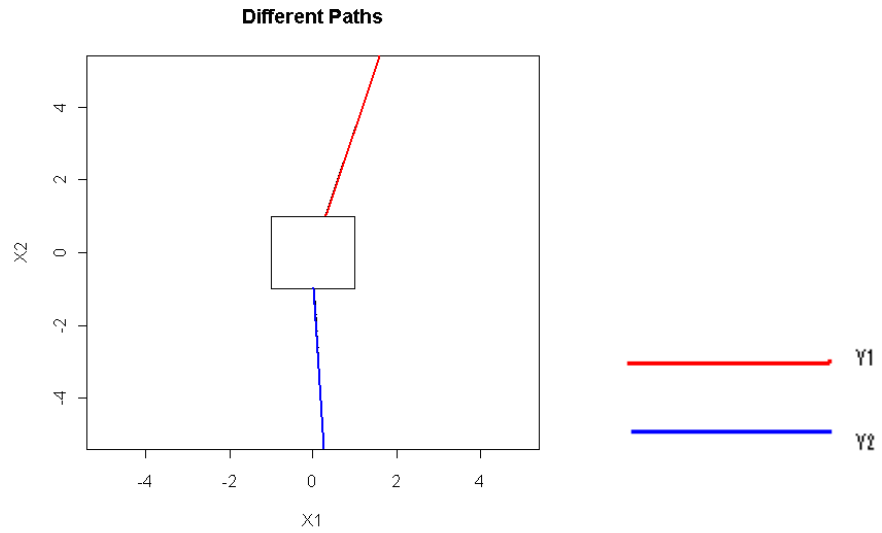


Figure 4.15 – Example 3: Possible different paths of steepest ascent

In this case, the convex combinations would have now been on the right side of the graph (the shaded region in Figure 4.16):

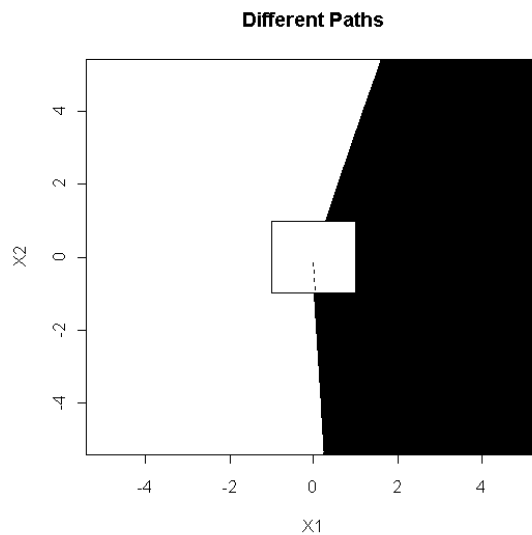


Figure 4.16 – Example 3: Possible different convex combinations of the paths of steepest ascent

Therefore, a different method should be used to calculate the path of interest. Utilizing our Bayesian reliability function, we consider all the points in the design region around the radius of $x'x=r^2$, or $\sqrt{2}$, to form our probability. We randomly sampled from the multivariate t distribution with mean 0, variance H^{-1} , and degrees of freedom, $\gamma=5$. From here we found the probability that a given point would create $y_1 \geq 850$ and $y_2 \leq 12$. These responses were arbitrarily chosen to represent the desired minimum trade off between the two responses. The ten points with the highest probabilities are shown in Table 4.11.

Table 4.11 – Example 3: 10 Highest Reliable Points

X1	X2	Reliability
-1.14412	0.831254	0.192
-0.94629	1.050966	0.18
-0.8511	1.129441	0.162
-0.96449	1.03429	0.152
-0.87068	1.114416	0.15
-1.0173	0.982395	0.144
-1.03429	0.964491	0.144
-1.08335	0.909039	0.144
-1.05097	0.946294	0.142
-1.06732	0.927808	0.142

The point with the highest probability is the point that we decided to move to, shown in Figure 4.17:

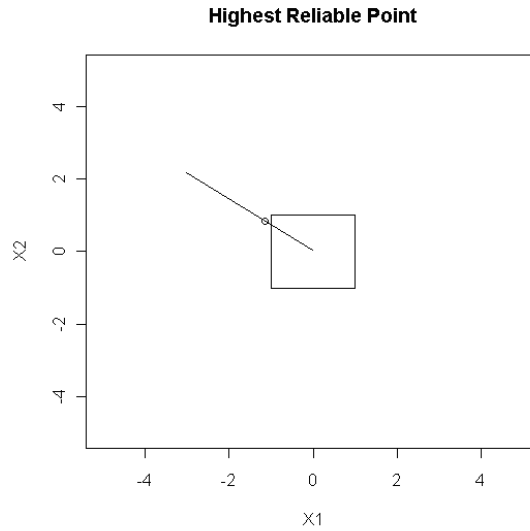


Figure 4.17 – Example 3: Highest reliable point

Therefore, the new path is $x_1 = -1.14$ and $x_2 = 0.83$ to maximize y_1 and minimize y_2 . Plugging these into our first order model yields:

$$\hat{y}_1 = 723.5667 - 44.85 * 1.14 + 151.4 * .083 = 685.0039$$

$$\hat{y}_2 = 20.12333 - 8.19 * 1.14 + 7.815 * .083 = 11.435375$$

In this case, the highest reliable point is not within our cutoff value for y_1 , which explains the low reliability of 0.192. However, the cutoff value for y_2 is satisfied. In order to increase the reliability, we might consider changing our cutoff value. This is especially true because the point that was selected did not satisfy our criteria.

4.4 Example 4

Now let's consider a simulation example. In this case, we want to see how the highest reliable point is affected as the error variance increases. Consider the following β 's:

Table 4.12 – Example 4: True Values of the Parameter Estimates

β	Y_1	Y_2
β_0	28.75	35.25
β_1	8	-8
β_2	4	-4.5

From the X matrix in Table 4.13, we can then simulate data using these β 's:

Table 4.13 – Example 4: X Matrix

Intercept	X_1	X_2
1	-1	-1
1	1	-1
1	-1	1
1	1	1
1	0	0
1	0	0
1	0	0
1	0	0

From here, we can see where the path of steepest ascent is in relation to our design region, shown in Figure 4.18:

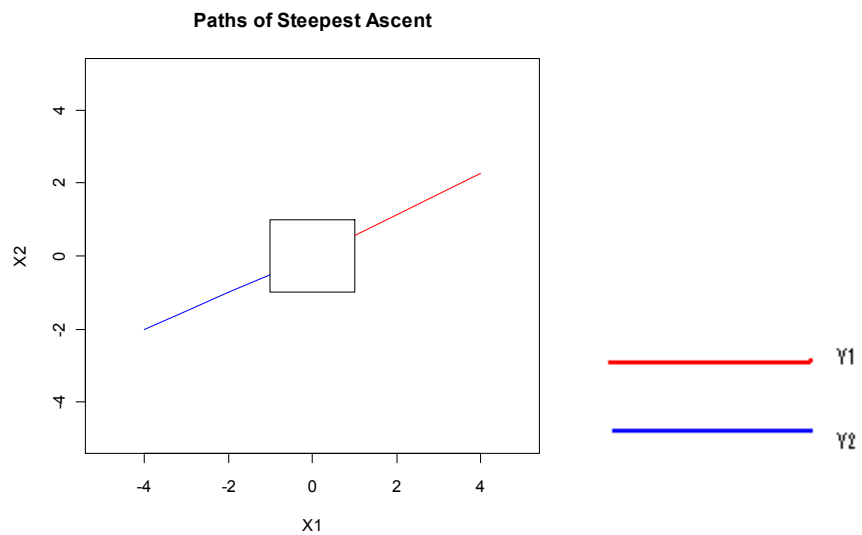


Figure 4.18 – Example 4: Paths of steepest ascent

Therefore, looking at the two paths, we know the Pareto Optimal point will be a convex combination, and thus between the two paths, shown in the shaded region in Figure 4.19.

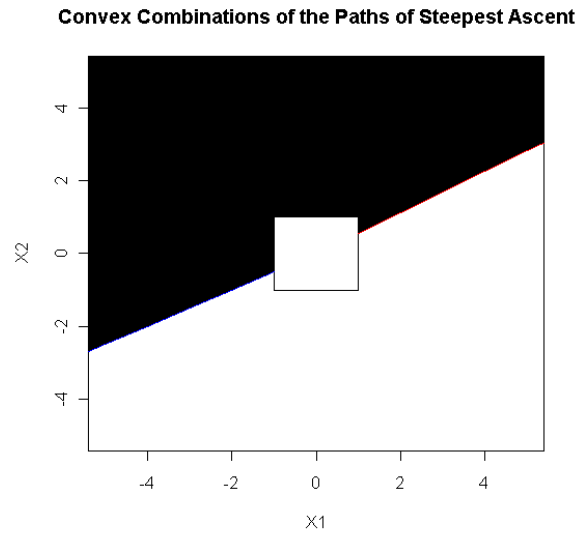


Figure 4.19 – Example 4: Convex combinations of the paths of steepest ascent

Notice the two paths are almost completely disparate directions. From here, we can calculate our \hat{y} s from the equations:

$$\begin{aligned}\hat{y}_1 &= 28.75 + 8x_1 + 4x_2 \\ \hat{y}_2 &= 35.25 - 8x_1 - 4.5x_2\end{aligned}$$

To simulate data, we looked at 1000 trials for two different error variances from a random error vector. Thus, we choose to look at two different models for each response:

$$\begin{aligned}\hat{y}_{11} &= 28.75 + 8x_1 + 4x_2 + \varepsilon \sim N(0,1) \\ \hat{y}_{21} &= 35.25 - 8x_1 - 4.5x_2 + \varepsilon \sim N(0,1) \\ \hat{y}_{12} &= 28.75 + 8x_1 + 4x_2 + \varepsilon \sim N(0,3) \\ \hat{y}_{22} &= 35.25 - 8x_1 - 4.5x_2 + \varepsilon \sim N(0,3)\end{aligned}$$

We then ran through the same code as before, sampling from around the design region a radius of $\sqrt{2}$. To save computation time, we only looked at every other degree

(e.g. $1^\circ, 3^\circ$ etc.). For each sampled point, we found a probability for $y_1 \geq 31$ and $y_2 \geq 31$ for $n=1000$. The values for the responses were arbitrarily chosen based on the first order models and where the responses are optimized. The point with the highest reliability was selected and placed into a matrix. For each of the 1000 trials, there are 1000 points with each point being the highest reliability for that particular trial. Figure 4.16 shows these points graphically for when the random noise is distributed normally with a mean of 0 and variance of 1.

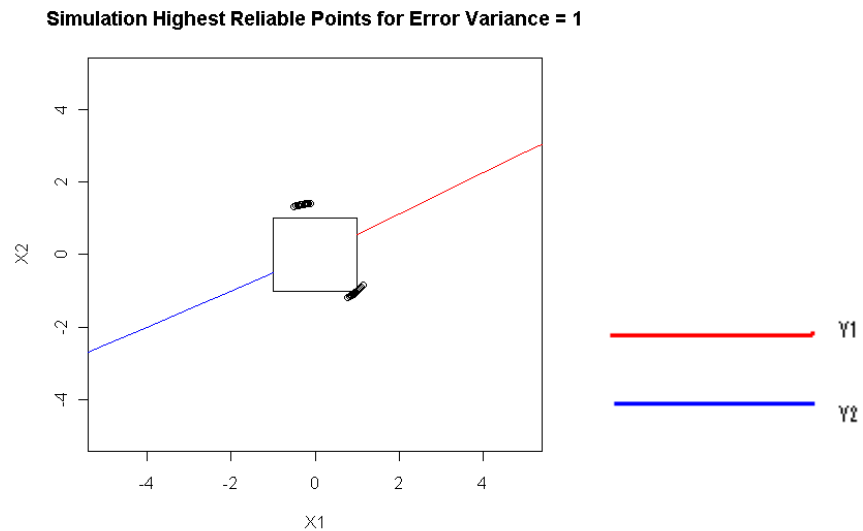


Figure 4.20 – Example 4: Simulation highest reliable points for error variance equal to 1

The lines represent the true paths of steepest ascent. As we can see from the graph, the highest reliable points can be on either side of the paths, and they do not necessarily have to be convex combinations of the true paths of steepest ascent.

Similarly, Figure 4.17 shows the highest reliable point when the error is normally distributed with a mean of 0 and a variance of 3:

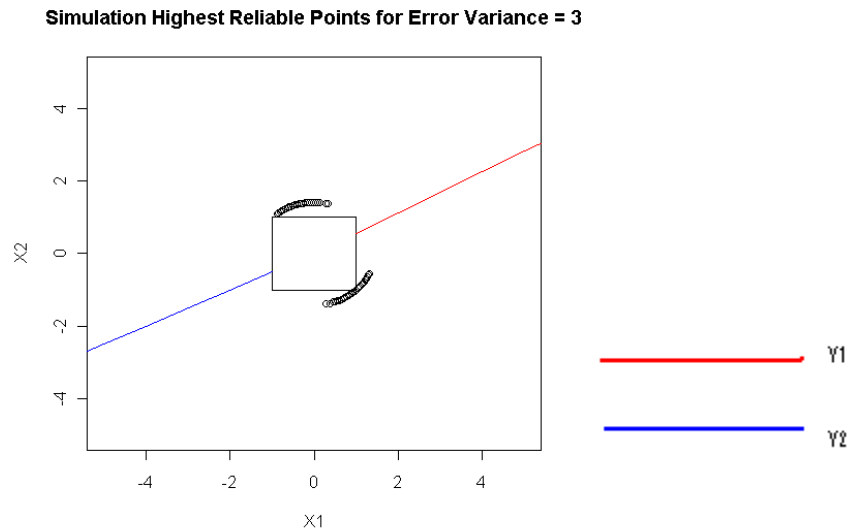


Figure 4.21 - Example 4: Simulation highest reliable points for error variance equal to 3

We can see the points are more spread out but in the same general area. This is expected, as now we have a higher variance for the error. On the other hand, once again the points can be on either side of the true paths of steepest ascent and are not necessarily a convex combination of these paths.

These simulations illustrate that using the highest reliable point gives a direction to move in for the path of steepest ascent, and this method takes into account the variances of the predictions to adequately provide this direction. Using the highest reliable point, we are able to glean the direction where future responses can be optimized. Also, because the points are on either side of the shaded region, the uncertainties of the parameter estimates are considered.

However, what these simulations do not show is whether the highest reliable points are convex combinations of the simulated paths of steepest ascent. For instance, in each simulation new estimates were found and thus new paths of steepest ascent were generated. From these paths it is impossible to know if the highest reliable point is still a

compromise direction. Therefore, we ran twenty additional simulations, each time looking at the graph for the simulated paths of steepest ascent and determining if the highest reliable point was a convex combination. From these trials we found that the highest reliable point was always a convex combination; so it appears that Mee and Xiao's method of finding the Pareto Optimal point does have some merit in that the compromise direction will always be a convex combination. However, MX's method gives the entire spectrum of points that are convex combinations and forces the experimenter to choose a point based on the best tradeoff. In our method we give a specific point that leads to the direction for the path of steepest ascent. This leaves no doubt from the experimenter as to if he chooses the right direction in which to move.

4.5 Example 5

Now let's consider an example with more than two factors. The following example was originally in Videvogel and Sandra (VS) (1991), but Mee and Xiao utilized the same experiment to show their results. VS conducted a five-factor, eight run fractional factorial design involving six responses to study the electrokinetic chromatography for separation of testosterone esters. Each run resulted in a chromatogram, and six characteristics of the chromatograms were measured. Table 4.14 lists the experimental runs, as well as the responses, two of which are reported here:

Y_{Res} , the resolution for distinguishing the second and third esters
 $Y_{Rate} = 1 / Y_{Time}$, where Y_{Time} is the eluting time for the fourth ester.

Table 4.14 – Example 5: Experimental Runs and Responses

Intercept	X ₁	X ₂	X ₃	X ₄	X ₅	Y _{Res}	Y _{Rate}
1	1	1	-1	1	-1	1.44	0.07315
1	1	1	1	-1	-1	0.2	0.09434
1	-1	1	1	1	1	0.5	0.08489
1	-1	-1	1	1	-1	0.6	0.08503
1	1	-1	-1	1	1	1.96	0.05163
1	-1	1	-1	-1	1	0.73	0.10215
1	1	-1	1	-1	1	0.6	0.07994
1	-1	-1	-1	-1	-1	0.84	0.099

Large values for Y_{Res} and Y_{Rate} are desired, and the primary objective of the experiment is to simultaneously optimize both responses. We first need to calculate $\hat{\beta}$, which yields:

Table 4.15 – Example 5: Estimates for the Responses

$\hat{\beta}$	Y ₁	Y ₂
$\hat{\beta}_0$	0.85875	0.083766
$\hat{\beta}_1$	0.19125	-0.009001
$\hat{\beta}_2$	0.14125	0.004866
$\hat{\beta}_3$	-0.38375	0.002284
$\hat{\beta}_4$	0.26625	-0.010091
$\hat{\beta}_5$	0.08875	-0.004114

In this case, because we have more than two factors, grid searching around a circle is not logical because we have more than two dimensions. Therefore, we implemented Peterson's search algorithm for multiple dimensions in order to sample around the design region. We used the same formula for the radius as before, $x'x=r^2$, yielding $\sqrt{5}$. From this, we were able to find the highest probability for maximizing both responses. In this instance, we chose to look for $y_1 \geq .86$ and $y_2 \geq .083$, as these are the means from VS's experiment. The results shown in Table 4.16 represent the ten highest reliable points to move along for the highest reliability.

Table 4.16 – Example 5: 10 Highest Reliable Points

X1	X2	X3	X4	X5	Reliability
-1.208153	0.783016	-1.710922	3.67E-17	9.81E-17	0.82
-1.208153	0.783016	-1.710922	7.39E-17	-7.42E-17	0.808
-1.208153	0.783016	-1.710922	-1.48E-17	1.04E-16	0.796
-1.208153	0.783016	-1.710922	1.02E-16	2.21E-17	0.794
-1.208153	0.783016	-1.710922	-1.04E-16	-7.41E-18	0.792
-0.158173	0.928202	-1.779876	-4.66E-01	-8.53E-01	0.79
-0.158173	0.928202	-1.779876	-8.18E-01	-5.25E-01	0.79
-0.158173	-0.157777	-1.952515	-8.98E-01	-5.76E-01	0.786
-0.158173	0.928202	-1.779876	-8.84E-01	4.05E-01	0.786
-1.208153	-0.133098	-1.647111	-4.31E-01	-7.90E-01	0.784

Using the highest reliable point, we can see the estimated values for each response by entering in the values from the point into our first order model. This yields:

$$\hat{y}_1 = .85878 - .19125 * 1.2 - .14125 * .783 + .38375 * 1.71 + .26625 * 0 + .08875 * 0 = 1.1749$$

$$\hat{y}_2 = 0.083766 + .009001 * 1.2 + .004866 * .783 - .002284 * 1.71 - .010091 * 0 - .004114 * 0 = .0945$$

Both responses are indeed above our cutoff values.

CHAPTER 5 Conclusion

Past approaches to finding the path of steepest ascent have not considered several key aspects: uncertainty in the model parameters, correlations and variances among the responses, and the variability of the process distribution. These elements are extremely significant with quality assessment, and future predictions could be jeopardized by not taking them into account. We can see from the examples that the Bayesian reliability method to calculate the path of steepest ascent provides a complete way to assess this quality. This approach is currently the only one that takes the variance-covariance structure and model parameter uncertainty into account. In addition, this method easily allows for the experimenter to measure the effect of changing the variance of the process being studied. Similarly, this method is flexible in that the experimenter can easily change the criteria for selecting a path of steepest ascent.

Computing a Bayesian reliability $p(x)$ for a specific point takes very little computational time, so searching for a specific point outside of the design region for two factors takes little more with high precision. When increasing the number of factors, the routine increases in runtime exponentially; this is why Peterson's polar coordinate transformation was utilized.

As we can see with Example 4.1, the Bayesian reliability method matches up with the overlap of the confidence cone as well as this point is a convex combination of the paths of steepest ascent. This is what we expected and proves that Mee and Xiao were

correct in stating that the compromise path of steepest ascent is a convex combination. The difference with our method is that we can find a specific direction to take rather than choosing from a set of trade offs for the best path to take.

For disparate paths of steepest ascent, we can see from Example 4.2 the Pareto Optimal point is not the best route to take. Because Pareto Optimality criteria do not include the correlations of the responses the predictions could be negatively affected. Therefore, the optimal method is shown through the Bayesian reliability approach, which considers the correlations of the future responses and produces the highest point for accurate and precise results.

This method is also not limited dimensionally. For more than two factors a point can be found where the highest reliability is found, although the result cannot be illustrated graphically. This point will lead us in the direction of the path of steepest ascent (descent), in which case sequential experimentation can be utilized to find a second-order model and continue optimization.

When looking at the reliability, we can see from the simulation example that the path of steepest ascent for two variables can vary dramatically when the error variance is high. The simulation also highlights that it is possible to have higher reliability all around the design region and not just in the convex combinations for the true paths of steepest ascent. However, it is impossible to know if the highest reliable points are not convex combinations of the simulated paths of steepest ascent from the Figures shown in the Examples section.

Future work can assess this reliability more accurately. For example, a researcher can utilize this method in an actual experiment where one can fully see if the highest

reliable point gives the path for optimized responses. Testing along this direction, one can see where the optimization occurs, and sequential experimentation can be performed, including generating a second-order model. By utilizing the Bayesian reliability approach an experimenter can be sure that the responses generated consider the correlations between responses as well as parameter uncertainty.

Future research will address how we can increase reliability as well as addressing the uncertainty of the model itself by using Bayesian model averaging. To increase the reliability we might consider changing the cutoff values for the responses. Along the same lines, we could check the simulated trials to determine if the highest reliable point was indeed a convex combination of the paths of steepest ascent.

Similarly, when we looked at more than three dimensions, a coarse grid search was performed. In the future, one could perform a coarse grid search first, but then as we find an area where the highest reliabilities are located, one could search a finer grid around that area. This can be repeated until the highest reliability can be found. In the same situation, we might try optimization techniques like the Nelder-Mead Simplex method.

In addition, sensitivity analysis can be performed on the path of steepest ascent generated by the Bayesian reliability approach, i.e. finding a specific reliability. One might want a reliability no smaller than 0.8. In this case, we would have to search the around the design region as well as change the cutoff values for the responses to achieve this reliability.

Using the Bayesian reliable approach that we did, we only looked at the vague prior to calculate our posterior distribution. We could have updated the prior with some

knowledge of the parameters, such as adding a normal distribution to the betas, and then calculated the posterior predictive distribution to use for our sampled point. This could have provided a higher reliability or might have lead us in a different direction, so it is worth exploring in the future.

Bibliography

Bibliography

- Ames, A.E.; Mattucci, N.; Macdonald, S.; Szonyi, G.; and Hawkins, D.M. (1997), "Quality Loss Functions for Optimization Across Multiple Response Surfaces," *Journal of Quality Technology*, 29, 339-346.
- Biles, W.E. (1975), "A Response Surface Method for Experimental Optimization of Multi-Response Processes," *Ind. Eng. Chem., Process Des. Dev.*, 14, 152-158.
- Box, G. E. P. and Draper, N. R. (1987), *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York, NY.
- Box, G. E. P., and Wilson, K.B. (1951), "On the Experimental Attainment of Optimum Conditions," *Journal of the Royal Statistical Society, Ser. B*, 13, 1-45.
- Chatterjee, Sanjit, Laudato, Matthew, and Lynch, "Genetic algorithms and their statistical applications: an introduction," *Computational Statistics & Data Analysis*, 22, 633-651.
- Del Castillo, E. (1996), "Multiresponse Process Optimization via Constrained Confidence Regions," *Journal of Quality Technology*, 28, 61-70.
- Del Castillo, E. (2007), *Process Optimization: A Statistical Approach*, New York: Springer.
- Del Castillo, E. and Montgomery, D.C. (1993), "A Nonlinear Programming Solution to the Dual Response Problem," *Journal of Quality Technology*, 25, 347-353.
- Derringer, G. C., and Suich, R. (1980), "Simultaneous Optimization of Several Response Variables," *Journal of Quality Technology*, 12, 214-219.
- Harrington J. (1965). "The desirability function," *Ind. Qual. Control*, 21, 494-498.
- Heller, Nelson B., and Staats, Gleen E., "Response Surface Optimization When Experimental Factors Are Subject to Costs and Constraints," *Technometrics*, 15, 113-123.
- Khuri, A.I., and Conlon, M. (1981), "Simultaneous Optimization of Multiple Responses Represented by Polynomial Regression Functions," *Technometrics*, 23, 363-375.

- Mee, R.W., and Xiao, J. (2008), "Steepest Ascent for Multiple-Response Applications," *Technometrics*, 50, 371-382.
- Myers, R. H. and Carter, W.H. (1973), "Response Surface Techniques for Dual Response Systems," *Technometrics*, 15, 301-317.
- Myers, R.H. (1999). "Response Surface Methodology – Current Status and Future Directions," *Journal of Quality Technology*, 31, 30-43.
- Myers, R. H., and Montgomery, D.C. (2002), *Response Surface Methodology*, New York: Wiley.
- Nelder, J.A., and Mead, R. (1964), "A simplex method for function minimization," *Comput J*, 7, 308-313.
- Peterson, J.J. (1993), "A General Approach to Ridge Analysis With Confidence Intervals," *Technometrics*, 35, 204-214.
- Peterson, J.J. (2004), "A Posterior Predictive Approach to Multiple Response Surface Optimization," *Journal of Quality Technology*, 36, 139-153.
- Raiffa, Howard, and Schlaifer, Robert. (2000), *Applied Statistical Decision Theory*. John Wiley & Sons, New York, NY.
- Taguchi, G. (1986). *Introduction to Quality Engineering*. UNIPUB/Kraus International Publications. White Plains, NY.
- Tang, L.C., and Xu, K. (2002), "A Unified Approach for Dual Response Surface Optimization," *Journal of Quality Technology*, 34, 437-447.
- Vindevogel, Johan and Sandra, Pat (1991), "Resolution Optimization in Micellar Electrokinetic Chromatography: Use of Plackett-Burman Statistical Design for the Analysis of Testosterone Esters" *Analytical Chemistry*, 63, 1530-1536.
- Vining, G. G. and Gohn. L. (1998). "Response Surfaces for the Mean and the Variance Using a Nonparametric Approach", *Journal of Quality Technology*, 30, 282-291.

Vita

Jeffrey Norman Fuerte, Jr. was born June 23, 1986 in Portsmouth, Virginia and is an American citizen. He graduated from J.R. Tucker in 2004, and received his Bachelor of Science in Statistics from Virginia Polytechnic Institute and State University, Blacksburg, Virginia in 2008. For the past two years, he has taught the lab in Statistics 208 at Virginia Commonwealth University.