An Efficient Methodology for Analysis of Stochastic Computer Simulation Experiments

---

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Doctor of Philosophy

by

Suzanne M. DeLong

May

2012

APPROVAL SHEET

The dissertation

is submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Suzanne M. DeJong
AUTHOR

The dissertation has been read and approved by the examining committee:

Advisor

Accepted for the School of Engineering and Applied Science:

James H. Taylor

Dean, School of Engineering and Applied Science

May

2012

ACKNOWLEDGMENTS

I would like to thank my husband, Vaughn, and children Keagan and Ryan. Thank you for enduring my long hours, and family separation so I could pursue this goal. Special thanks to Dr. Joyce A. Nagle and Dr. Matthew S. Gerber for your time in reviewing my work and providing valuable feedback. Thanks to Yue Zang for working with me on the Repast simulation, Matlab coding and the many discussions we had on the police staffing study. Thank you to my committee for sticking with me through this process over the past several years. Your patience, willingness, and guidance were very much appreciated. I would also like to thank my advisor, Professor Donald E. Brown for giving me direction. You were always available to look over my work and provide feedback whatever the hour. Finally, I must thank Constance Brown for giving me encouragement and a cup of tea when I needed both.

Abstract

Much research was done to improve the efficiency of the experimental design process through methods that improve and extend response surface methods (RSM). This research will take ideas used in two different areas of research and combine these ideas and apply them unconventionally to a new area. The first idea comes from the medical research community where there is a need for the efficient conduct of clinical trials accomplished through interim data monitoring and analysis. (Jennison and Turnbull, 2006, Karrison, Hua, and Chappell, 2003, Burington and Emerson, 2004) The second idea comes from the area of industrial processes where there is a desire to increase productivity efficiently through the use of Evolutionary Operation (EVOP). (Box and Draper, 1969) RSM is the vehicle primarily used in industrial processes to achieve this efficiency. Methods used in these two areas provide the basic methodology in which to improve the efficiency of simulation experimentation, where complex systems are examined.

The purpose of this dissertation is to develop a new experimental design process for large, complex stochastic simulations that maximizes one or more output performance measures while meeting constraints for cost, time and number of replications required for maintaining an absolute error of at most β. The number of replications constraint will save time by reaching a near optimal solution and may result in allowing the experiment to end early, start follow-up studies earlier or to terminate futile studies. This research develops of a new methodology termed the Efficient Computer Experiment Methodology (ECEM) and is shown to achieve efficiencies in replications of 40% or greater when compared to traditional methods.

This methodology achieves efficiencies in several areas. First, establishing a data monitoring plan and using power analysis ensures that only the necessary replications are conducted to achieve statistically significant results. EVOP, a method traditionally used in the industrial process, is used to improve factor settings during the course of simulation experimentation in order to improve the response(s) and allow greater experimental efficiency. The use of orthogonal and nearly orthogonal Latin hypercube designs (OLHD/NOLHDs) is advantageous and efficient as they are a form of stratified random sampling which gains efficiencies through minimizing variance. A Gaussian process model is fit in order to use simulated annealing (SA) to optimize the response prediction formula which results in another savings in replications. Interim analysis, used during the conduct of clinical trials, is inserted throughout simulation experimentation in order to determine when to end an experiment so as not to waste resources and to effectively draw conclusions through the sequential analysis of hypotheses. This methodology is a new approach to efficient experimentation and uses a new EVOP approach using OLHDs/NOLHDs.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

ANOVA – Analysis of Variance

CART – Classification and Regression Trees

CLT – Central Limit Theorem

DOE – Design of Experiments

ECEM – Efficient Computer Experiment Methodology

EVOP – Evolutionary Operation

GA – Genetic algorithm

GLM – General linear model

GSM – Group sequential methods

GST – Group sequential tests

iid – Identically and independently distributed

LHD – Latin hypercube design

LHS – Latin hypercube sampling

MARS – Multivariate adaptive regression spline

MLSM – Moving least squares method

MRSM – Modified Response Surface Methodology

NMSS – Nelder-Mead Simplex Search

NN – Neural networks

NOLHD – Nearly orthogonal Latin hypercube design

OAT – One at a time

OLHD – Orthogonal Latin hypercube design

RBF – Radial basis functions

REVOP – Random EVOP

ROVOP – Rotating square EVOP

RSM – Response surface methodology/Response surface model

RST – Repeated significance test

SA – Simulated annealing

SB – Sequential bifurcation

SPRT – Sequential probability ratio test

SRS – Simple random sampling

SVM – Support vector machines

TS – Tabu search

WLS – Weighted least squares

1 Introduction

Systems in a variety of fields to include production, transportation, economics, biology, ecology, law enforcement, military operations, behavioral and social sciences and in the medical field (Graham, 2004) have become extremely complex and increasingly expensive to develop. In order to avoid wasting resources while maintaining the ability to draw meaningful conclusions, the need for efficient experimentation developed. A recent trend in experimentation is the use of computer simulation. (Fang, Li, and Sudjianto, 2006) As computing capabilities increased, the application of computer simulation grew as did the demand for greater fidelity representation in models. As a result, model complexity has increased. The increased complexity resulted in an increase in required resources and also the need for efficient design in computer simulation experimentation. (Giunta, Wojtkiewicz and Eldred, 2003) The effort to model real-world phenomena as closely as possible has resulted in large scale computer simulations that are time consuming to run and involve a large number of input variables. The original problem that occurred with traditional experimentation now exists and there is a need for even greater efficiency when using computer simulation experimentation. (Fang et al., 2006)

Schamburg (2004) states that, "current approaches to analysis are not well structured for human learning through the use of complex simulations. For analysis, often times the systems engineer is forced to make numerous simplifying assumptions and then select a few variables over a limited domain". This dissertation will address how the analyst can structure the experimental design process with a general approach that will improve analytical methods for complex, stochastic computer simulation experiments and

create an efficient computer experiment methodology (ECEM). This research extends current research in several areas and focuses on stochastic computer simulations with multiple responses. The methodology starts by specifying the level of accuracy desired and aims to achieve results while meeting the desired accuracy in an efficient manner through power analysis. Orthogonal or nearly orthogonal Latin hypercube designs (OLHD/NOLHD) are applied and the results are modeled with Gaussian process models during the Evolutionary Operation (EVOP) process. The response prediction formula found through Gaussian Process modeling is optimized with simulated annealing (SA) and multiple objective optimization. Interim data monitoring and analysis typically found during the conduct of clinical trials is applied to guide the experimentation process. The methodology presented here is referred to as the efficient computer experiment methodology (ECEM).

This dissertation is laid out as follows. Chapter 2 defines the problem and gives the problem statement and scope. Chapter 3 provides background information and a literature review on efficient experimentation and discusses the main components of the methodology developed. Chapter 4 explains the steps of the methodology and the necessary assumptions of the ECEM. Chapter 5 presents the properties of the ECEM. Chapter 6 demonstrates the methodology on the application of a chemical mixing experiment and on a police staffing study. Chapter 7 states the conclusions and contributions of this research. Chapter 8 addresses future work in this area.

2 Problem Statement and Scope

Current methods for improving the efficiency of computer simulation experimentation include methods that improve and extend response surface methods (RSM) and methods that find the global optimum while decreasing the number of replications. This dissertation will seek to develop alternative methods to supplement current methodologies to accomplish this objective. The alternative methods involve the use of EVOP while using OLHDs/NOLHDs, using RSM and applying SA and multiple objective optimization to the response prediction formulas, and using interim data monitoring and analysis tools typically used in the conduct of clinical trials. The applications tested are large, complex, stochastic computer simulations with multiple responses.

The purpose of this dissertation is to develop a new methodology that analyzes stochastic computer simulations efficiently while optimizing one or more output performance measures and meeting constraints such as cost, time and\or number of replications required for maintaining an absolute error of at most β. The number of replications constraint will save time by reaching a near optimal solution that is within tolerance of the upper bound of the optimal solution and may result in allowing the experiment to end early, start follow-up studies earlier or to terminate futile studies. This will be accomplished through the application of interim data monitoring and analysis employed during the conduct of clinical trial analysis. In addition, response surface methods will be employed during the EVOP process with OLHD/NOLHD schemes while applying SA to the response prediction formula in order to find the factor settings that optimize the responses under study.

The stochastic optimization in this work will use the following notation:

Maximize\Minimize E($\mathbf{R}(v_1, v_2, \ldots v_k)$) (2.1)

Where

E($\mathbf{R}(v_1, v_2, \ldots v_k)$) is the expected value of **R** given input factors v_k

such that

$l_1 \leq v_1 \leq u_1$

$l_2 \leq v_2 \leq u_2$

.

.

$l_k \leq v_k \leq u_k$

$t_{i\text{-}1,\ 1\text{-}\alpha/2} \sqrt{S^2(n)/i} \leq \beta$

$l_n \leq n \leq u_n$

$g(n) \leq g$

$c(n) \leq c$

R = output performance measure(s)

v_f = value of input factors where f = 1..k

l_f = lower bound on the factors where f = 1..k

u_f = upper bound on the factors where f = 1..k

n = number of replications

n_l = lower bound on (n) number of replications required

n_u= upper bound on (n) number of replications required

$\boldsymbol{S}^2(n)$ = variance estimate of **R** based on n

i = increment (replication) number

α = significance level (Type I error)

β = maximum absolute error (Type II error)

$t_{i\text{-}1,\ 1\text{-}\alpha/2}$ = test statistic

$g(n)$ = time to meet objective function in n replications

g = time constraint

$c(n)$ = cost ($) of n replications

c = total amount budgeted

In matrix form this problem can be expressed as follows*:*

Maximize E[**R**(***V***)] (2.2)

Where

E[***R***(***V***)] is the expected value of **R** given input factors ***V***

such that

$\boldsymbol{L} \le \boldsymbol{V} \le \boldsymbol{U}$

$t_{i\text{-}1,\ 1\text{-}\alpha/2}\ \sqrt{S^2(n)/i}\ \le \beta$

$l_n \le n \le u_n$

and\or $g\ (n) \le g$

and\or $c(n) \le c$

where

R = output performance measure(s)

V = value of input factors 1..k

L = lower bound on factors 1..k

U = upper bound on factors 1..k

n = number of replications

n_l = lower bound on (n) number of replications required

n_u= upper bound on (n) number of replications required

$\boldsymbol{S}^2(n)$ = variance estimate of **R** based on n

i = increment (replication) number

α = significance level (Type I error)

β = maximum absolute error (Type II error)

$t_{i\text{-}1,\ 1\text{-}\alpha/2}$ = test statistic

$g(n)$ = time to meet objective function in n replications

g = time constraint

$c(n)$ = cost ($) of n replications

c = total amount budgeted

The single objective problem is difficult to solve because if the number of factors k is large, then searching for an optimal solution in k-dimensions increases the size of the experimental design as a power or two or more. This involves evaluating the objective function many times and could become very costly. Where the objective function is evaluated and how many times it is evaluated needs to be determined. As a result of the complexity of the problem, the objective function cannot be evaluated by simply plugging a set of possible decision variables into a closed-form equation. The simulation must be run to produce an output **R**. If the simulation is stochastic, replications are required to evaluate the objective function. For complex systems, this means complex simulations with a large number of input variables and would therefore require an

extremely large number of replications. (Law and Kelton, 2000)

There is significant literature on the design and analysis of experiments with computer simulations and on stochastic computer simulations which contain references to how efficiencies can be gained. Many of the techniques and assumptions rely heavily on replications in order to get an accurate estimate of the variance and these methods could result in a lot of costly simulation time to achieve these estimates. This research extends the literature in the area of the design and analysis of stochastic computer simulation experiments and presents an efficient computer experiment methodology (ECEM). This dissertation will demonstrate the new methodology's applicability to a chemical mixing process simulation as well as to a police staffing simulation study.

This dissertation will seek to answer the following research questions: 1) Will the statistical analysis tools employed during clinical trials and applied to stochastic simulation experiments improve the process? 2) How can complex, stochastic simulation computer experiments be conducted efficiently? 3) Can current EVOP and RSM be extended to apply more sophisticated DOE to gain efficiency?

3 Background

This literature review includes a review of efficient methods that are currently available in the design of experiments (DOE) and RSM. EVOP is reviewed to give the background behind this method that was established in the 1950's. Techniques for model fitting are addressed as well as alternative methods to search the design space and optimize the responses. Finally, clinical trial analysis is reviewed with the intentions that this research will leverage the use of a data monitoring plan, error spending functions and stopping boundaries such as those developed by Pocock, and O'Brien, and Flemming (Dmitrienko, Molenberghs, Chuang-Stein and Offen, 2007) to determine how many replications need to be done to meet desired accuracy and if and when the experiment should be stopped while using the EVOP process.

3.1 Efficient Experimental Design

The following four sections describe how efficiencies can be gained during the experimental design process. The selection of the experimental design, sampling techniques, variance reduction methods and model fitting are addressed with respect to how efficiencies can be gained in each of these areas.

3.1.1 Designs of Experiment

Experimental designs are considered efficient when they maximize information gained from an experiment with as few experimental trials as possible. This is important to achieve when the trials are expensive or the experiment trial runs are time-consuming. Experimental designs that have been developed for efficiency include two- and three-level fractional factorial designs, Plackett-Burman designs, and response surface designs

such as Box-Behnken and Central Composite designs. (Myers and Montgomery, 2002) In the case where there are restrictions on randomization, nested and split-plot designs are used. Additionally, RSM and methods to extend or improve RSM have been the primary means to improve upon the efficiency of experimentation. (Myers et al., 2002, and Schamburg, 2004) These techniques include the use of meta-models, variable screening, domain reduction and the use of space-filling designs.

Variable screening is accomplished either with individual or group screening techniques and allows one to determine which factors are significant so that the experiment can focus on the factors which cause the greatest effect on the responses studied. An individual screening technique is one-at-a-time (OAT) variable screening. (Box et al., 1969) Given many factors, this method would be highly resource intensive. Group screening techniques, also referred to as supersaturated designs; include iterated fractional factorial design and sequential bifurcation (SB) where interaction effects are only important if the corresponding main effects are important. Some screening designs are fractional factorials and Plackett-Burman designs. (Meyers et al., 2002)

Domain reduction is the process of testing smaller regions within the initial design space to gain knowledge about the entire region. This is a sequential procedure where successive tests of the design region may lead to discovering the global optimum. Section 3.3 will elaborate on some techniques to conduct the search of the design region while finding global optimums.

Space-filling designs scatter the design points within the design region rather than in clusters, at corners or on the surface such as is seen in classical designs. (Cioppa, 2002) The space-filling design is a new DOE concept used for modeling deterministic

systems when there is little or no information about the effects of the factors on the responses and the relationship is thought to be non-linear. (Santner, et al., 2003) The aim of space-filling designs is to spread the points as evenly as possible around the operating space. These designs fill out the *n*-dimensional space with points that are in some way regularly spaced. According to Husslage, Rennen and VanDam (2006), deterministic computer experiment designs should be space-filling. Section 3.1.2 describes sampling techniques and space-filling designs in more detail.

This research will use space-filling designs to analyze stochastic computer experiments. Space-filling designs employ sampling techniques such as Latin square sampling, orthogonal sampling, simple random sampling, stratified random sampling and uniform sampling. Some space-filling designs are distance-based designs such as maximin and minimax designs. Space-filling designs are desirable for computer experiments when prediction accuracy is of importance. (Santner, et al., 2003) The various sampling techniques are reviewed next.

3.1.2 Sampling Techniques

Simple random sampling consists of randomly selecting design points within the design region. Each design point has an equal probability of being selected and, therefore, it is possible that the design points could reside next to each other or in the same row or column of another design point since the columns and rows are not specified within the random design. Figure 3.1 demonstrates an example of simple random sampling within a design space and how it is the least restrictive of all the sampling techniques discussed here.

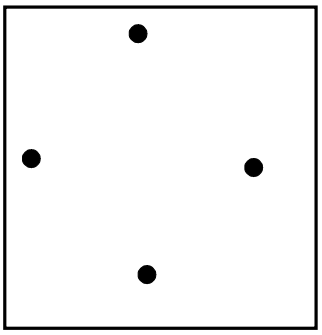


Figure 3.1 Simple random sampling within a design space randomly selects points with no regard as to where the previous design points were selected.

Latin square sampling is a statistical sampling technique where a square grid containing sample points contains only one sample in each row and each column. This differs from simple random sampling as the design points are guaranteed to be distributed across the region since the entire region is divided into grids and selected randomly within each row and column only once. Latin square sampling is more restricted than random sampling. Figure 3.2 shows a sample space divided into rows and columns and how samples can be taken from this space and exhibit the properties of Latin square sampling.

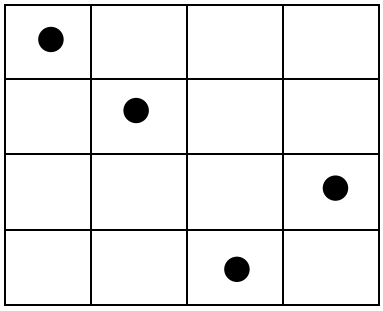


Figure 3.2 This is an example of a sampling technique called Latin Square Sampling where only one sample is contained in each row and in each column.

Orthogonal sampling is a result of sampling evenly over the entire region by

dividing the region into equally probable subsections and taking one sample point from each section as demonstrated in Figure 3.3. Orthogonality is more difficult to achieve as orthogonal sampling ensures that the selection of random design points is a very good representative of the real variability. (Steinberg and Lin (2006), Tang (1993), and Cioppa (2002) This is similar to Latin squares in that one sample comes from each row and column, however the grid is further divided into sections where only one point resides in the subsection as well as only one point in each row and column. Orthogonal sampling is more restricted than Latin square sampling.

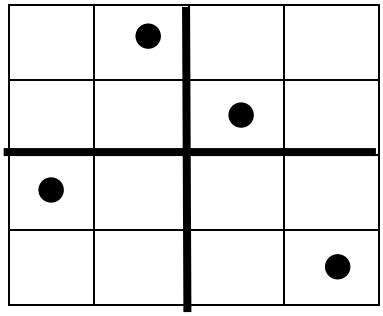


Figure 3.3 Demonstration of orthogonal sampling within a design space. The bold lines shows the addition of subspaces which further restrict where the design points are taken.

Stratified random sampling (Figure 3.4.) obtains a set of design points as a result of equally dividing the experimental region into *n* strata and randomly selecting one point from each stratum. Equally spaced points are selected across the experimental region using a uniform distribution. (Santner et al., 2003) This is similar to Latin squares in that sampling is done only once from within each stratum. The rows or columns of the Latin square are considered to be strata. This differs from Latin squares as row or column independence may not exist as is found in Latin squares.

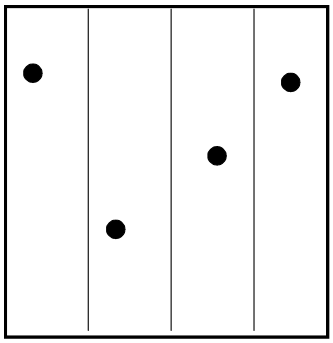


Figure 3.4 Stratified random sampling where the design space is divided into sections and one design point is selected within each stratum.

Uniform sampling distributes the design points uniformly across the design space. The distribution of the design points are compared to the uniform distribution. (Santner et al., 2003) One of the major disadvantages of uniform designs is difficulty in finding a design for many combinations of variables and runs, thus severely restricting the number of uniform designs readily available for use. (Cioppa, 2002) Santner et al. (2003) discusses work by Fang et al., (2006) where it is shown that uniform designs may be orthogonal thus making them attractive for use in computer experiments. Uniform sampling differs from simple random sampling in that uniform sampling guarantees that the design points are uniform across the sample space by design and not as a result of chance that could occur in simple random sampling.

Distance-based designs are based on the actual distance between design points. These designs ensure that no one point is too close to another point. Two distance based designs are maximin and minimax designs. In the maximin distance design the minimal distance between any two points is maximized. Maximin designs have good space-filling properties but are not always non-collapsing. (Husslage et al., 2006) The minimax distance design is where the maximum distance between any two points is minimized.

(Santner, et al., 2003) These designs are similar to Latin squares but the design points are spread evenly by a distance measure rather than by generic rows and columns. This distance measure can be considered as a way to delineate rows and columns. Some measurement criteria in evaluating the distribution of design points across the design region are: maximin and minimax distances, maximum entropy and discrepancy. The latter two are very difficult to calculate. For a more detailed description of these concepts see Santner, et al. (2003).

Other space-filling designs discussed by Santner et al. (2003) include grid designs, lattice designs and nets. All of these designs are beneficial when the sample size is large. These all appear similar to Latin square sampling in the sense that the design region is being divided into equal regions and then being sampled based on these regions and whether they are on a grid, lattice or net.

An experimental design which makes use of one of these sampling techniques is the Latin hypercube design (LHD) which employs Latin squares in "*n*" dimensions. In a Latin hypercube design, only one design point is taken in each row and column but also only within each plane in the *n*-dimensional design space. Not all LHD's are space-filling and there are actually only subsets of these designs that are considered space-filling. Santner et al. (2003), Schamburg (2004) and Cioppa (2002) summarize LHD's. Improvements to these space-filling designs are called orthogonal LHD's (OLHD) (Ye, 1998) and nearly orthogonal LHD's (NOLHD) (Cioppa, 2002). Due to the issues with multi-collinearity that arise with LHD's (Cioppa, 2002) this research will seek to leverage the benefits of OLHDs/NOLHDs.

In implementation, one must decide in advance how many sample points to take

and then record from which rows and columns they are taken. For each input variable, v_k, the range is divided into "*n*" strata all of equal marginal probability and each stratum are sampled only once. (Cioppa, 2002) An example of a space-filling LHD in two dimensions is shown in Figure 3.5. It is similar to Latin square sampling. As the third dimension is added, design points cannot come from the same strata as the other planes.

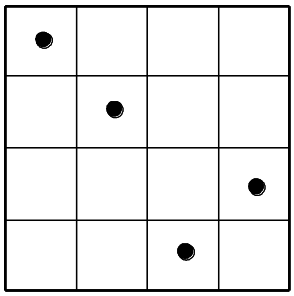


Figure 3.5 A space-filling LHD is similar to Latin square sampling and looks identical when viewing in two dimensions. As the third dimension is added the design points cannot overlap.

Figure 3.6 shows an example of a non-space-filling LHD in two dimensions. Note the presence of an apparent pattern in this non-space-filling design. To be space-filling the sampling should not exhibit any type of pattern and should appear random. (Santner, et al., 2003)

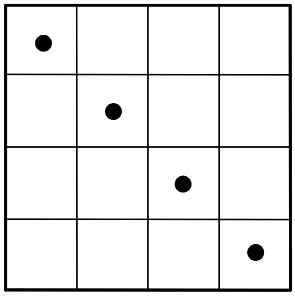


Figure 3.6 A Non-space-filling LHD exhibits a pattern as shown here with the design points lined up along the diagonal of the design space.

Orthogonal Latin hypercube design (OLHD) is an extension of the LHD where the orthogonality minimizes variance and ensures independence among the coefficient estimates in a regression model. (Ye, 1998) Cioppa (2002) states that "orthogonality enhances the ability to analyze and estimate as many effects, interactions, and jump discontinuities as possible." Orthogonal LHD's employ the concepts behind orthogonal sampling and are therefore more restricted than Latin squares or LHS. (Ye, 1998)

3.1.3 Variance Reduction

Reducing variance gains greater precision given the same amount of simulation time or achieves a specified precision with less simulation time. (Law and Kelton, 2000) There are several variance reduction techniques in the literature concerning stochastic simulation. Asmussen and Glynn (2007) and Ripley (1987) discuss the use of variance reduction techniques and how applying these techniques can greatly increase the efficiency of the simulation. Some of these techniques include control variates, antithetic sampling, common random numbers, importance sampling and stratification,. According to Ripley (1987), using more than one variance reduction technique may actually counter act the effects of the other or cause a larger variance and therefore it is not recommended to use multiple variance reduction techniques.

The use of control variates is the most widely used variance reduction technique where the expected value is known and correlated with the response Y. A sufficient number of replications should be completed so that the response Y and control variates distribution is normal. Caris and Janssens (2005) discuss how control variates can be used in a multiple response simulation with a small number of replications. Antithetic sampling forces the selection of samples so that they are not identically and

independently distributed (iid). A large number of samples are needed to get accurate results. Common random numbers utilizes a common stream of random numbers to cause correlation in the model and to reduce noise and improve computational efficiency.

Importance sampling and stratification are reviewed here as they are similar to the concepts of Latin square sampling that is used in this research. Importance sampling simply means that the samples are drawn from the state space where the most contribution to the response is found. Therefore some prior knowledge of the state space is required or is found through screening, pilot runs or previous trials. As discussed in the previous section, stratification is similar to Latin square sampling; the former is just slightly more restrictive in that the points cannot lie in the same rows or columns. Ripley (1987) states that it is the experimenter's responsibility to consider employing variance reduction during the course of a large simulation study. Given this information, this research achieves variance reduction through the inherent properties found when using OLHDs/NOLHDs and therefore gains efficiency and is discussed in more detail in section 5.2.2.

3.1.4 Model Fitting

After the design has been determined and the response data collected, a predictive model is fit to the data. Traditional methods of model fitting include linear regression where parameters are estimated using least squares, cubic and quadratic fits (polynomial RSM). More advanced methods include multiple regression analysis, multivariate adaptive regression splines (MARS), classification and regression trees (CART) meta-model, general linear models (GLM), Gaussian process models, Spatial correlation models such as Kriging, neural networks (NN), radial basis functions (RBF), support

vector machines (SVM) and moving least squares method (MLSM). Others who have applied meta-modeling techniques are Simpson, Korte and Mistree, 1998; Wang, 2003; Schamburg, 2004 and Crino, 2006. All have shown some success in improving experimental results although have not been entirely efficient.

The method of least squares estimates the linear coefficients of the response prediction formulas that minimize the sum of squared errors between the regression model and the prediction formula. First-order response prediction formulas are:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_k x_k \tag{3.1}$$

and second-order response prediction formulas are:

$$\hat{y} = b_0 + \sum_{i=1}^{p} b_i x_i + \sum_{i=1}^{p} b_{ii} x_i^2 + \sum \sum_{i<j}^{p} b_{ij} x_i x_j \tag{3.2}$$

Where $\hat{y}$ is the predicted response, the factors are denoted by x_i and b_0, b_i and b_{ij} are the regression coefficients which are unbiased estimators of the true parameters β_0, β_i, β_{ij}. Cubic, quadratic and higher order polynomials are derived from the Taylor series expansion and are detailed in Myers and Montgomery (2002).

CART analysis is used to better understand the relationship between factors and responses and to assist in the selection of the next design region to explore. The responses are prioritized resulting in a CART for each response of interest where new and sometimes conflicting design regions are identified. Schamburg (2004) applies this meta-modeling technique in his dissertation where an enormous number of runs were

completed and efficiency was not addressed.

Space-filling designs are often analyzed through kriging (spatial correlation or Gaussian) models. (Crino, 2006) When the simulation is deterministic and the function is known, kriging gives an exact interpolator which results in the kriging prediction equal to the observed output.

Neural networks are mathematical models inspired by biological neural networks of the central nervous system. Nodes are connected together to form a network where a function approximation is capable of learning from observed data. Training an algorithm however on a new data set is not efficient and may require a significant amount of experimentation. (Stern, 1996) Neural networks are applicable when prediction is more important and there is no mathematical formula that relates the input factors to the output variables. (Samoilenko and Osei-Bryson, 2010)

Multivariate adaptive regression splines (MARS) was introduced by Friedman (1991) and is a linear model with a forward stepwise algorithm to select model terms followed by a backward procedure to prune the model. The approximation bends to model curvature at "knot" locations, and one of the objectives of the forward stepwise algorithm is to select appropriate knots. After selection of the basis function is completed, smoothness to achieve a certain degree of continuity is applied. (Crino, 2006)

Gaussian process models are interpolation models that develop a probability model through Bayesian prediction in order to model the relationship between the inputs and outputs. Gaussian process models are well suited for computer experiments where the response and predictors are continuous. There is much literature on the use of Gaussian process models for deterministic computer simulations. (Santner, et al., 2003)

The parameters of a Gaussian process model are fit through maximum likelihood and use the product exponential covariance or correlation function with a power of two as the estimated model. The assumption in using this model is that the response Y is normally distributed with mean μ and standard deviation σ^2M. The M matrix is composed of elements as shown in Equation 3.3.

$$m_{ij} = exp\left(-\sum_k \theta_k \left(x_{ik} - x_{jk}\right)^2\right) \tag{3.3}$$

Where

θ_k =correlation parameters

x_{ik} and x_{jk} = elements of the covariate matrix **X**

In order to predict the value of $y(x_{new})$, the process must exhibit some regularity over the design space therefore a space-filling design is preferred. All pairs of locations x_1 and x_2 must have a common orientation so that the inter-point distance will have the same covariance. "The Gaussian process model is an example of a probabilistic non-parametric model that provides information about prediction uncertainties which are difficult to evaluate appropriately in non-linear parametric models." (Kocijan, Murray-Smith, Rasmussen and Girard, 2004) Santer et al. (2003) states that deterministic computer experiments are well suited for a Gaussian process model. The use of Gaussian process models will be explored for the stochastic computer experiments examined in this dissertation.

It is important to select a model that will appropriately fit the data. This model can be used for predicting the response without having to conduct any replications to find the

results. The more appropriate the model fit the better the prediction results. In general, it is recommended that the simplest model be used to fit the data in order to avoid over fitting the data. This also reduces the level of error and can increase statistical power. This research will use Gaussian process modeling in an effort to gain efficiencies by developing good response prediction formulas. These response prediction formulas will be optimized using SA and multiple objective optimization to find factor settings that optimize the responses. This will result in a decrease in the number of trials that will need to be conducted to find these same factor settings. Both Gaussian process modeling and SA are part of the methodology presented here and are discussed in more detail in section 3.3.

3.2 Evolutionary Operation (EVOP)

The concept of EVOP was introduced by Box and Draper in 1957 (Meyers et al., 2002), although Box et al. (1969) state that EVOP was actually established as early as 1954. EVOP is a method permanently applied to a full scale process in order to continually monitor and improve the operating parameters of the process in order to produce improved operating conditions. EVOP is a systematic way of conducting experimentation that is similar to the ideas of sequential analysis. Box el al. (1969) cite that one drawback of EVOP is that it must rely on many replications to separate the signal from the noise. Perhaps by employing an efficient OLHD/NOLHD a tradeoff can be made between the increase in the number of replications required by the EVOP process and the stochastic nature of the computer simulation. Additionally, applying power analysis and interim data monitoring and analysis ensures that only the necessary number of replications needed for statistical analysis will be completed thus helping to

gain efficiencies. The process of power analysis and interim data monitoring and analysis is discussed in detail in section 3.4.

3.2.1 EVOP Techniques

EVOP can be applied to a number of process variables, however only two or three variables are usually considered. While examining all the variables at one time is an efficient methodology, Box et al. (1969) discourages the use of too many variables because "in actual operation, the making of changes on too large a number of variables can merely lead to chaos". Given the improved computing capabilities of today, this research will explore a large number of variables while employing EVOP and demonstrate that it is possible to go beyond the two or three variables recommended by Box, et al. (1969).

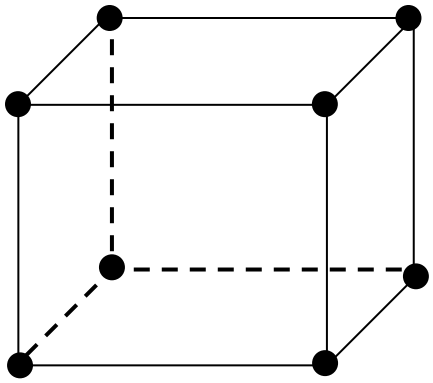


Figure 3.7 Location of the design points for a 2^3 factorial depicted on the unit cube.

Typically EVOP designs use a 2^k factorial design as shown in Figure 3.7. Box and Draper (1969) discuss the advantages of using simple factorials during the EVOP process and layout the steps for conducting the process. Given that the analysis of the EVOP

process was conducted manually when Box and Draper wrote about this procedure in the 1950's and 1960's, it is apparent why this conclusion is made about using simple factorials.

In an attempt to explore more complex EVOP procedures, other EVOP techniques employed in an attempt to optimize include Rotating Square EVOP (ROVOP), Random EVOP (REVOP) and Simplex EVOP. (Box et al., 1969) RSM is discussed in conjunction with EVOP to "be too complicated to be used under the circumstances in which EVOP is employed but can be utilized if a proficient technician is available". (Box et al., 1969) With today's computing resources this challenge is not as daunting as originally proposed in 1954. These more complex EVOP techniques are described next.

The ROVOP (Figure 3.8) was developed by Dr. Edwin Harrington in 1964 for use with only two variables. Cycle one consists of square 1. If no improvements are detected, then you rotate the square 45 degrees from the original square and fit square 1 into square 2. If no improvements are detected then cycle 3 repeats in the same manner as cycle 2. The process repeats itself until an improvement is detected and the process moves in the direction of the improvement. (Box et al., 1969) REVOP by Satterthwaite (1959), makes use of completely random points rather than a planned factorial design. (Box et al., 1969)

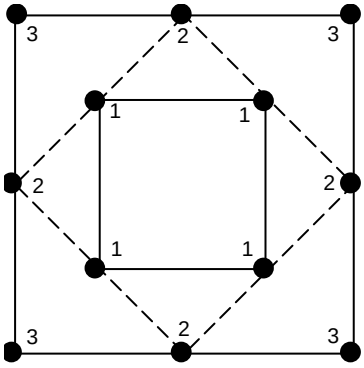


Figure 3.8 Location of the design points on the unit square for three successive ROVOPs.

Meyers and Montgomery (2002) discuss a Simplex EVOP process (Figure 3.9) however it is not a recommended design. In the Simplex EVOP, the initial design forms an equilateral triangle. A new vertex is added opposite the lowest yield forming a new equilateral triangle after each cycle. The process continues until an improved response is no longer achieved or the process results in a return to the same vertex repeatedly. (Spendley, Hext and Himsworth, 1962)

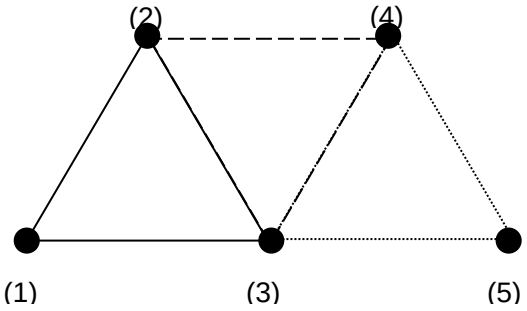


Figure 3.9 Location of the design points in a Simplex EVOP design in 2 Dimensions.

3.2.2 The EVOP Process

An experimental design is selected and the process or simulation is run at these design points and is called a cycle. After at least two cycles, the range of produced values provides the estimate for the standard deviation (σ) which is updated with each subsequent EVOP cycle. A cycle can consist of a single simulation run, multiple simulation runs or batched simulation runs. These runs are all conducted at the same operating conditions leading to an average response representative of the operation at those fixed conditions. It may be necessary to complete some initial simulation runs to get an estimate of σ. In an industrial process, this initial estimate of σ can be obtained from plant records, therefore it may be possible to obtain an estimate based on previous simulation runs or pilot runs during the course of the same study. Variable screening as discussed in the Chapter 3.1.1 can help narrow the pool in determining which factors are significant. (Butler, 2001)

As EVOP is used throughout the study and goes through these cycles and phases, estimates of the effects, variables, function and the region will begin to take shape and morph as the process continues in the direction of improvement using the statistical techniques of RSM and other methods of searching the design region. Box and Draper caution about making comparisons of EVOP to optimization techniques because EVOP "is an experimental technique for seeking the preferable" factor settings that achieve desired response objectives. Optimization is therefore not a goal of the EVOP process. Rather than seeking the optimal solution, EVOP is concerned with finding solutions that are an improvement over current operating conditions and are conducted sequentially to continue to find further improvements. In EVOP, it is desirable to find continuous

improvement without the pressure of promising to find the absolute best settings. EVOP usually starts with a process that has variables of "interest". The EVOP process begins by examining these variables. As the EVOP program evolves, "the dimensionality of the experimental space develops in ways that could not have been predicted" initially. (Box et al., 1969) It may be possible that as the process moves in the direction of improvement, the original limits of the variables are tested and result in going beyond the preconceived limits. Additional variables of interest may even be found that weren't initially thought to be significant.

New literature on EVOP is basically non-existent. Some believe that EVOP was abandoned as a result of RSM and Taguchi methods made popular in the 70's. (Holmes and Mergen, 2006) When EVOP was first introduced, the method was conducted using a worksheet methodology with manual calculations. If computers were available, the EVOP calculations could be handled with computer automation. This dissertation will seek to develop new interest in EVOP, while analyzing more than three factors, through the use of sophisticated experimental designs and aided by today's use of computers.

3.3 Optimization

While EVOP was not intended as an optimization technique, this research will leverage optimization techniques to find the best factor settings that optimize the response prediction formulas found through Gaussian process modeling within the EVOP process. There are several techniques that can be used. (Ghani, 1995, Barton and Ivey, 1996, Humphrey and Wilson, 2000; and Meyers et al., 2002) The method of steepest ascent, tabu search (TS), simulated annealing (SA), the genetic algorithm (GA), and the Nelder-Mead simplex search (NMSS) algorithms are examples of local and global search

techniques. (Shang, Wan, Fromherz and Crawford, 2001)

The method of steepest ascent is one of the methods which use the gradient of the function to determine the direction in which to search for maxima. One disadvantage of this method is in the case of a discontinuous function where the derivative cannot be computed and no new search direction can be determined. This method typically works well only on functions that are unimodal. On multimodal functions this method may not be able to escape a local maxima to check for other maxima in the function to find a global optima.

Tabu search (TS) is a pseudo random, deterministic search algorithm. The TS algorithm has memory to store past solutions and the history of the swaps that were made. A pointer indicates past changes and keeps track of the swaps. This memory forces the search to explore new areas within the space and is not constrained to local optimums like the method of steepest ascent. (Michalewicz and Fogel, 2000)

SA was inspired by the heating and cooling of metals to reduce defects and is a pseudo random, stochastic search algorithm that finds optimal values numerically. It is a search method as opposed to a strictly gradient based algorithm. It chooses a new point where all uphill points are accepted and while some downhill points are also accepted depending on probabilistic criteria. This iterative algorithm randomly generates a solution and then moves probabilistically to new solutions. SA may move to solutions of lesser quality so that the search can move to new areas thus allowing a better quality solution. Moves are done probabilistically according to a temperature procedure, where the higher the temperature value the more likely the lesser solution is accepted. SA has many strengths, among them are 1) it can deal with highly nonlinear models, chaotic and noisy

data and many constraints, 2) it is a robust and general technique and has main advantages over other local search methods, 3) it is flexible and has the ability to seek out global optimal solutions. (Michalewicz et al., 2000) Figure 3.10 outlines the heuristic for conducting SA.

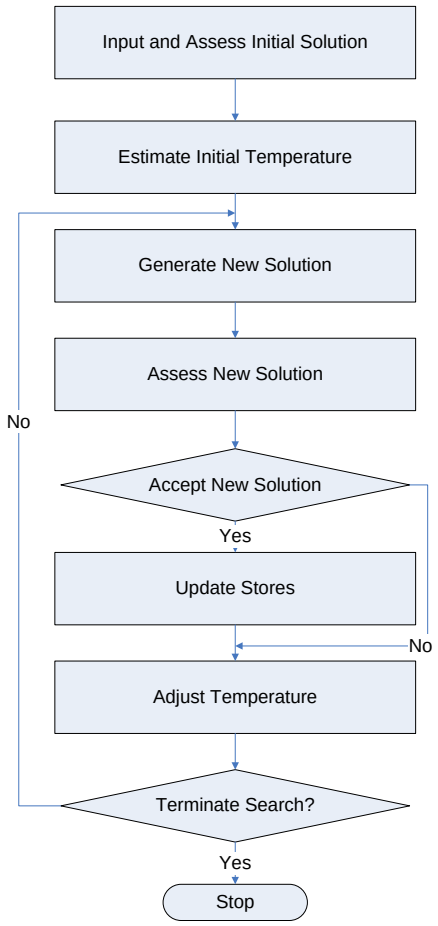


Figure 3.10 The steps for applying the simulated annealing heuristic to find the global optimal solution to a problem.

A genetic algorithm (GA) is a global search technique that finds optimal solutions and is derived from the processes of natural evolution. A GA allows for a large number of parameters and can deal with noisy data. Weaknesses of a GA's are their inaccuracies

and that they are computationally intense to implement. (Michalewicz et al., 2000)

The Nelder-Mead simplex search (NMSS) method was originally designed for unconstrained optimization of deterministic functions but it can also be applied to optimize stochastic simulations. (Barton et al., 1996) According to Barton and Ivey this is the most popular direct search method. Barton et al. (1996) describe 1) the original Nelder-Mead method, 2) the Nelder-Mead Simplex Method, and 3) a modification to the Nelder-Mead Simplex Method which shows greater improvements over the two previous methods.

The advantages of SA over the other methods discussed, especially the ability to escape local optima, make it an ideal choice for use within the ECEM. Once the factor settings that result in an optimal response have been found using SA, they are used in the computer simulation experiment to collect data and are the "treatment" under investigation. Initially *n* replications are completed to collect an adequate amount of data to base the first interim analysis. The variable *n* is determined through the use of power analysis which is completed during the development of the monitoring plan. These concepts are discussed in Section 3.4.

3.4 Clinical Trial Analysis

There are similarities that can be made during the course of planning and executing a clinical trial and planning and executing a stochastic simulation experiment. The selection of participants, monitoring the results and analyzing the data are all common tasks. While the simulation experiment under study may not be a matter of life and death, it is of great concern to be efficient especially when the simulation is complex and resource intensive. Efficiencies during the conduct of clinical trials are a result of

planning for the experiment, statistically monitoring the data collection process and conducting analysis so that as soon as results become apparent during the conduct of the trials, the experiment can either be stopped early or a decision can be made to continue the experiment based on the results gathered. This plan and how the data are monitored and analyzed are discussed in Section 3.4.1.

3.4.1 The Monitoring Plan and Protocols

The conduct of interim data monitoring and analysis is based on the monitoring plan or protocol developed. This plan specifies: 1) The hypotheses to be tested; 2) significance level α (Type I error probability); 3) maximum number of planned analyses so that the maximum sample size can be determined; 4) best sample size using power analysis and, therefore, determining the Type II error probability, $ß$; 5) an estimate of the effect of the protocol on the outcome and its associated *p-value*; and 6) the type of stopping boundary to be applied. (Everitt, 2004) These boundaries are summarized in section 4.2.7 and are used for testing hypotheses and summarizing evidence simultaneously. The process is fully sequential as the evidence is assessed after each patient/observation or groups of patients/observations. "…a clinical trial is stopped as soon as enough information is accumulated to reach a conclusion …whether the drug is superior or inferior to the control". (Dmitrienko et al., 2007)

Power analysis is used to compute sample size through the proper design of the experiment and has been in use long before clinical trial analysis was developed. One must select the appropriate power for the situation as too high a power can result in trivial effects becoming too significant during testing and too low a power will cause the erroneous rejection even though the data fit perfectly. Section 5.1 discusses the properties

of power analysis and how it leads to efficiencies through determining the appropriate sample size.

3.4.2 Interim Data Monitoring and Analysis

Medical researchers conduct “clinical trials which are a type of research study that uses human volunteers to test new methods of screening, prevention, diagnosis, or treatment of a disease to answer specific health questions”. (Jennison et al., 2006; Karrison et al., 2003) Statistical methods have been developed in the biostatistics field to reduce the number of required subjects, save time and money and to improve the outcomes of treatment protocols. (Hagino et al., 2004) Group sequential methods (GSM) were developed to allow early termination of studies on the basis of interim comparisons known as interim efficacy analysis. (Burington et al., 2004) In clinical trials, interim data monitoring and analysis is performed after study groups of subjects have completed the trial. GSM may allow early closure of one or more treatment arms based on the results of interim analyses. The GSM approach defines a critical value to be used for each interim analysis so that the overall Type I error rate is maintained at a level defined in the clinical trial protocol. By enabling early closure, GSM protects patients from unnecessary exposure to a potentially unfavorable treatment. (Whitehead, 1999) During the course of the sequential testing the trial is ended when sufficient data have been collected to determine if one treatment protocol is better than another protocol or if a protocol is found to be unsafe to the patient.

Sequential data monitoring strategies consist of repeated significance tests and stochastic curtailment tests. Repeated significance tests are used a great deal in clinical

trials during interim data monitoring and analysis. These tests make use of error spending functions which determine the shape of the stopping boundaries. Error spending functions are discussed in detail in section 3.4.3. Stochastic curtailment tests focus on futility and are based on the likelihood of observing a given treatment effect based on predictive inference. These tests span the realm of frequentist in nature to Bayesian while some are mixed frequentist and Bayesian. (Dmitrienko et al., 2007) The repeated significance test may use a variety of possible error spending approaches such as design-based, ten-look, Lan-DeMets, the Jennison-Turnbull, and the Hwang-Shih-DeCani family of error spending functions. Stochastic curtailment futility rules, based on conditional power, were established by Lan-Simon-Halperin and Pepe-Anderson-Betensky. Futility rules, based on predictive power, were established by Dmitrienko et al. (2007) and Wang (2003). Dmitrienko et al. (2007) describes the use of each of these methods in great detail.

There are several concerns with interim data monitoring. First, the preservation of Type I error can be an issue but can be controlled through the use of an alpha spending function which will then specify how the Type I error will be spent. Second, there may be a penalty for taking interim looks such as a stricter p-value required or a larger sample size. Again, this can be controlled by the choice of error spending function used. Lastly, if the trial requires a rigid schedule with respect to the number and timing of interim looks, than the timing of the interim looks may need to be specified over the course of the trial as opposed to being determined statistically.

3.4.3 Stopping Criteria

There is much literature on the use of stopping criteria for an experiment. Some

criteria are simply when a threshold has been reached or when a goal such as a percent achievement has been reached. Using interim data monitoring and repeatedly testing data can result in an inflated false positive error rate if not conducted properly. Clinical trials use group sequential methods and the application of stopping boundaries. An error spending function, a frequentist approach, is used to define the stopping boundary in order to control this error rate. The selection of an error spending function should be appropriate to the needs of the trial. The error spending function should be conservative when its use is not intended for early efficacy stopping but is appropriate for administrative looks, futility stopping or sample size adjustment. The error spending function should be aggressive primarily when early efficacy stopping is desired. The important quantities to fix in the design are the choice of error spending function and the maximum sample size. Formal futility boundaries are an alternative approach. The β-spending function allows for the preservation of the Type II error. "Interim data monitoring allows us to estimate the current information about the treatment effect, δ, from the actual data of the trial, re-estimate sample size, and preserve the power of the study". (Mehta, 2004)

Error spending functions determine the rate at which Type I or Type II errors are spent. This is achieved by designing the stopping boundary in the design stage based on an error spending function. This function is used to generate adjusted critical values at the specified points in time. Some common stopping boundaries and error spending functions used in clinical trials are reviewed. The following variables are defined:

$\alpha(t)$ = Type I error spending function

t = Total sample size

α = Prespecified Type I error rate

$\Phi(x)$ = Cumulative probability function of the standard normal distribution

z = z-statistic

ρ = Shape parameter (Can only be a positive numbers, larger values result in a lower Type I error spending rate at the beginning of the experiment.)

1) Design-based error spending functions allow for the selection of the number and timing of the interim analyses.

2) Ten-look error spending functions are simply ten equally spaced interim analyses.

3) The O'Brien-Flemming boundary is characterized by a non-constant critical value applied at each interim analysis and is more desirable since it is conservative in continuing the clinical trial.

4) The Pocock boundary maintains a constant critical value at each interim analysis and requires the largest sample size in order to achieve the desired power. The p-value used depends upon the number of interim analyses.

5) Lan-DeMets 1 (1983) generates conservative boundaries that are similar to the O'Brien-Fleming through applying the error spending function in Equation 3.4.

$$\alpha(t) = 4 - 4\Phi\left(\frac{z\alpha/4}{\sqrt{t}}\right) \quad (3.4)$$

6) Lan-DeMets 2 (1983) generates aggressive boundaries that are similar to the Pocock (1977) boundaries through applying the error spending function in Equation 3.5.

$$\alpha(t) = \alpha \log\{1 + (e - 1)t\} \quad (3.5)$$

7) Jennison-Turnbull error spending functions are summarized in Equation 3.6. Note that when ρ = 3 this error spending function is identical to the O'Brien-Flemming plan and when ρ = 1 approximately, this error spending function is identical to the Pocock plan.

$$\alpha(t) = \alpha t^{\rho} \tag{3.6}$$

8) Hwang-Shih-DeCani error spending function is described in Equation 3.7.

$$\alpha(t) = \frac{\alpha(1-e^{-\rho t})}{(1-e^{-\rho})} \text{ if } \rho \neq 0 \text{ and } \quad \alpha(t) = \alpha(t) \text{ if } \rho = 0 \tag{3.7}$$

Every time data are evaluated to determine if the clinical trial protocol should be halted, the chance of falsely rejecting the null hypothesis is introduced or a chance of introducing a Type I error. If the data is looked at multiple times, and α = 0.05 as the criterion for significance, then there is a 5% chance of stopping each time. Heuristically, a "statistical boundary" is established and the trial is stopped if that boundary is crossed. Other statistical boundaries are the Pampallona, Tsiatis and Kim, and Lan-Simon-Halperin. (Dmitrienko et al., 2007) Application of an error spending function to a statistical boundary was introduced by Lan-DeMets, Jennison-Turnbull and Hwang-Shih-DeCani and is discussed in more detail in Everitt (2004), Dmitrienko et al. (2007) and Cook et al. (2007). This research will apply stopping boundaries that will use efficacy and futility monitoring simultaneously.

3.5 Frequentist and Bayesian Views

The frequentist view for interpreting probability is defined as the limiting frequency of occurrence of an event in an infinite number of trials. This approach avoids the use of prior probabilities and thus avoids the use of Bayes' rule for the purpose of assigning probabilities to parameters. θ is considered fixed in the frequentist view while it is a random variable in the Bayesian view. Frequentists focus on planning an experiment and choosing an appropriate sample size.

Bayesian statistics attempts to treat all statistical inference as a probabilistic inference. The Bayesian view of probability is that the probability of an event occurring is subjective based on a personal belief of that event occurring. The Bayesian method implements the notation using Bayes' rule.

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \tag{3.8}$$

Where p(θ) is the prior probability, p(θ|x) is the posterior probability and p(x|θ) is the likelihood function.

Clinical trial statistics deal primarily in the frequentist view although new research in the conduct of clinical trials is also using the Bayesian view or in some cases mixing both Bayesian and frequentist views. The methodology presented here leverages both frequentist and Bayesian views in order to gain efficiencies. Discussing frequentist and Bayesian views are addressed here because in the past statistician seemed to take a stance in one view or the other and stuck with it. In today's complex world it is no longer sufficient to be versed in one view or the other. Today's modern problems benefit from

the application of both the frequentist and Bayesian approaches in order to find solutions and this research has applications in both the frequentist and Bayesian views.

3.6 Summary

This section reviewed literature on experimental designs and model fitting, EVOP, global optimization techniques, and statistical methods used during the conduct of clinical trials. Sufficient evidence exists to support the use of OLHDs/NOLHD's during an EVOP scheme that is designed to analyze large, complex, stochastic simulation experiments with multiple responses especially given the improved computing capabilities of today's computers. The Gaussian process model is a powerful interpolation tool for non-linear model fitting and has been cited as a good tool for deterministic computer experiments. The use of SA as a global optimizer and the application of power analysis and interim data monitoring and analysis are tools that can help gain efficiencies during the experimentation process. Each of these tools reviewed provide a benefit that could also be advantageous to improving efficiency in stochastic simulation experimentation.

The literature reviewed lacks a comprehensive plan for optimizing stochastic computer simulation experiments. Most comparable studies focus on one aspect or another within the experimentation process working with computer experiments. Saab and Rao (1991) developed the stochastic evolutionary (SE) algorithm which uses local and global searches within an inner and outer loop heuristic for an optimal design. Jin, Chen and Sudjanto (2003) extend the SE algorithm with the enhanced stochastic evolutionary (ESE) algorithm which finds an efficient global optimal search algorithm by constructing an optimal LHD and evaluates optimality criteria. Chantarawong,

Rungrattanaubol and Na-udom (2012) further extend ESE with their “Enhancement of ESE” called EESE where SA is applied.

This research proposes an end to end methodology for completing a large stochastic computer simulation experiment efficiently. The use of EVOP with OLHDs/NOLHDs is not discussed in the current literature and the application of clinical trial statistical tools, in particular the unique stopping boundaries and error spending function approach found in clinical trials is not found anywhere outside the clinical trial literature. This methodology is unique in that it walks the experimenter through the process of determining how many replications are needed, use of OLHDs/NOLHDs during the EVOP process and application of clinical trial stopping boundaries and error spending functions to determine whether or not to stop the experiment or to continue based on the planned protocol.

The next chapter describes how Gaussian process modeling, SA, and interim data analysis with clinical trial type stopping boundaries are employed in a methodology for efficient stochastic computer experiments and chapter 5 discusses the theoretical properties of the methodology.

4 The Efficient Computer Experimentation Methodology (ECEM)

This section lays out the steps to the ECEM and addresses the tools and assumptions made in conjunction with using the methodology.

4.1 The Process

The following methodology is developed to pursue efficient stochastic computer simulation experimentation. This approach consists of the best procedures to gain efficiencies while achieving good results. The methodology consists of first planning the experiment through the development of the interim data monitoring plan. This plan determines the number of interim analyses, the maximum sample size needed, and states the hypotheses to be tested. Next, the OLHD/NOLHD is selected based on the range of feasible factor settings and the necessary replications are completed. A Gaussian process model is fit and the response surfaces are generated. SA and multiple response optimization are applied to the response prediction formulas and these factor settings are analyzed in the response surface plots to select regions of interest that optimize all responses. This process continues while conducting interim data monitoring and analysis to determine if and when to terminate the experiment. In the case of a simulation experiment not applied to a clinical trial, efficacy is defined as the desired improvement in the response as a result of the treatment. Efficacy is given by:

$$P\{Z_k > u_k \text{ for any } k = 1,\ldots,m \mid \mu_1 - \mu_2 = 0\} = \alpha \text{ (reject } H_o) \quad (4.1)$$

Where

P = Probability

Z_k = Test statistic on the kth look

u_k = Upper stopping boundary on the kth look

k = Number of times the data is looked at

m = Number of the last look

α = Prespecified Type I error

H_o = Null hypothesis

μ_1 = True Mean for treatment #1

μ_2 = True Mean for treatment #2

In this case, there is a treatment difference and the treatment that yields the better results should be used to continue the experiment and seek further improvements. Futility is given by:

$$P\{Z_k < l_k \text{ for any } k = 1,\ldots,m \mid \mu_1 - \mu_2 \neq 0\} = \beta \text{ (do not reject } H_o) \quad (4.2)$$

Where

l_k = the lower stopping boundary on the kth look

β = Prespecified Type II error

In this instance, there is not a treatment difference and the new treatment applied did not result in a statistical improvement over the current treatment. Stop the trial so as not to continue to waste resources.

The detailed steps to the general methodology are:

1) Develop the monitoring plan or protocol. (Includes doing the power analysis, determining *n* and establishing the stopping criteria for the interim analysis). If current factor settings exist these settings are treatment #1, go to step 8.

2) Select the range of factor settings to explore. (The DOE)

3) Set up and run the simulation given the EVOP experimental design

chosen. (constitutes a complete cycle)

4) Run *x* cycles based on time and budget constraints of the EVOP experimental design. (constitutes the first phase)

5) Fit model to the data and estimate the parameters in the Gaussian process model.

6) Apply simulated annealing to optimize each of the responses given the response surfaces from 5 and apply multiple response optimization.

7) Examine the response surface of the design and select a new range for the factor settings based on these results.

8) Repeat steps 2 – 6. The new factor settings found at step 6 constitute Treatment #2.

9) Given Treatment #1 and the new factor settings found as Treatment #2, conduct *n* replications of each treatment as prescribed in the monitoring plan.

10) Conduct interim analysis as prescribed in the monitoring plan until the stopping criteria has been met. Once the stopping criteria has been met, assign the better of the two treatments to Treatment #1 and go to step 2.

Figure 4.1 summarizes the heuristic for the methodology. The boxes in yellow and orange are the steps that constitute Phase P and Phase P + 1 respectively of the EVOP process. The blue boxes are the steps where the concepts from examining clinical trials are completed or where the data is examined to determine if the experiment should continue.

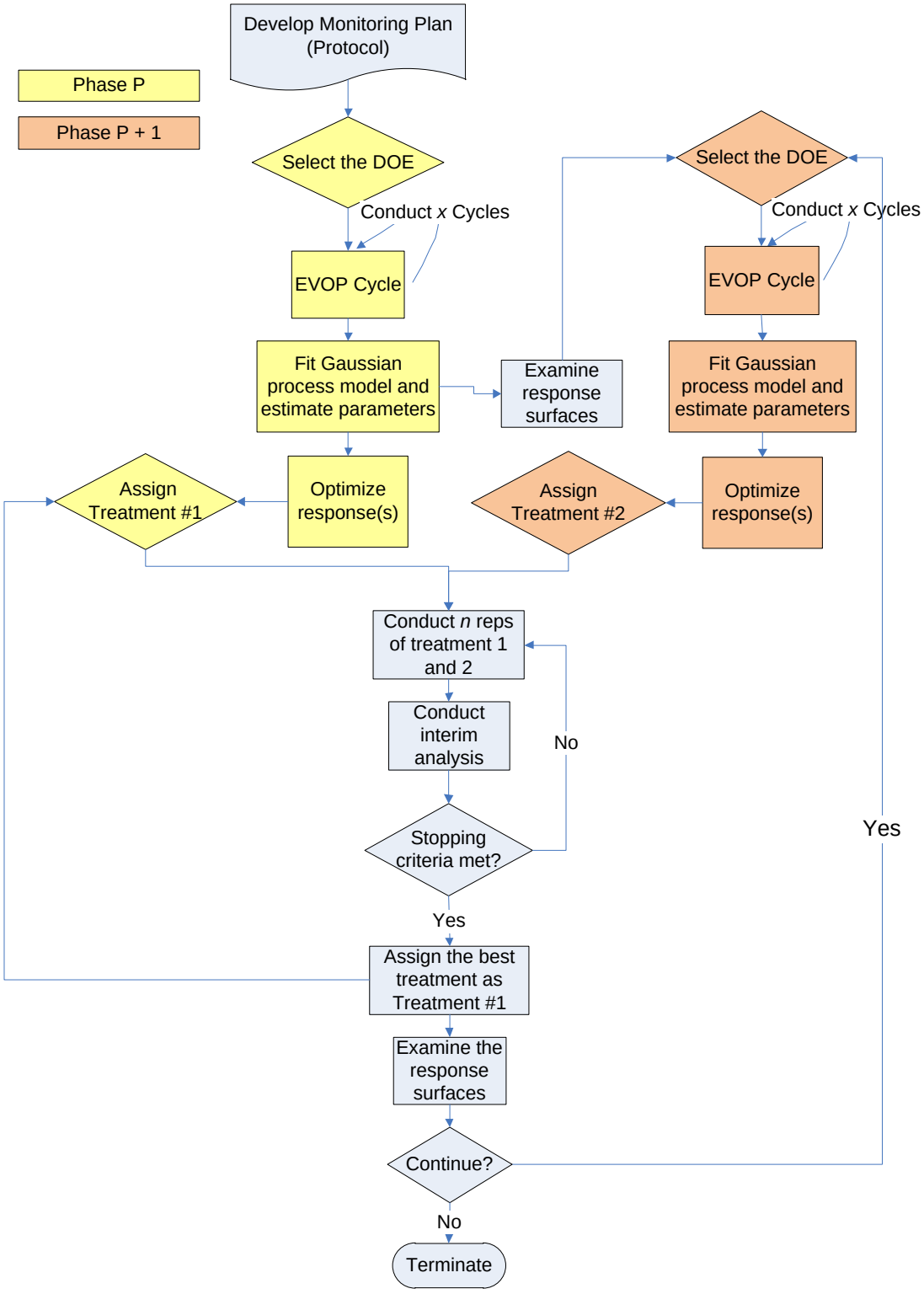


Figure 4.1 ECEM for Analysis of Stochastic Computer Simulation

As the methodology repeats, the design region morphs as depicted in Figure 4.2. The interim analysis completed determines whether or not there is a statistical difference between the two treatments and if the search of the design region for additional improvements to the responses should be continued.

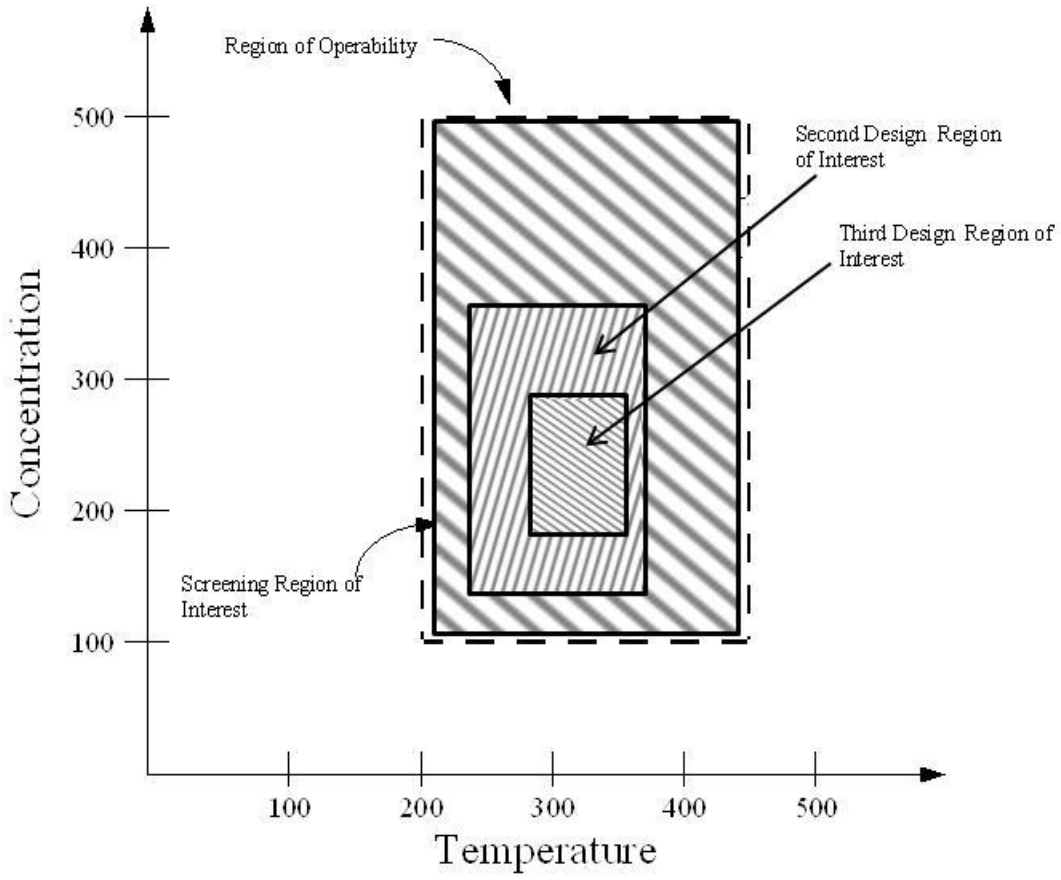


Figure 4.2 Morphing of the Design Region

4.2 Tools and Assumptions

The tools used within the methodology consists of: sample size determination with power analysis, hypothesis testing, DOE and RSM, sequential analysis using EVOP, model fitting with Gaussian process modeling, optimizing the response prediction formula, and interim data monitoring and analysis to determine if the stopping criteria is

met. The following assumptions are made with respect to the application of the ECEM and each tool inherent in the methodology.

4.2.1 Sample Size Determination and Power Analysis

First, to determine the sample size, the Type I error is assumed and the effect is estimated to be low, medium or high. Section 5.1 discusses sample size determination and effect size through the use of power analysis. Once results are obtained, the estimated effect can be checked to determine if the original estimate was accurate. If the effect estimate is not correct, the sample size can be recalculated based on the actual effect. If the new sample size is larger, the trial should continue using the larger sample size. The Type II error is calculated using power analysis.

"Power is a function of the significance level, reliability and variability of the sample data and the size of the treatment effect". (Everitt, 2004) In comparing the effectiveness of two treatments, assume the response is normally distributed with expected values of μ_1 and μ_2 respectively. The outcome variable is assumed to be normally distributed and has constant variance. If this is not the case, transform the data. The most common transformation that can be applied is logistic regression which will result in a more realistic assumption about the distribution.

4.2.2 Hypothesis Testing

The assumptions for hypothesis testing and data analysis are that randomization is met, the experiment is reproducible and the data meet the normality assumption where decisions are based on tests of hypotheses as well as the reliability of estimates where confidence intervals are used**.**

Assume the state for H_o and H_1. It will be assumed here that:

H_o: $\mu_1 = u_2$ (no treatment difference) (4.3)

H_1: $\mu_1 \neq u_2$ (4.4)

In hypothesis testing: H_o: $\mu_1 = \mu_2$ has three assumptions:

1) There are two independent populations where equal sample sizes are drawn. This is important in having homogeneous variance.
2) The dependent variable is continuous.
3) All samples are randomly selected from the population and randomly assigned to the different condition. Each of them appears in one and only one combination of conditions and the samples come from normal distributions. In conducting the simulation experiment, the OLHD/NOLHD schemes are employed as the sampling methodology.

Clinical trials are often times not reproducible, however this is not the case with computer experiments where the results can be reproduced and replications can be completed. The basic assumptions within clinical trials are as follows: 1) the random samples from treatment one, $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ are from a population with mean μ_1 and variance σ_1^2 2) the random samples from treatment two, $\mathbf{Y} = \{Y_1, Y_2, \ldots, Y_n\}$, are from a population with mean μ_2 and variance σ_2^2 and 3) the samples from treatment one and two are independent of one another. It is assumed that each treatment has an equal number of samples taken. In order to estimate μ_1 - μ_2, use $\bar{\mathbf{X}} - \bar{\mathbf{Y}}$, the difference between the corresponding sample means, as a natural estimator. It is proposed that the expected value of $\bar{\mathbf{X}} - \bar{\mathbf{Y}}$ is μ_1 - μ_2, so $\bar{\mathbf{X}} - \bar{\mathbf{Y}}$ is an unbiased estimator of μ_1 - μ_2. The standard

deviation of $\overline{\mathbf{X}} - \overline{\mathbf{Y}}$ is given by

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \tag{4.5}$$

The Central Limit Theorem (CLT) provides support to the normality assumption. The CLT states that as the sample size *n* becomes sufficiently large, the sampling distribution of the mean tends towards a normal distribution. The problem with applying the CLT is determining when *n* is large enough. If the underlying distribution is close to a normal density curve than the CLT approximation will be sufficient for a small *n*. If the distribution is far from normal, then a large *n* is required. Therefore, the probabilities of the errors of Type I and Type II are not affected severely by moderate departures from normality.

In the case where the sampling distribution of the mean is not normal, a transformation of the data is done to ensure near normality for the distribution of the transformed data. In this research, power analysis is done to ensure that *n* is sufficient. Power analysis is accomplished as part of the planning phase to the experimentation process and therefore *n* is determined such that these assumptions are not violated. Significance level, α, is specified along with the effect size to determine the appropriate sample size that supports the assumptions discussed above. This is discussed in more detail in section 5.1.

4.2.3 DOE and RSM

DOE is a systematic approach to examining the effects of many factors on a

response(s). An empirical model is fit and the analysis of variance (ANOVA) helps determine what the significant factors are on the response(s). RSM is a "collection of statistical and mathematical techniques for developing, improving and optimizing processes." (Meyers et al., 2002) The analysis is enhanced through the use of a graphical representation of the relationship between the factors and the response(s).

DOE and RSM assume that there is white noise. White noise implies that the error is normally distributed with constant variance. This assumption is often unrealistic as the experimental output may have variability that is not constant when the input combination changes. It is important to test that the white noise assumption holds. This is done by examining a plot of the residuals and verifying that the distribution of the residuals is normal. "If the assumption does not hold, the analysis methods may need to be adapted using, for example, weighted least squares (WLS), which weights the experimental outputs based on their variability. Further analysis is also possible through computer-intensive methods, such as 'bootstrapping' and 'cross-validation' ". (Kleijnen, 2008)

The errors are also assumed to be identically and independently distributed. That is, the error of an observation is not correlated with that of another error. This can be achieved with the use of proper randomization of the design points. The use of OLHD\NOLHD achieves this result and is shown is section 5.2.1.

Employing a LHD is an ideal design to ensure the proper random sampling and assists in the reduction of errors through the employment of a properly planned randomized design. The advantages of LHDs help gain efficiencies as they require fewer subjects and eliminate systematic biases through counterbalancing. If necessary, the removal of error variances can be accomplished through two-way blocking. (Hoshmand,

2006)

Homogeneous variance between treatment groups is assumed. Applying the treatments uniformly to the treatment groups leads to the tendency to stabilize the variances between treatment groups. Minor departures from these assumptions do not greatly impact conclusions drawn. Should major departures from these assumptions exist, again a data transformation is necessary. In increasing the accuracy of the experiment, the number of replications can be increased which will also decrease the error variance.

4.2.4 Sequential Analysis and EVOP

Sequential analysis is the statistical theory and methods where the sample size is determined based on the accumulation of data. EVOP is a sequential analysis procedure in the area of industrial processes where DOE is incorporated into the sequential procedure. The DOE assumptions therefore apply to the EVOP process.

Group sequential methods (GSM), developed for clinical trial analysis, are where groups of accumulated data are analyzed. GSM use can be traced back to the 1920's (Ghosh, 1970) and is utilized in the EVOP process as described by Box, et al. GSM were later employed in multi-stage plans developed by Columbia University Research Group to accept or reject batches of items in proportion to defective items found. This became the United States Military Standard MIL-STD105D. GSM in clinical trials didn't come into favor until the 1970's. Pocock is credited with energizing the use GSM's in clinical trials in 1977, followed by O'Brien and Flemming shortly after. (Dmitrienko et al., 2007)

Group sequential tests (GST) are centered on the sum of observations collected and the Central Limit Theorem (CLT) as the observation sums are typically approximately normally distributed. GSTs are applied when a sequence of statistics can be fixed in a

Brownian motion. A Brownian motion is a mathematical model used to describe the random movement of particles in a liquid or gas and this concept is extended to other real world stochastic processes. The accumulation of data in groups and analyzing it after each group is collected is referred to as interim analysis. (Ghosh, 1970)

In the ECEM, interim analyses are accomplished in comparing two treatment designs. Simulation batch runs are completed on each design and the results are compared using the test for hypotheses described in section 4.2.2.

4.2.5 Gaussian Process Modeling

Gaussian processes are a stochastic process that when applied to a sampling function will give normally distributed results. This model is used for deterministic simulation as discussed by Santner et al. (2003) and handles the output as a random multivariate normal stochastic process.

A Gaussian process model is given by: (Jones and Johnson, 2009)

$$y(X) \sim Normal(\mu 1_n, \sigma^2 M(X, \theta)) \quad (4.6)$$

Where

Y is *n x 1* vector of data

X is *n x p* matrix of continuous covariates

μ and σ are scalar mean and variance parameters

θ is a *p x 1* vector of correlation parameters

M(X, θ) is an *n x n* correlation matrix

The model correlation function is

$$m_{ij}(X,\theta) = exp\left(-\sum_k \theta_k \left(x_{ik} - x_{jk}\right)^2\right) \quad (4.7)$$

where $\theta_k \geq 0$.

If $\theta_k = 0$, then the correlation is 1 across the range of the kth factor and the fitted surface is flat. Large θ_k corresponds to low correlation in the ith factor and the fitted function exhibits many humps in that direction. Maximum likelihood (ML) is a generalization of least squares used to estimate parameters from a wide class of models and μ,σ, and θ may benefit with maximum likelihood.

Gaussian process modeling is used to find a response prediction formula. Model fitting assumptions that must hold are that the response(s) must be continuous in order to model continuous predictors. Observing the actual versus predicted plots, goodness of fit can be measured by checking that a linear relationship between the actual and predicted responses exists. In applying Equation 3.3 as the estimated model, it is assumed that the responses are normally distributed with mean μ and standard deviation σ^2**M.** The output of the model report is essentially an analysis of variance table, but the variation is computed using a function-driven method.

4.2.6 Optimization of the Response Prediction Formulas

SA is used to optimize the response prediction formula. When there are multiple responses, multiple response optimization is used in conjunction with SA. It must be

assumed that the response prediction formula is continuous and concave or convex over the interval so that a global optimum can be located. The response prediction formula as a result of Gaussian process modeling has the form (Jones et al., 2009)

$$\hat{y}(x) = \hat{\mu} + r'(x, \hat{\theta}) M^{-1}(X, \hat{\theta})(y - \hat{\mu} 1_n) \tag{4.8}$$

Where, $r_i(x, \hat{\theta})$ is an n x 1 vector of estimated correlations of the unobserved y(x) at a new value of the explanatory variables with the observations in the data, y(x). The form of $m_i(x, \hat{\theta})$ is

$$m_i(x, \hat{\theta}) = exp\{-\sum_{k=1}^{p} \theta_k (x_k - x_{jk})^2\} \tag{4.9}$$

where m has the same form as the correlation matrix. Replacing the x vector with the X data matrix, m becomes M which cancels with M^{-1}. Therefore, $\hat{y}(x) = y$ and the Gaussian process models interpolates the data.

The prediction variance, discounting error in estimating the parameters, is

$$\frac{Var(\hat{y}(x))}{\sigma^2} = 1 - r'(x, \hat{\theta}) M^{-1}(\boldsymbol{X}, \hat{\theta}) + \frac{\left(1 - \mathbf{1}' M^{-1}(\boldsymbol{X}, \hat{\theta}) m(x, \hat{\theta})\right)^2}{\mathbf{1}' M^{-1}(\boldsymbol{X}, \hat{\theta}) \mathbf{1}} \tag{4.10}$$

In the case of multiple objectives, the criteria or objectives are often times in conflict with each other. In this case, the design which is the Pareto optimal design is chosen as the best design. Multiple objective optimization methods include multiple objective linear

or non-linear programming, preemptive optimization, weighted sum and goal programming. (Lee, 2002)

Multiple objective linear programming results in the formation of an efficient frontier. The efficient frontier consists of all the solutions that meet the criteria of the multiple objective problem. This strategy is used when the constraints and objectives are linear. There is not usually a unique solution but rather a preferred solution. Linearity requires the following assumptions. 1) A change in a variable is proportionate to the change in the variables contribution to the value function. 2) Additivity of each term in the value function. 3) The decision variables are non-integer. If this assumption does not hold, then integer programming techniques must be used. 4) The coefficients are known and constant. If linearity does not exist, then non-linear programming is used. (Lee, 2002)

Preemptive optimization considers each objective one at a time based on established priorities. Once the objective is optimized, a bound or optimal objective value is obtained and is used to set that objective to this constant. This process is repeated on all the objectives. The final solution is an efficient point of the original multiple objective model. (Lee, 2002)

Weighted sum strategy converts multiple objectives into a single objective using weights and sums. Again the objectives are prioritized and weighted accordingly and then summed. It is assumed that the single objective value function exists. (Lee, 2002)

Goal programming converts the objective function from a minimization or maximization problem and sets a target value to each objective. These goals have adjustable constraints and can be altered if needed. (Lee, 2002)

Given the many applications of this methodology, it is left to the discretion of the

analyst which multiple objective optimization strategy to use. The selection of the strategies used in this research is identified for each application in Chapter 6.

4.2.7 Interim Data Monitoring and Analysis

In conducting interim data monitoring and analysis, the following assumptions are made. This research will use a sequential design that combines efficacy and futility testing. The O'Brien-Flemming stopping boundary is used for efficacy, upper stopping boundary, and the Pocock boundary is used for futility, lower stopping boundary. A design based error spending function is applied. The use of stopping boundaries and error spending functions were discussed in more detail in section 3.4.3.

The assumptions that were presented with respect to hypothesis testing hold for interim monitoring and analysis with the application of stopping boundaries. Futility and efficacy testing relies on comparing the *Z-statistic* to a critical value. Futility is defined when the *Z-statistic* is compared to a critical value that becomes a lower stopping boundary. Efficacy testing compares the *Z-statistic* to a critical value that becomes the upper stopping boundary. The two designs are compared and the following rules are given: 1) if $Z_k > u_k(\alpha, ß)$, reject the hypothesis H_o, 2) if $l_k(\alpha, \beta) \leq Z_k \leq u_k(\alpha, ß)$, continue the experiment and collect additional observations, 3) if $Z_k < l_k(\alpha, ß)$, fail to reject the hypothesis H_o. Chapter 5 discusses the properties of the methodology in more detail.

5 Properties of the Methodology

This chapter addresses the properties of the methodology which make the ECEM more efficient than traditional methods of computer experimentation analysis. Efficiency is defined as follows. In statistics, efficiency is used to compare various statistical procedures. In experimental design or hypothesis testing, efficiency refers to a measure of optimality. A more efficient estimator or hypothesis test needs less samples to achieve a given performance. In experimental design, variance or mean squared error defines efficiency. In significance testing, sample size to achieve a given power defines efficiency.

There are several techniques used throughout the methodology that lends itself to the efficiency of the methodology. The efficiencies within the methodology come from 1) hypothesis testing and minimizing sample size through power analysis, 2) space-filling experimental design allowing for the inclusion of the global optimum, 3) variance reduction through stratification (OLHDs/NOLHDs), 4) interim data monitoring and analysis, and 5) EVOP and sequential analysis.

Section 5.1 addresses power analysis and determining minimum sample size which is conducted during the development of the monitoring plan. Section 5.2 discusses the selection of a space-filling design, the LHD and its properties and finding the global optimum. This section also addresses the normality assumptions and minimizing variance. Section 5.3 explains RSM with fitting the Gaussian process model in order to obtain the response prediction formula. This formula is optimized to find the best factor settings thereby reducing the number of replications needed. RSM is extended by using SA to optimize the response prediction formula. Section 5.4 addresses the conduct of

sequential analysis (EVOP) and the tools found during the conduct of clinical trial analysis. (Interim data monitoring and analysis) Each of these techniques is discussed in detail and how efficiencies are gained as a result of their implementation. Finally, the methodology is used on known test functions (section 5.5) and on a test function with noise (section 5.6) and compared to SA and GA in order to demonstrate the efficiency of the ECEM.

5.1 Minimum Sample Size

The monitoring plan consists of the following steps:

1. Specify the hypotheses to be tested.
2. Define the significance level
3. Define the maximum number of planned analyses and the maximum sample size possible based on resources.
4. Compute the best sample size using power analysis.

Power analysis is used to compute the minimum number of samples that must be collected in order to detect an effect. The alternative is to calculate the effect based on the given sample size. Power analysis can be done either at the beginning of the study or at the end. This research will conduct power analysis at the beginning in order to determine the number of replications necessary. Power analysis is dependent on 1) statistical significance level desired, 2) the magnitude of the effect and 3) sample size. Detailed discussions about power analysis can be found in Cohen (1988), Murphy, Myors and Wolach (2009).

Significance level is selected based on the strength of evidence the experimenter wishes, i.e. the Type I error, α, and Type II error, β to determine the confidence interval

to base the results. Effect size is the degree to which the observable fact is present in the population. The larger the effect size stated, the smaller the sample size needed to detect it. The effect size, d, can be estimated or if the mean and standard deviation is known, it can be calculated as shown in Equation 5.1. This equation assumes equal variance of the two populations and independent samples.

$$d = \frac{\mu_1 - \mu_2}{\sigma} \tag{5.1}$$

The non-centrality parameter δ is given by Equation 5.2. When determining *n*, δ is found using a table based on the desired power and the experimenter's specified significance level α.

$$\delta = d\sqrt{\frac{N}{4}} \text{ where } n = n_1 + n_2 \tag{5.2}$$

The desired *n* is then found by Equation 5.3.

$$n_{desired} = \left(\frac{\delta}{d'}\right)^2 \tag{5.3}$$

When the mean and standard deviation or variance is not known such as during the planning for an experiment and no prior information is known, the effect size, d, is estimated. Cohen (1988), Dattalo (2008), Murphy, et al. (2009) each addresses effect size, how to estimate it and the meaning of small, medium and large effect size. Sample

size should be appropriate so that statistical analysis of the results is valid.

A power of 0.8 is the typical standard for most practitioners when the primary concern is to correctly reject the null hypothesis. (Gehan and Lemak, 1994) By conducting power analysis, the sample size can be determined so that the statistical results are valid. Note that if the sample size exceeds the maximum sample size possible based on available resources, the power would have to be decreased or additional resources used so that power can be preserved and the maximum sample size can meet the necessary sample size. Power analysis will force the researcher to consider not only if the statistics are significant but also quantifies the strength of the effect. (Murphy, et al., 2009)

5.2 Experimental Design Properties

Two experimental design properties are discussed here and include the experimental design coverage in the unit cube and the normality assumption and minimum variance as a result of employing LHD.

5.2.1 Experimental Design Coverage

This section compares the proposed methodology's LHD to the experimental designs used with traditional DOE. LHDs are space-filling designs that span the unit hypercube with an equal number of design points while traditional experimental designs cover only a subspace of the unit hypercube with a varying number of design points depending upon the type of experimental design chosen. Both Schamburg (2004) and Crino (2006) show that the coverage of the design space for a space-filling design is greater than the coverage offered by traditional experimental design. The notation from Schamburg

(2004) is used here where the fractional coverage coefficient δ is defined as the portion of the edge length that each design point should cover, given the desired number of observed levels *m* and edge length *c*, $\delta = \frac{c}{m}$.

The methodology using LHD proposes model development over domain Ω while the traditional experimental design uses domain Θ, where both of these domains exist in Ψ and the methodology presented here initially considers $\Omega = \Psi$.

Lemma 1 *(Schamburg) (Experimental Design Coverage in the Unit Cube): For the unit cube, let H equal the fractional hyper-volume coverage of the traditional experimental design. Also for the same unit cube, let V equal the fractional hyper-volume coverage of the space-filling experimental design. Let n (=m) represent the number of levels observed for each variable in the Latin-Hypercube Design and o represent the number of levels observed by the traditional experimental design, therefore assuming m ≥ n, then*

$$V \geq H. \tag{5.4}$$

See Schamburg (2004) for the proof.

Lemma 2 *(Schamburg) (Inclusion of the Global Optimal within the Experimental Design): Let $x_i^*, \ldots, x_p^*$ represent the global optimal solution in domain space Ψ . If an optimal exists, the probability of covering the optimal with the LHD is greater than or equal to the probability of covering it in the traditional experimental design. Furthermore, the LHD is certain to cover the optimal when the LHD considers the full*

decision space Ω when compared to the traditional, smaller, experimental design space Θ. Therefore, assuming $c \geq b$,

$$[Pr(x_1^*, \ldots, x_d^* \in \Theta) = \theta] \leq [Pr(x_1^*, \ldots, x_d^* \in \Omega) = \omega = 1] \quad (5.7)$$

See Schamburg (2004) for the proof.

This property is important to the methodology as SA is applied to the response prediction formula in searching for the global optima. Therefore, having good coverage is a necessary condition for the success of the methodology. LHD therefore results in the desirable properties of good coverage of the unit cube so that the chances of finding the global optimal are improved.

5.2.2 Normality and Minimum Variance

LHS provides additional properties as defined in Santner et al. (2003) and supports the assumptions of normality discussed in section 4.2.

Definition 1:

$$\mu = \int_{\chi} y(x) dF(x) \quad (5.12)$$

$$\text{and } \alpha_j(x_j) = \int_{\chi_{-j}} [y(x) - \mu] dF_{-j}(x_{-j}), \quad (5.13)$$

For $1 \leq j \leq d$. Then μ is the overall mean, the $\{\alpha_j(x_j)\}$ are the main effect functions corresponding to the coordinates of x, and

$$r(x)=y(x)-\mu-\sum_{i=1}^{d}\alpha_i(x_i) \qquad (5.14)$$

is the residual (from additivity) of y(x).

The following theorem compares Latin hypercube sampling (LHS) to simple random sampling (SRS) in order to show that a LHS scheme results in a smaller variance than just employing a SRS scheme.

Theorem 1 *(Santner): As n* $\rightarrow\infty$*, under Latin hypercube sampling (LHS) and simple random sampling (SRS)*

$$Var_{LHS}\{\bar{Y}\} = \frac{1}{n}\int_{\chi} r^2(x)dF(x) + o(n^{-1}) \qquad (5.15)$$

and

$$Var_{SRS}\{\bar{Y}\} = \frac{1}{n}\int_{\chi} r^2(x)dF(x) + \frac{1}{n}\sum_{i=1}^{d}\int_{a_i}^{b_i}\alpha_i^2(x_i)dF_i(x_i) + o(n^{-1}), \qquad (5.16)$$

respectively.

This theorem states that unless all $\alpha_j(\bullet)$ *are identically 0 in the limit, then LHS has a smaller variance than SRS. Stein (1987) provides the proof.*

Theorem 2: *(Santner) If y(x) is bounded under LHS Minimizes , then* $\sqrt{n}(\bar{Y}-\mu)$ *tends in distribution to*

$$N(0,\int_{\chi} r^2(x)dx) \text{ as } n \rightarrow \infty \qquad (5.17)$$

The selection of LHS supports the assumption that the samples are independent and the above theorems support the assumptions that the distribution is normal distributed with a small variance. Owen (1992) provides the proof.

Given definition 1 an theorem 1 and 2, it follows that the use of LHS results in a distribution that is normal and a smaller variance over a SRS scheme. Having a smaller variance improves efficiency and helps minimize the number of samples needed. Having a distribution that is normal is a prerequisite for many statistical analysis tools as shown in section 4.2.

5.3 Optimization Properties

In the ECEM, RSM is extended with the use of SA and multiple objective optimization. After the Gaussian process model is fit to the data, SA is applied to the response prediction formula. Using SA to find the factor settings that optimize the response(s) minimizes the number of experimental runs as the prediction model is used rather than running additional experimental runs. SA also helps determine the optimal response which can be used as an upper or lower bound in applying linear or non-linear optimization.

Given the case of multiple objectives, the analyst must determine the decision maker's priorities in optimizing the output variables. Tradeoffs and compromises must be made in coming up with these preferences. The optimization may result in a set of possible solutions. Figure 5.1 shows the best solutions within this set of solutions which make up the Pareto set, an efficient frontier of solutions within the set of possible

solutions.

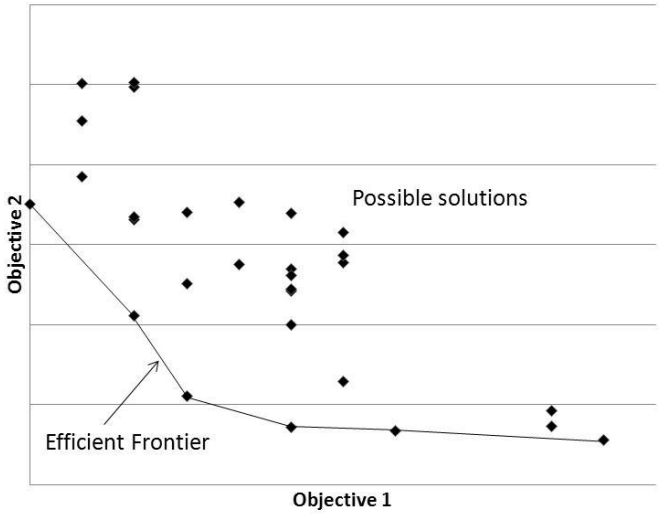


Figure 5.1 Multiple objective optimization requires tradeoffs. Identifying the efficient frontier helps define the best solution given these tradeoffs.

For a problem with more than two objectives, the efficient frontier is considered for each pair of objectives. This can be complex as the number of objectives grows and therefore the analyst should consider one of the four methods identified for multiple objective optimization in section 4.2.6.

Since this methodology makes use of Gaussian process modeling, we know the general form for the prediction formula and know that it is a second order model. As a result we must use a nonlinear optimizer. Using the upper or lower bound determined through the application of SA, the highest priority objective is optimized while the other objectives are soft constraints for the non-linear programming problem.

If preemptive optimization or goal programming is the chosen optimization method, SA can be the optimization tool in which to optimize each objective while carrying out the preemptive or goal programming process. Similar to its use in the case of applying non-linear optimization, SA provides the bound in order to convert the objective

into a constraint. Applying the weighted sum strategy, the objectives should first be made into a single objective function which can then be optimized using SA. Once the optimized solution is determined, it becomes the treatment under investigation and the treatment that additional simulation runs are spent.

5.4 Properties of Sequential Analysis

Sequential analysis began with the efficient testing of anti-aircraft gunnery and other weapon systems during WWII. Due to time constraints, it was highly desirable to conduct fewer inspections but still achieve accurate results. The use of these methods were classified in the mid 1940's, however after the war these methods were expanded to inventory, queuing, reliability, lifecycle tests, quality control, design of experiments and multiple comparison problems. (Mukhopadhyay and de Silva, 2009) These methods came into favor during the conduct of clinical trials in the 1960's and 1970's and incorporated adaptive designs and optimal stopping rules.

Wald developed the traditional methods of sequential analysis after coworkers suggested a general approach which resulted in Wald's book on sequential analysis in 1947. Wald (1947) states that the use of sequential analysis helps minimize sample size. Table 5.1 and Table 5.2 show the average sample size savings given α and β.

Table 5.1 Average % savings in sample size when H_1 is true

α\β	.01	.02	.03	.04	.05
.01	58	60	61	62	63
.02	54	56	57	58	59
.03	51	53	54	55	55
.04	49	50	51	52	53
.05	47	49	50	50	51

Table 5.2 Average % savings in sample size when Ho is true

α/β	.01	.02	.03	.04	.05
.01	58	54	51	49	47
.02	60	56	53	50	49
.03	61	57	54	51	50
.04	62	58	55	52	50
.05	63	59	55	53	51

In the 1950's, modifications to Wald's sequential probability ratio test (SPRT) were made to make use of sequential analysis during the conduct of clinical trials. The alternative to the SPRT is the repeated significance test (RST). Interim analysis methods became more powerful based on stochastic curtailment and statistical tests using the O'Brien-Flemming statistical boundaries. (Lai, 2001) Both the RST and stochastic curtailment test are used in this methodology during the sequential analysis of the data using interim data monitoring.

5.5 Test Functions

The ECEM was applied to four nonlinear test functions in eight dimensions. These

test functions are commonly used in testing global optimization methods. The results were compared to traditional response surface design and global optimization techniques. The four test functions are described in Molga and Smutnicki (2005) and are summarized below.

DeJong's first function is the simplest of the four functions. It is unimodal and convex and has the form shown in equation 5.18.

$$f(x) = \sum_{i=1}^{8} x_i^2, \text{ where } -5.12 \leq x_i \leq 5.12 \qquad (5.18)$$

The global minimum is f(x) = 0 when $x_i = 0$, $i = 1,\ldots,n$.

Rosenbrock's valley, also known as the banana function or DeJong's second function, has a long parabolic shaped valley which makes convergence to the global optimum difficult. The function is of the form shown in equation 5.19.

$$f(x) = \sum_{i=1}^{n-1}[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2], \text{ where } -2.048 \leq x_i \leq 2.048 \qquad (5.19)$$

The global minimum is f(x) = 0 for $x_i = 1$, $i = 1,\ldots,n$.

Rastrigin's function is a modification of DeJong's with a cosine function to replicate multiple local minima. The function is shown in equation 5.20.

$$f(x) = 10n + \sum_{i=1}^{n}[x_i^2 - 10cos(2\pi x_i)] \text{ , where } -5.12 \leq x_i \leq 5.12 \qquad (5.20)$$

The global minimum is f(x) = 0 for $x_i = 0$. $i = 1,\ldots,n$.

Schwefel's function is multimodal and function evaluations usually converge in the wrong direction. The function is of the form shown in equation 5.21.

$$f(x) = \sum_{i=1}^{n}\left[-x_i sin\left(\sqrt{|x_i|}\right)\right], \text{ where } -500 \leq x_i \leq 500 \quad (5.21)$$

The global minimum is f(x) = -418.9829n for $x_i = 420.9687$, $i = 1,\ldots,n$.

These functions were evaluated using SA, GA, the EVOP with an NOLHD and SA, EVOP with NOLHD and GA, EVOP with full factorial design and SA and EVOP with full factorial design and GA. The results of applying each of the above to the four test functions are summarized in Table 5.3. Each of these results is discussed in detail below.

Table 5.3 Test functions summary of results

Test Function		**SA**	**GA**	EVOP NOLHD SA	EVOP NOLHD GA	EVOP Full Factorial SA	EVOP Full Factorial GA
DeJong's Fcn	Optimal fcn eval	0.0205	0.0150	1.85E-10	0.00807	32	32
	iterations	261	83	33	33	512	512
	error	0.0205	0.0150	1.85E-10	0.00807	32	32
Rosenbrock's Valley	Optimal fcn eval	3.4412	2.0222	-55.202	-52.714	∞	∞
	iterations	444	96	66	66	512	512
	error	3.4412	2.0222	55.202	52.714	∞	∞
Rastrigins Fcn	Optimal fcn eval	41.5076	2.8664	-8.81E-13	7.16E-4	156.624	156.624
	iterations	424	104	66	66	512	512
	error	41.5076	2.8664	-8.81E-13	7.16E-4	156.624	156.624
Schwefels Fcn	Optimal fcn eval	-1416.62	-31.5621	-1492.74	-622.166	∞	∞
	iterations	2110	54	66	66	512	512
	error	0.1546	0.9806	0.1092	0.6284	∞	∞
DeJong's Fcn With Noise	Optimal fcn eval	0.8049	-2.7019	-0.2037	-0.2032	∞	∞
	iterations	1013	1013	66	66	512	512
	error	0.8049	2.7019	0.2037	0.2032	∞	∞

First, each test function was optimized with SA and GA and the results show that SA has a comparable or lower error with comparable iterations except in the case of Rastrigin's functions where GA had better results. When NOLHD and SA where applied to the functions, the errors were improved over the SA and GA results with far fewer experimental iterations. One benefit of applying the methodology is that the response surfaces generated helps to narrow the search area in which SA or GA is applied, therefore focusing the search to the global optimum as indicated on the response surface.

Next, each test function was evaluated by applying the iterative process of EVOP with the NOLHD and full factorial design schemes. In applying the EVOP process with

NOLHD, the significant factors can be found and the model is reduced to few factors thus simplifying the model and the optimization process as well. The NOLHD prediction formula resulted in less significant factors than the full factorial design in both the phase I and phase II design process. Table 5.4 shows the significant factors for each design given the test functions. Note that there were no significant factors in the full factorial for Rastrigin's function. This is a result of the same function evaluation for each of the design points and therefore a constant plane for each of the pairwise response surfaces.

Table 5.4 Phase I significant factors for each test function

Test Function	**NOLHD**		**Full Factorial**	
	Significant Factors	#	Significant Factors	#
DeJong's Fcn	x1, x5, x6, x8	4	x1,x2,x3,x4,x5,x6,x7,x8	8
Rosenbrock's Valley	x2,x4,x5,x8	4	x1,x2,x3,x4,x5,x6,x7,x8	8
Rastrigin's Fcn	x3,x4,x8	3	none	0
Schwefels Fcn	x1,x4,x5,x8	4	x1,x2,x3,x4,x5,x6,x7,x8	8
DeJong's Fcn w\noise	x1,x3,x4,x8	4	x1,x2,x3,x4,x5,x6,x7,x8	8

Each of these designs were then optimized with SA and GA. Table 5.5 summarizes the Phase I function evaluation results which show that the NOLHD design with SA and GA found similar optimal function evaluations, however SA resulted in factor settings that were closer to the actual optimal factor settings and was therefore more accurate in finding the global minimum design points. It was also found that Rosenbrock's valley and Schwefel's function did not converge to a solution given the full factorial design and Rastrigin's function using the NOLHD did not converge when GA

was applied.

Table 5.5 Phase I global optimization of common test functions using NOLHD and Full Factorial designs and SA and GA optimization techniques

Test Function	NOLHD SA Fcn eval	NOLHD GA Fcn eval	FF SA Fcn eval	FF GA Fcn eval
DeJong's Fcn	1.85E-10	0.00807	209.72	209.72
Rosenbrock's Valley	1396.40	1396.44	∞	∞
Rastrigins Fcn	1.98	∞	231.40	231.40
Schwefels Fcn	-1033.83	-936.11	∞	∞
DeJong's Fcn w\noise	-0.7778	-0.7940	∞	∞

The four evaluated test functions were compared to the actual optimal function evaluation and the error was calculated by equation 5.22.

$$Rel.Error = \frac{|f(x) - \hat{f}(x)|}{1 + |f(x)|} \times 100\% \qquad (5.22)$$

The results summarized in Table 5.6 show that the NOLHD design resulted in a smaller error. The SA and GA errors are close but the SA optimal solution was more accurate in that it found the optimal factor settings in the DeJong's and Rastrigin's function.

Table 5.6 SA and GA error comparisons for phase I global optimization of common test functions using NOLHD and Full Factorial designs.

Test Function	NOLHD SA error	NOLHD GA error	FF SA error	FF GA error
DeJong's Fcn	1.85E-10	0.00807	209.72	209.72
Rosenbrock's Valley	1396.40	1396.44	∞	∞
Rastrigins Fcn	1.98	1.98	231.40	231.40
Schwefels Fcn	0.3829	0.4412	∞	∞
DeJong's Fcn w\noise	0.7778	0.7940	∞	∞

The response surfaces generated in Phase I were examined and new factor ranges were selected for the new factor settings. New NOLHD and full factorial designs were generated and Table 5.7 summarizes the factors that were found to be significant in Phase II.

Table 5.7 Phase II Significant factors

Test Function	NOLHD		Full Factorial	
	Significant Factors	#	Significant Factors	#
DeJong's Fcn	x1,x2,x5,x6,x7	5	x1,x2,x3,x4,x5,x6,x7,x8	8
Rosenbrock's Valley	x1,x2,x3,x6,x8	5	x1,x2,x3,x4,x5,x6,x7,x8	8
Rastrigins Fcn	x2,x5,x6,x7	4	none	0
Schwefels Fcn	x1,x4,x5,x8	5	x1,x2,x3,x4,x5,x6,x7,x8	8
DeJong's Fcn w\noise	x2,x4,x5,x7,x8	5	x1,x2,x3,x4,x5,x6,x7,x8	8

Table 5.8 summarizes the Phase II function evaluation results which show that the NOLHD design with SA achieved a lower error and was more accurate in finding the

global minimum design points over the full factorial design.

Table 5.8 Phase II global optimization of common test functions using NOLHD and Full Factorial designs and SA and GA optimization techniques

Test Function	**NOLHD SA** Fcn eval	**NOLHD GA** Fcn eval	**FF SA** Fcn eval	**FF GA** Fcn eval
DeJong's Fcn	0.00186	0.0128	32 For all xi	32 For all xi
Rosenbrock's Valley	-55.202	-52.714	∞	∞
Rastrigins Fcn	-8.81E-13	7.16E-4	156.624 For all xi	156.624 For all xi
Schwefels Fcn	-1492.74	-622.166	∞	∞
DeJong's Fcn w\noise	-0.2037	-0.2032	∞	∞

Table 5.9 shows a decrease for the error in achieving the global optimum function evaluations as the factor settings are narrowed in phase II.

Table 5.9 SA and GA error comparisons for phase II global optimization of common test functions using NOLHD and Full Factorial designs.

Test Function	**NOLHD SA** error	**NOLHD GA** error	**FF SA** error	**FF GA** error
DeJong's Fcn	0.00186	0.0128	32	32
Rosenbrock's Valley	55.202	52.714	∞	∞
Rastrigins Fcn	-8.81E-13	7.16E-4	156.62	156.62
Schwefels Fcn	0.1092	0.6284	∞	∞
DeJong's Fcn w\noise	0.2037	0.2032	∞	∞

5.6 Interim Analysis and Test Function with Noise

To demonstrate the interim analysis portion of the methodology, the DeJong test function modified by adding a noise term with a low signal to noise ratio is examined and is shown in equation 5.22.

$$f(x) = \sum_{i=1}^{8} x_i^2 + Gaussian\ (0,3), \text{ where } -5.12\ \leq x_i \leq 5.12 \qquad (5.22)$$

The global minimum is E[f(x)] = 0 when $x_i = 0$, $i = 1,\ldots,n$.

The ECEM was applied to the function. The NOLHD was simulated with five replications and a response prediction formula was found. SA was applied to the response prediction formula to find the optimal factor settings that optimize the formula. These factor settings became the phase I results. The factor settings were then narrowed as a result of examining the response surface and a new NOLHD was found and simulated with five replications. A new response prediction formula was found and SA was applied. The optimal factor settings became the phase II results. The phase I and phase II results were then used in the stochastic DeJong test function and simulated while interim analysis was applied. Table 5.10 summarizes the results for the data collected after ten and 20 replications.

Table 5.10 The data collected on the noisy DeJong's test function for phase I and phase II given ten and 20 replications.

n	Phase I Mean	Phase I standard dev	Phase II Mean	Phase II standard dev
10	-1.04	3.97	-0.37	2.79
20	-0.35	3.33	.77	3.5

After 20 replications the stopping criteria was met and the trial was stopped for futility. While there was an improvement in the optimal value from phase I to phase II, the statistics indicate that there is not a statistical difference between the phase I and phase II results. Table 5.11 shows the test statistic at each interim analysis and the upper and lower stopping boundaries.

Table 5.11 Interim analysis results after ten and 20 replications.

Analysis	Test statistic	Lower stopping boundary (test statistic scale)	Upper stopping boundary (test statistic scale)	Decision
1	-.4366	-1.046	3.6321	Continue
2	-1.037	-.4576	3.6084	Stop trial for futility

Figure 5.2 graphically depicts the interim analysis results and that the trial should be stopped for futility as the phase I and phase II results are not statistically different.

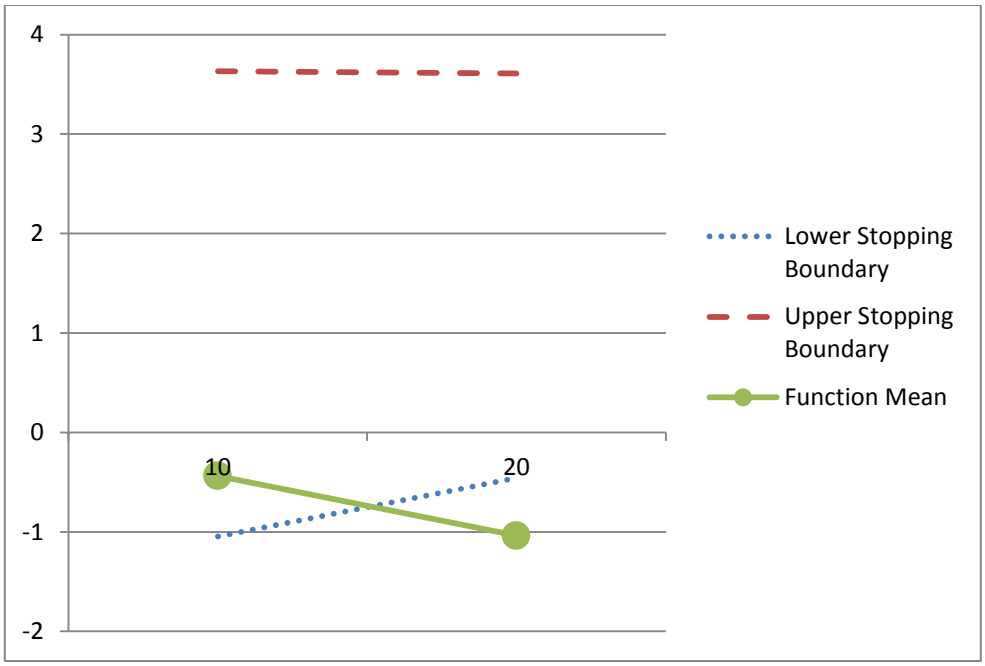


Figure 5.2 Efficacy and Futility Monitoring Boundaries for the interim analysis of the noisy DeJong test function.

5.7 Summary

The theoretical properties discussed in this chapter provide sufficient evidence for the use of the ECEM. Each of the areas discussed either provide necessary conditions for statistical analysis or help achieve efficiency in time, cost or number of replications. It was shown here that the ECEM uses techniques that help minimize sample size with power analysis and the use of a space filling design, the NOLHD. The experimental design also achieves efficiency as it also minimizes variance.

The ECEM was demonstrated on four test functions plus the DeJong's test function with noise. The results are compared and show the ECEM to be more efficient over SA and GA. Interim analysis is demonstrated on The DeJong's test function to demonstrate how the ECEM applies to a stochastic process. This methodology is an improvement over traditional methods and is demonstrated in Chapter 6 with a chemical mixing experiment and a police staffing simulation.

6 Application of the Methodology

Section 6.1 describes the chemical mixing problem simulation used to initially test the ECEM. The methodology is applied and the results are discussed. Section 6.2 describes the police staffing study simulation. The methodology is applied to this simulation and results are presented.

6.1 The Chemical Mixing Problem

The methodology developed in this research was applied to the simulation of a chemical mixing process. (Hill, 1998) The output of a chemical process in pounds per hour is thought to be influenced by a subset of five factors: temperature, reactant concentration, catalyst feed rate, pressure, and reaction time. The range of operation for each of these factors is shown in Table 6.1:

Table 6.1 Phase I Feasible Factor Settings

Process variable	Minimum	Maximum
Temperature (°C)	200	450
Concentration (g\l)	100	500
Catalyst feed rate (g\m)	200	600
Pressure (psi)	140	200
Reaction time (hrs)	1	6

These limits represent the absolute limits of operation for this process. Outside these limits the process is likely to produce unacceptable results or unsafe operating conditions for plant personnel. However, large segments of this region are unexplored, and since this is a full-scale process, care must be taken in adjusting any of these variables.

The current operating conditions on these variables are summarized in Table 6.2.

Table 6.2 Current Operating Condition Factor Settings.

Process variable	Setting
Temperature (°C)	262.5
Concentration (g\l)	200
Catalyst feed rate (g\m)	300
Pressure (psi)	155
Reaction time (hrs)	2.25
Simulated Avg Yield (lbs\hr)	33.43
Simulated Avg Viscosity	0.78
Simulated Avg. Molecular Weight	350.78

This is a multiple objective problem with the following objectives for the three responses:

Table 6.3 Response Objectives

Yield	Viscosity	Molecular Weight
Maximize	< 450	350 ≤ Molecular Weight ≤ 550

Yield is defined as the output in pounds of the chemical mixture per hour, viscosity is the thickness of the resulting chemical mixture and molecular weight is the sum of the atoms of the molecules in the chemical mixture. The simulation is run at faster than real time so the issue with this experiment and being too time consuming or using too much computing resources is not an issue. Three replications at each design point are completed in order to get the average output. The expected value for each output would be the long-run average value over many independent replications of the experiment. Attempting to predict results in as few replications as possible it is not likely that the average value is the expected value, but as the design region is morphed and the factor ranges are reduced it is likely that the average could approach the expected value as the number of times the process is replicated. If the simulation is capable of reaching steady-

state conditions, then this would be the expected value for the output parameters. The chemical mixing problem being investigated has a fixed time interval.

6.1.1 Application of the Methodology to the Chemical Mixing Problem

The steps for the methodology are presented with the resulting work shown. In step 1, the monitoring plan or protocol is developed. This plan consists of the following six steps:

1) Choose the group sequential plan that reflects the objectives of the experiment and specify the hypotheses to be tested: H_o: $\mu_1 = u_2$ (no treatment difference) and H_1: $\mu_1 \neq u_2$ (meaningful treatment difference). Examine the current treatment and the subsequent treatment with hypothesis testing. Once the hypothesis H_o is met, the experiment is terminated. Expanding the hypotheses to account for the multiple responses use the terminology $\boldsymbol{H_o}$: $\boldsymbol{\mu_1} = \boldsymbol{u_2}$ and $\boldsymbol{H_1}$: $\boldsymbol{\mu_1} \neq \boldsymbol{u_2}$ where $\boldsymbol{H_o}$, $\boldsymbol{H_1}$, $\boldsymbol{\mu_1}$ and $\boldsymbol{u_2}$ is the matrix notation for the set of null and alternative hypotheses and mean values for all the responses under consideration and the subscript 1 and 2 refer to the current and improved data sets respectively under comparison.
2) A two-sided test with a significance level of $\alpha = 0.05$
3) The maximum number of planned analyses and the maximum sample size are determined based on available resources such as time and\or money.
4) The best sample size is computed based on power analysis and is dependent on the estimate of the treatment effect. Figure 6.1 shows the results of the power analysis.

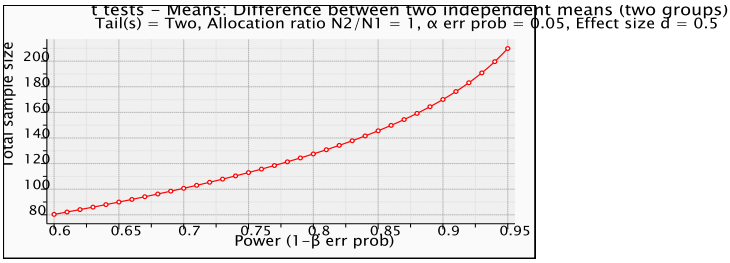


Figure 6.1 Total Sample Size vs. Power

5) Estimate the treatment effect. The results of these calculations are shown in Table 6.4. To start this process a "medium" treatment effect is assumed. Figure 6.2 depicts the distribution of the errors.

Table 6.4 A Priori Required Sample Size

Analysis: A priori	**Compute required sample size**
Input: Tail(s)	2
Effect size d	0.5 (assumed medium)
A error probability	0.05
Power (1- β error probability)	0.90
Allocation ratio N2/N1	1
Output: Noncentrality parameter δ	3.278719
Critical t	1.974017
Df	170
Sample size treatment 1	86
Sample size treatment 2	86
Total sample size	172

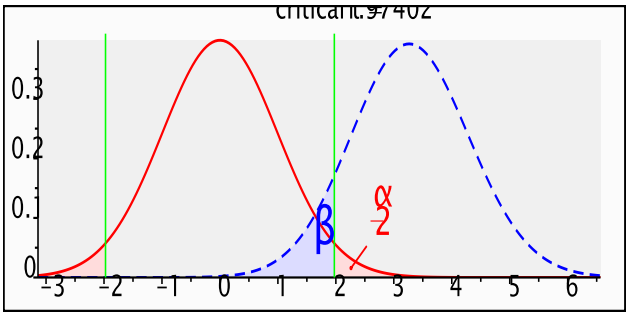


Figure 6.2 Error Distribution

As a result of the power analysis and estimating the treatment effects, conducting three interim analyses with 29 samples for each treatment would achieve 174 total samples.

6) The type of stopping boundary and the stopping criteria. A group sequential design that combines efficacy and futility testing is selected. The O'Brien-Flemming stopping boundary is used for efficacy, upper stopping boundary, and the Pocock boundary is used for futility, lower stopping boundary. A design based error spending function is applied. Stopping boundaries are selected to protect Type I and Type II error. Simultaneous efficacy and futility is applied as follows:

    1) Stop the trial for efficacy if $Z_k > u_k(\alpha, ß)$
    2) Continue the trial if $l_k(\alpha, ß) \leq Z_k \leq \mu_k(\alpha, ß)$
    3) Stop the trial for futility if $Z_k < l_k(\alpha, ß)$

For the first treatment, the current operating conditions will be applied as shown in Table 6.2.

Step 2, select the DOE (range of factor settings to explore). The OLHD will be used due to its efficiency and space-filling properties. The OLHD uses the full range of allowable factor settings in order to initially examine the entire region of feasible operating conditions since there is no prior knowledge on how the process performs over varying operating conditions. Given the five factors, the OLHD EVOP has only 17 design points.

Step 3, the experiment is set up and run given the EVOP experimental design chosen (constitutes a complete cycle).

Step 4, three replications are conducted and the data are collected. (constitutes the phase). Replications are needed as a result of the stochastic nature of the simulation and ensure the fitted model is accurate and will have higher predictive power.

Next in step 5, a Gaussian process model is fit to the results given the three responses and the response prediction formula for each of these is estimated. Appendix A details the response prediction formula for each response as a result of fitting the Gaussian process model. A Gaussian process model is applied to the collected data and results in a prediction formula for each of the three responses: yield, viscosity and molecular weight. The Gaussian process model report can be found in Appendix B. This report highlights the significant factors and interactions. It was found that temperature, concentration and the temperature * concentration interactions were the most significant factors affecting the yield of the product. Temperature, feed rate, time and their interactions were the most significant factors affecting the viscosity of the product. Temperature, concentration, feed rate, pressure and all of their interactions were the most significant factors affecting the molecular weight. Table 6.5 summarizes the significant

factors for each response.

Table 6.5 Phase II Significant Factors by Response

Response/Factor	Yield	Viscosity	Molecular Weight
Temperature	X	X	X
Concentration	X		X
Feed Rate		X	X
Pressure			X
Time		X	
Temp*Conc	X		X
Temp*Feed Rate		X	X
Temp*Time		X	
Temp*Press			X
Conc*Feed Rate			X
Conc*Press			X
Feed Rate*Time		X	
Feed Rate*Press			X

All of the RSMs with respect to the responses and their significant factor interactions were examined. The following Treatment #1 RSMs for each response with respect to temperature and concentration are demonstrated in Figure 6.3. These response surfaces can be used to find the temperature and concentration settings that optimize each of the three responses based on the objectives in Table 3.

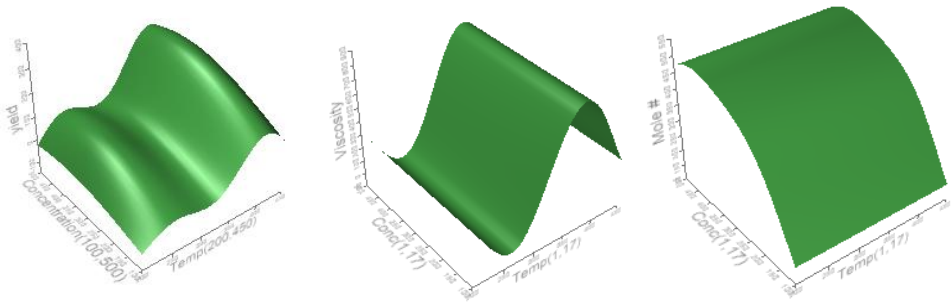


Figure 6.3 Phase II OLHD RSM

Figure 6.4, Figure 6.6 and Figure 6.8 demonstrates the adequacy of the model fit to the data. The model appears to be a good fit for the data for each response and no transformations of the data are needed.

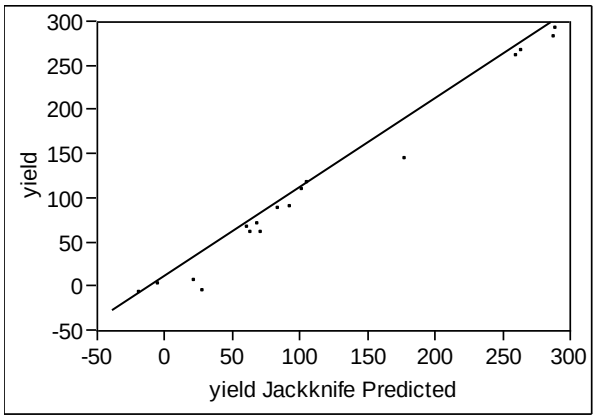


Figure 6.4 Phase II Yield Actual by Predicted Plot

The prediction profiler can be used to help optimize the yield and shows where the response is optimized given the varying factor settings. Figure 6.5, Figure 6.7 and Figure 6.9 shows the prediction profiler for each response.

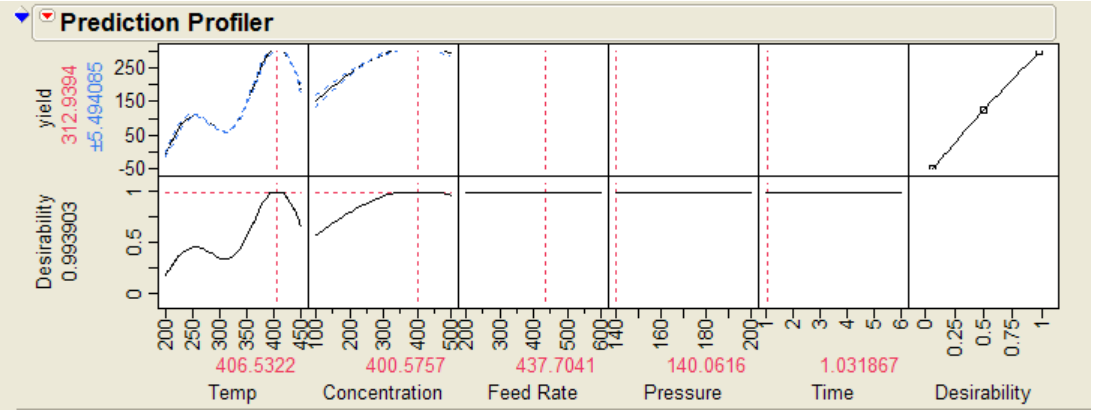


Figure 6.5 Phase II Yield Prediction Profiler

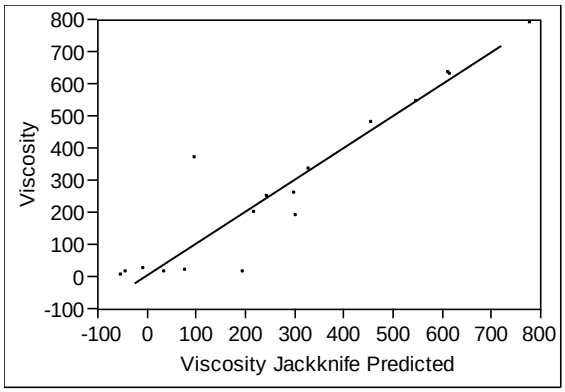


Figure 6.6 Phase II Viscosity Actual by Predicted Plot

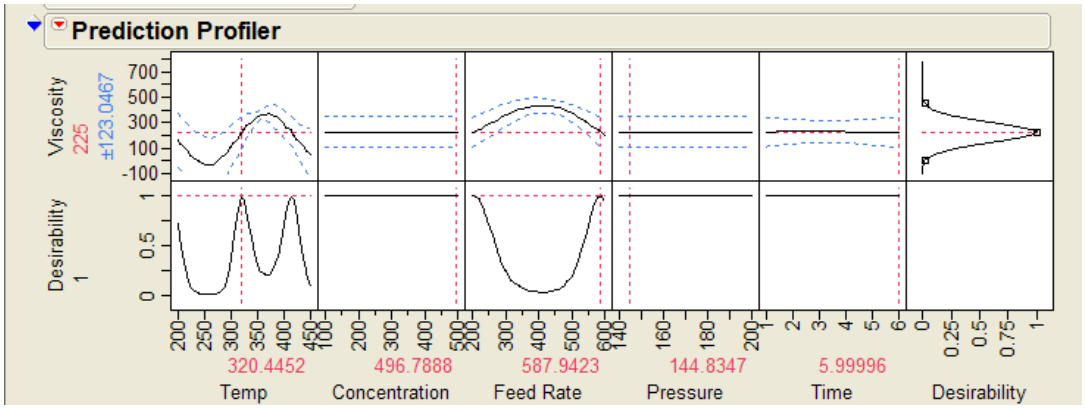


Figure 6.7 Phase II Viscosity Prediction Profiler

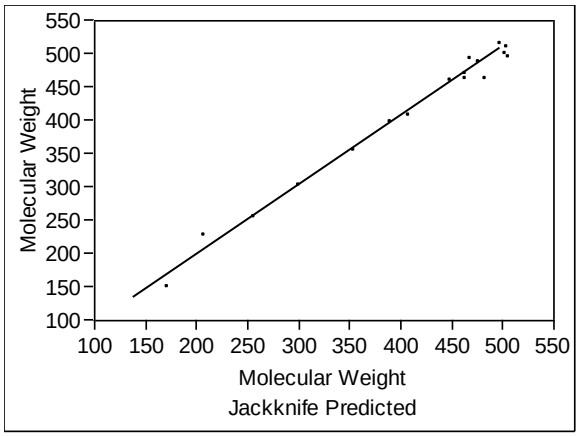


Figure 6.8 Phase II Molecular Weight Actual by Predicted Plot

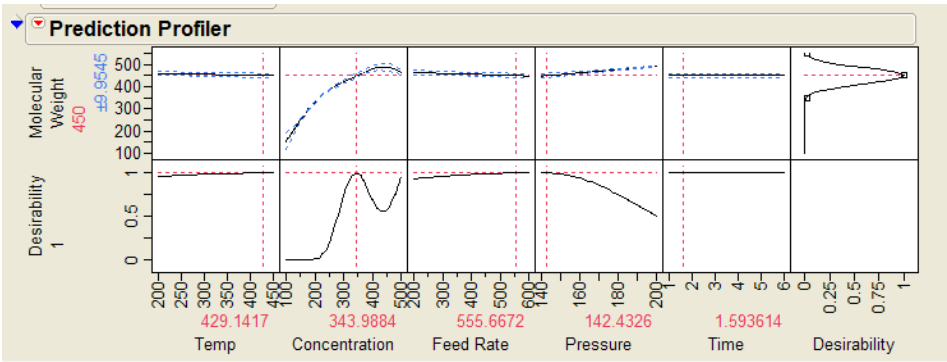


Figure 6.9 Phase II Molecular Weight Prediction Profiler

Step 6 optimizes the response prediction formula using SA. The response prediction formulas are optimized using SA to maximize the responses and obtain the optimal solution and factor settings for this model. Applying SA to the yield prediction formula indicates that a temperature of 407 ˚C and a concentration of 401 g\l will result in a maximum yield of 312.92 lbs\hr. In checking these factor settings in Figure 6.3 it can be seen that these are solutions that also meet the viscosity and molecular weight objectives as well. The three response prediction formulas, with the constraints on the responses, are solved simultaneously using non-linear programming to find a solution

that meets the three response objectives. Observing the three response surfaces aids in determining this result. The results of multi-objective optimization of the three response prediction formulas are shown in Table 6.6.

Table 6.6 Phase II Multi-Objective Optimization of Response Prediction Formulas

Response	**Predicted Value**
Average Yield (lbs\hr)	312.92
Average Viscosity	325.09
Average Molecular Weight	479.92

Step 7) Now that an optimal solution for treatment #2 is found, this step allows a new treatment to be chosen to see if the results can be improved upon. Search the design region by examining the response surface of this design and select a new range for the factor settings based on these results. A search of the design region was accomplished using RSM and a visual inspection of the response surfaces. Keeping in mind the three objectives (Table 6.3) for the three responses, these RSMs can help locate the feasible factor settings that meet each of the three objectives. The response surfaces help visualization of the feasible areas that meet the objectives of each of the responses simultaneously.

Step 8) The current operating conditions are designated as treatment #1. Steps 3 – 6 above define treatment #2. Go to step 9.

Step 9) Treatment #1 and treatment #2 are replicated 29 times. The results are summarized in Table 6.7.

Table 6.7 Comparison of Treatment Results

	Temperature ˚C	Concentration g\l	Average Yield lbs\hr	Average Viscosity	Average Molecular weight.
Treatment #1 Results	262.5	200	33.4202	.7872	350.7872
Treatment #2 Results	407	401	306.2495	276.2684	499.7740

In step 10, interim analysis is applied to the data with the following results. After the first interim analysis, the trial continued and each treatment was replicated 29 times more. After the second interim analysis the trial was stopped for efficacy. Treatment #2 was clearly superior to Treatment #1 and resulted in a much higher yield while also meeting the viscosity and molecular weight constraints. After the second interim analysis it was concluded that the experiment can be terminated due to efficacy of Treatment #2 over Treatment #1. The results of the interim analyses are summarized in Table 6.8.

Table 6.8 Interim Analysis Results

Analysis	Test Statistic	Lower Stopping Boundary	Upper Stopping Boundary	Decision
1 after 29 trials	2.2454	1.0992	2.3861	Continue
2 after 58 trials	2.25	1.9600	2.0726	Stop trial for efficacy

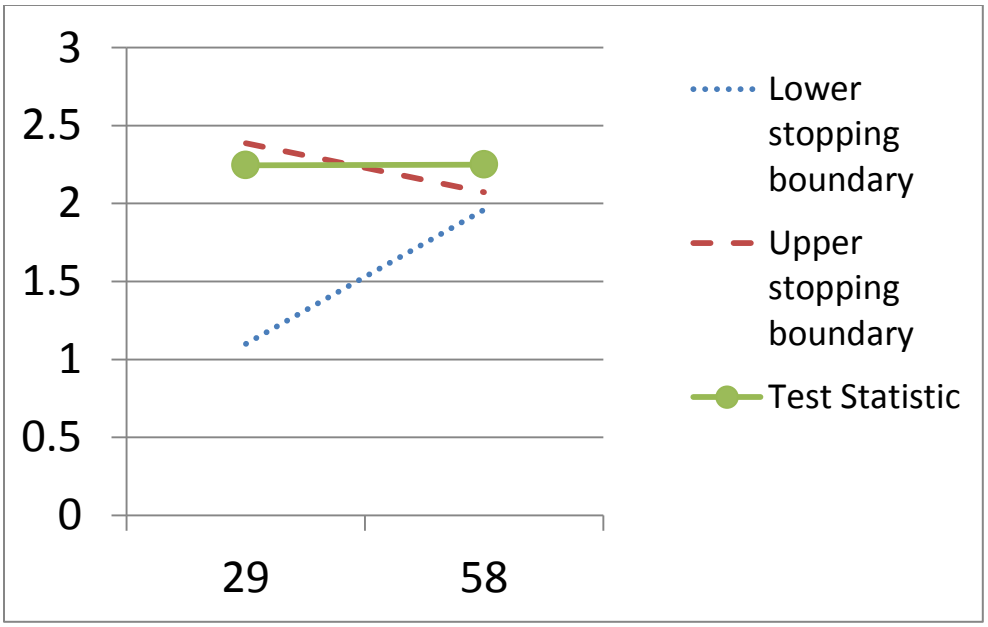


Figure 6.10 Phase I and Phase II Efficacy and Futility Monitoring Boundaries

Next, it is considered if the experiment should be continued or not. Given the results of step 7 of examining the response surfaces, it is concluded that a new range of factor settings be explored to see if greater yields can be achieved. Table 6.9 summarizes the new factor ranges in which to explore.

Table 6.9 Factor Setting Ranges for Phase III

Factor Name	Temperature °C	Concentration g\l	Feed Rate g\m	Pressure psi	Time hrs
Low Level	375	200	200	140	1
High Level	450	500	600	200	6

After setting up the design, the simulation was run with replications and, once again, a Gaussian process model was fit to the data (steps 2 – 5). Table 6.10 shows the significant factors by responses. Note there are addition and omission of factors as significant factors with respect to each response as from what was seen with Phase 1. An addition is annotated with an X*, an omission is annotated with brackets, ( ). The addition

of pressure as a significant factor with respect to yield was determined with Phase III. Although the effect is statistically significant it is small in comparison to the other factors that were determined significant. The response of molecular weight had the most significant change with respect to the significant factors and interactions.

Table 6.10 Phase III Significant Factors by Response

Response/Factor	Yield	Viscosity	Molecular Weight
Temperature	X	X	( )
Concentration	X		X
Feed Rate		X	X
Pressure	X*		( )
Time		X	X*
Temp*Conc	X		( )
Temp*Feed Rate		X	( )
Temp*Time		X	
Temp*Press	X*		( )
Conc*Feed Rate			X
Conc*Press	X*		( )
Conc*Time			X
Feed Rate*Time		X	X
Feed Rate*Press			( )

Figure 6.11 shows the results of the RSM for each of the three responses. The figure shows a dramatic increase in the level of detail in each of the response surfaces over the selected factor ranges and gives a more accurate representation of the RSM. This will help further define the area of study with respect to improving the range over which to examine the factor settings.

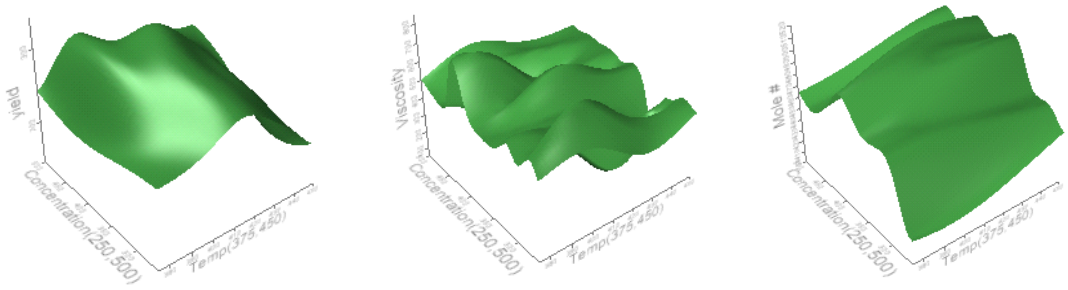


Figure 6.11 Phase III OLHD RSM

Again, the actual by predicted plots are examined to ensure the adequacy of the model fit. Figure 6.12, Figure 6.14 and Figure 6.16 shows the actual by predicted plot for each response respectively. Figure 6.13, Figure 6.15 and Figure 6.17 show the prediction profiler for each response.

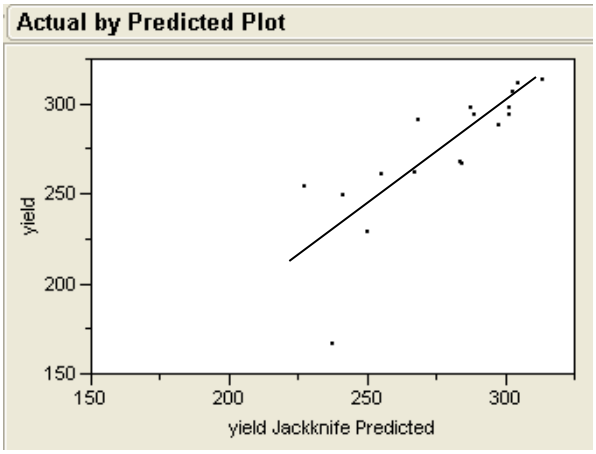


Figure 6.12 Phase III Yield Actual by Predicted Plot

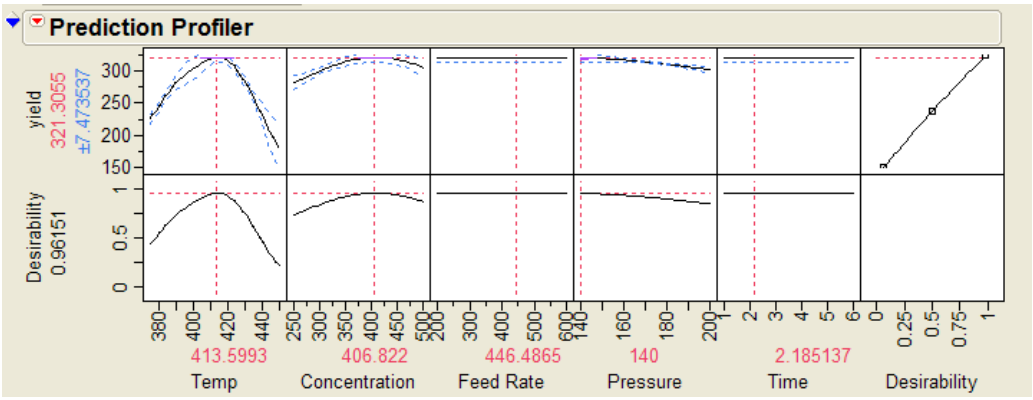


Figure 6.13 Phase III Yield Prediction Profiler

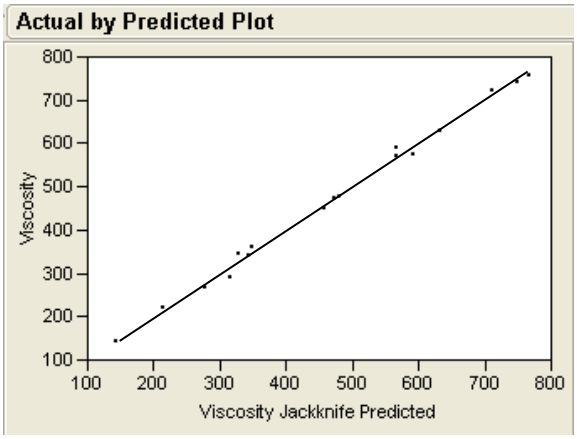


Figure 6.14 Phase III Viscosity Actual by Predicted Plot

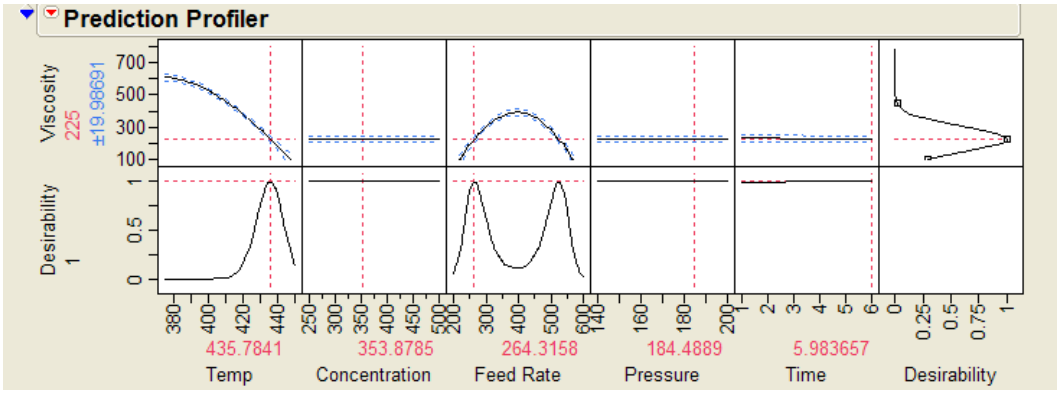


Figure 6.15 Phase III Viscosity Prediction Profiler

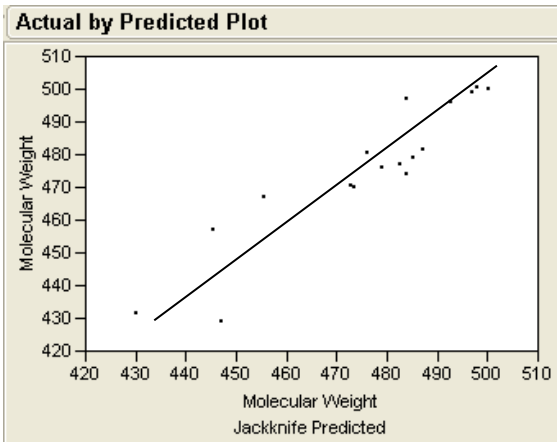


Figure 6.16 Phase III Molecular Weight Actual by Predicted Plot

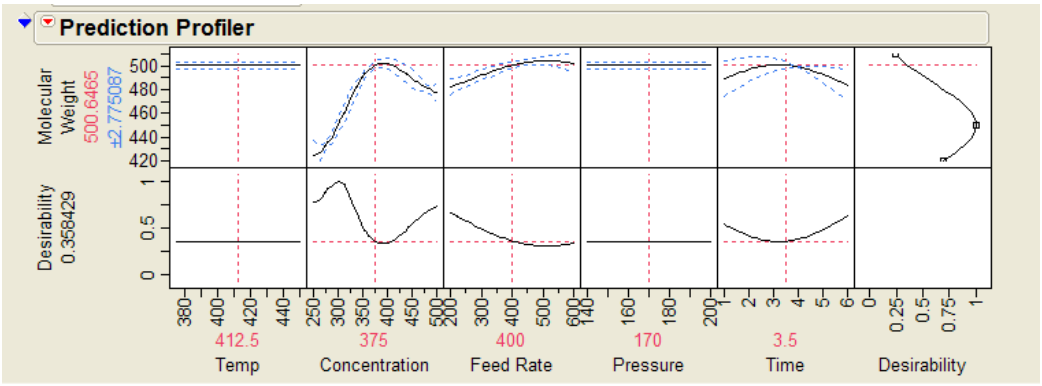


Figure 6.17 Phase III Molecular Weight Prediction Profiler

Step 6 optimizes the response prediction formulas using SA. Applying SA to the yield prediction formula indicates that a temperature of 412 ˚C and a concentration of 413 g\l will result in a maximum yield of 317.30 lbs\hr. In checking these factor settings in Figure 6.11 it can be seen that these are solutions that also meet the viscosity and molecular weight objectives as well. The three response prediction formulas, with the constraints on the responses, are solved simultaneously using non-linear programming to find a solution that meets the three response objectives. The results of multi-objective optimization of the three response prediction formulas are shown in Table 6.11.

Table 6.11 Phase III Multi-Objective Optimization of Response Prediction Formulas

Response	**Predicted Value**
Average Yield (lbs\hr)	317.30
Average Viscosity	264.64
Average Molecular Weight	484.35

Step 9, replications are conducted to collect the data to do the interim analysis on the current factor settings, Treatment #1, and the new factor settings, Treatment #2. The results of the two treatments after 29 replications are shown in Table 6.12.

Table 6.12 Comparison of Treatment Results

	Temperature ˚C	Concentration g\l	Average Yield lbs\hr	Average Viscosity	Average Molecular weight.
Treatment #1 Results	407	401	306.2495	276.2684	499.7740
Treatment #2 Results	413	412	305.5582	227.69	500.24

In step 10, interim analysis is applied to the data with the following results. The first interim analysis shows that each treatment should be replicated 29 more times for a total of 87 replications each. After the second interim analysis it was concluded that the experiment can be terminated due to a lack of Treatment #2 benefit over Treatment #1. The results of the interim analyses are summarized in Table 6.13.

Table 6.13 Interim Analysis Results

Analysis	Test Statistic	Lower Stopping Boundary	Upper Stopping Boundary	Decision
1 after 58 trials	2.2454	1.0992	2.3861	Continue
2 after 87 trials	1.4554	1.9600	2.0726	Stop trial for futility

As described in section 6.1.1, stop the trial for efficacy if $Z_k > u_k(\alpha, ß)$, continue the trial if $l_k(\alpha, ß) \leq Z_k \leq \mu_k(\alpha, ß)$ and stop the trial for futility if $Z_k < l_k(\alpha, ß)$. Figure 6.18 graphically depicts the interim analysis results which show after the second interim analysis the trial should be stopped for futility as $Z_k < l_k(\alpha, ß)$.

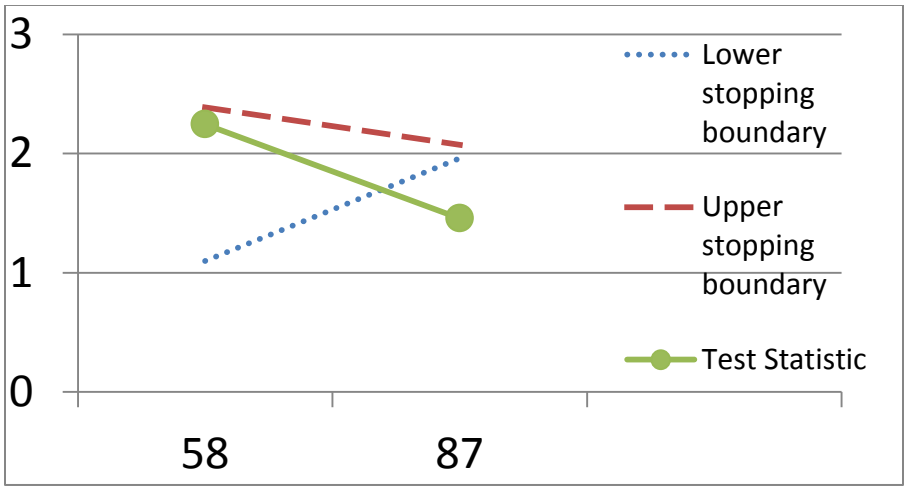


Figure 6.18 Efficacy and Futility Monitoring Boundaries

A box plot is used to graphically summarize and compare groups of data. Figure 6.19 shows a comparison of Treatment #1 and Treatment #2 after 87 trials.

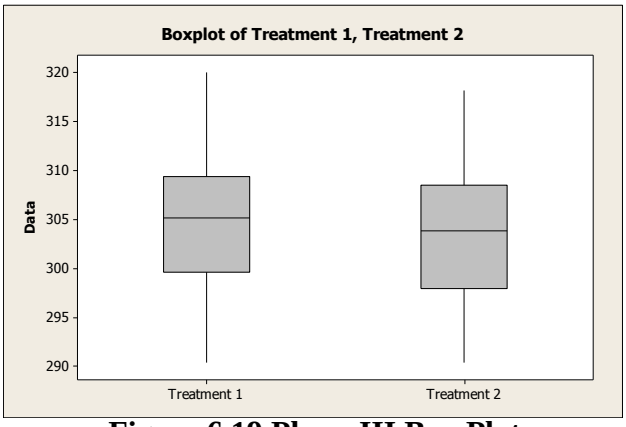


Figure 6.19 Phase III Box Plots

The ECEM results were compared to the full factorial EVOP scheme where after two full EVOP phases and two interim analyses in the second phase, the stopping criteria are not met and the maximum yield for the "best" treatment at that point is 240 lbs\hr. Measures of performance also include a comparison of the number of total design points, achievement or percent improvement of the objective function, and measurement error between the simulation results and the SA results.

6.1.2 Results of the Chemical Mixing Problem

The 2^5 full factorial has 32 design points compared to 17 design points in the OLHD. This is a 47% savings in the number of design points and replications if each design used the same number of replications. This is not the case however as the 2^5 full factorial requires additional replications and still does not achieve the same results as in the OLHD EVOP scheme. The 2^5 full factorial phase III EVOP scheme is run with an additional 154 replications, thus the savings in replications with the OLHD EVOP is really 64% to achieve similar results. After two EVOP phases and two interim analyses, the OLHD EVOP scheme improved the yield response from the initial operating conditions by 48% with a final average yield of 316.67 lbs\hr whereas the 2^5 full factorial resulted in an average yield of only 214.32 lbs\hr. The 2^5 full factorial improved the yield response by 34%. The OLHD resulted in a 32% improvement in the yield over the 2^5 full factorial. The approved solution presented for this problem is given by a 2_V^{5-1} with 4 center points. This design uses 20 design points and results in an average yield of 316.43 lbs\hr. The OLHD is a 15% savings in design points over this fractional factorial.

While both the OLHD and 2^5 full factorial EVOP schemes are clearly an improvement over the current operating conditions, the OLHD EVOP scheme resulted in

greater efficiency through achievement of a higher mean yield in fewer replications than the 2^5 full factorial EVOP scheme. The proposed methodology shows great potential and efficiency for application to stochastic computer experiments with a large number of factors and multiple responses.

6.2 The Police Staffing Simulation Study

The efficient computer simulation methodology is applied to a simulation that examines the staffing of police patrols in the City of Charlottesville, Virginia. This simulation is based on a prototype of RepastCity (Java Repast 1.2). (Malleson) RepastCity is an agent based simulation within a virtual city with a road network. ArcGIS was used to display the City of Charlottesville and its road network. The City of Charlottesville Police Department provided the crime data for the city from 2001 through 2006 which is used to simulate the calls for service within the simulation. The City currently consists of eight districts in which one police patrol is assigned for each shift. This simulation study will analyze various patrol policies with the current patrol policy to see if the performance of the police patrols can be improved. The performance output examined is the average response time, the percent of the total time the patrols are occupied and the cost of the patrols in terms of the distances driven and the number of cars. Minimizing the number of cars also leads to a minimal cost solution and will be taken into account during the evaluation of the competing alternatives. When the average response time and the percent of the total time the patrols are occupied are not statistically significant between competing policies, the policy which minimizes the number of cars is preferred as there are less cars to purchase and maintain and fewer patrols to pay wages. There are eight regions within the city that the patrols may be

deployed to. These regions are depicted by the eight districts shown in Figure 6.20.

Figure 6.20 Charlottesville Police Patrol Regions

The City currently employs one patrol for each district per shift for a total of eight patrols per shift. Given the current calls for service data, Figure 6.21 depicts the probability of a crime occurring within each district. District 3 has the highest crime probability followed by district 5 and 2. Districts 4, 7 and 8 have the lowest probabilities of crime.

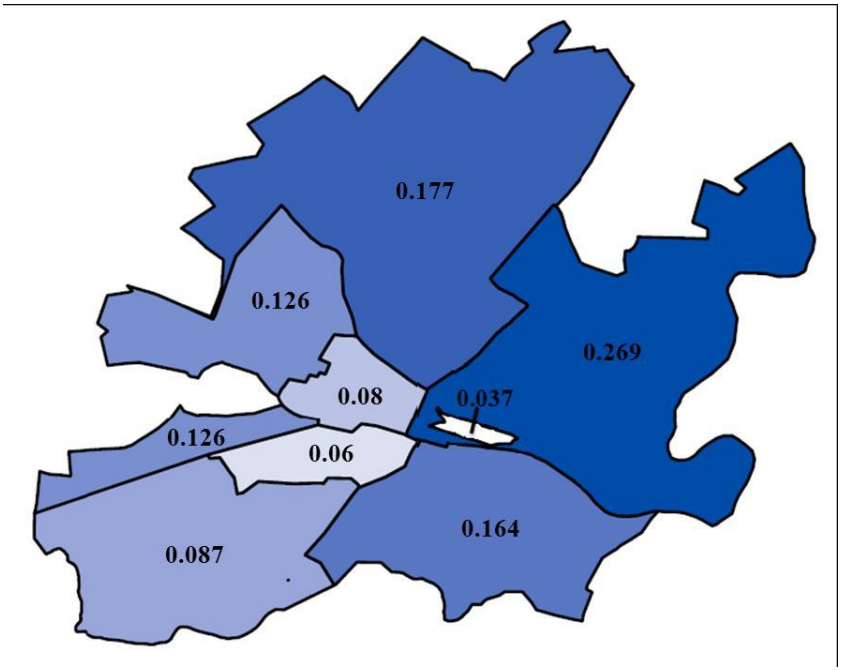


Figure 6.21 City of Charlottesville Crime Probability by District

This study will examine the effects of varying the number of patrols from zero to four in each district and examine the effects on the three outputs. The region of operability for these factors is shown in Table 6.14:

Table 6.14 Phase I Factor Settings

Factor	Minimum	Maximum
Number of Patrols in District 1	0	4
Number of Patrols in District 2	0	4
Number of Patrols in District 3	0	4
Number of Patrols in District 4	0	4
Number of Patrols in District 5	0	4
Number of Patrols in District 6	0	4
Number of Patrols in District 7	0	4
Number of Patrols in District 8	0	4

These limits represent the number of patrols that can be assigned to any one district. Given the size of these districts, it is assumed that placing more than four patrols in any one area would be exorbitant and would greatly exceed the need given the current levels of crime in these areas. Having as many as four patrols however gives a broad enough range in which to examine these factors and therefore a broader response surface in which to examine the factors and there interactions.

This is a multiple objective problem with the following goals for the three responses (Table 6.15):

Table 6.15 Response Objectives

Average Response Time	Average % Time Occupied	Average Total Cost
Minimize	Maximize	Minimize

Average response time is the average response time of all the calls the patrols respond to within the simulation and is measured by the time it takes the patrol to arrive at the scene once a call is received. The average percent total time occupied is the total time the patrol is busy conducting its duties during the course of the simulation run. The average total cost is measured by adding the entire distance all patrols cover over the entire city over the course of the simulation run. The end result of applying the factor settings will be to recommend how many patrols to position in each of the eight districts that best meets the response objectives. Initially, three replications at each design point are completed in order to get the average performance output. The expected value for each output would be the long-run average value over many independent replications of the experiment. Each simulation has a warm up period that allows the patrols to travel to their assigned district prior to receiving any calls. The experiment under investigation

here is configured to run for the same fixed time interval for each policy under investigation.

6.2.1 Application of the Methodology to the Police Staffing Study

The steps for the methodology are presented with the resulting work shown. In step 1, the monitoring plan or protocol is developed. This plan consists of the following six steps:

1) Choose the group sequential plan that reflects the objectives of the experiment and specify the hypotheses to be tested: H_o: $\mu_1 = u_2$ (no treatment difference) and H_1: $\mu_1 \neq u_2$. Examine the current treatment and the subsequent treatment with hypothesis testing. In this problem, the current treatment of one patrol in each district is used and is compared to the results the methodology produces in Phase I. Once the stopping boundaries are met for either efficacy or futility, the experiment is terminated. Expanding the hypotheses to account for multiple responses, the following terminology is used: $\boldsymbol{H_o}$: $\boldsymbol{\mu_1} = \boldsymbol{u_2}$ and $\boldsymbol{H_1}$: $\boldsymbol{\mu_1} \neq \boldsymbol{u_2}$ where $\boldsymbol{H_o}$, $\boldsymbol{H_1}$, $\boldsymbol{\mu_1}$ and $\boldsymbol{u_2}$ is the matrix notation for the set of null and alternative hypotheses and mean values for all the responses under consideration and the subscripts 1 and 2 refer to the two treatments under comparison.
2) A two-sided test with a significance level of $\alpha = 0.05$
3) The maximum number of planned analyses and the maximum sample size are determined based on available resources such as time and/or money.

4) The best sample size is computed based on power analysis and is dependent on the estimate of the treatment effect. Figure 6.22 shows the results of the power analysis.

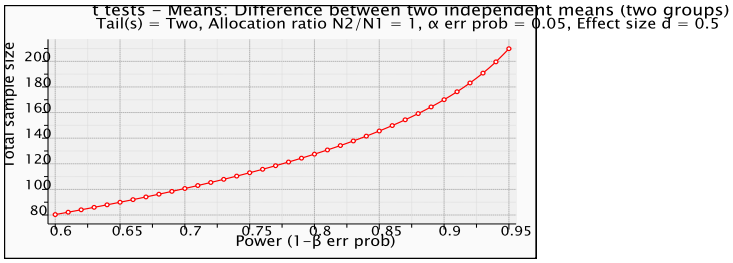


Figure 6.22 Total Sample Size vs. Power

5) Estimate the treatment effect. The results of these calculations are shown in Table 6.16. To start this process a "medium" treatment effect is assumed. Figure 6.23 depicts the distribution of the errors.

Table 6.16 A Priori Required Sample Size

Analysis: A priori	**Compute required sample size**
Input: Tail(s)	2
Effect size d	0.5 (assumed medium)
A error probability	0.05
Power (1- β error probability)	0.90
Allocation ratio N2/N1	1
Output: Noncentrality parameter δ	3.278719
Critical t	1.974017
Df	170
Sample size treatment 1	86
Sample size treatment 2	86
Total sample size	172

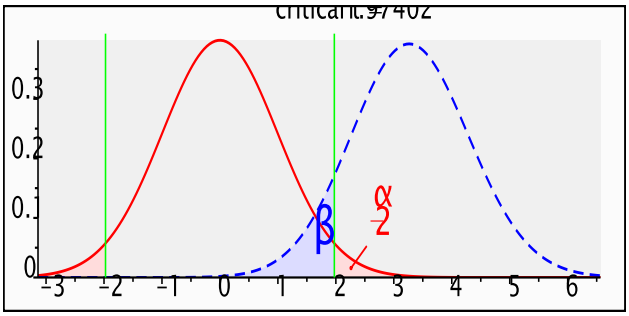


Figure 6.23 Error Distribution

As a result of the power analysis and estimating the treatment effects, conducting three interim analyses with 29 samples for each treatment would achieve 174 total samples.

6) Select the type of stopping boundary and the stopping criteria. A group sequential design that combines efficacy and futility testing is selected. The O'Brien-Flemming stopping boundary is used for efficacy, upper stopping boundary, and the Pocock boundary is used for futility, lower stopping boundary. A design based error spending function is applied. Stopping boundaries are selected to protect Type I and Type II error. Simultaneous efficacy and futility is applied as follows:

    1) Stop the trial for efficacy if $Z_k > u_k(\alpha, ß)$
    2) Continue the trial if $l_k(\alpha, ß) \leq Z_k \leq \mu_k(\alpha, ß)$
    3) Stop the trial for futility if $Z_k < l_k(\alpha, ß)$

In step 2, select the DOE (range of factor settings to explore). For the first treatment, the current operating conditions will be applied which employs one patrol in each district. The NOLHD design will be used due to its efficiency and space-filling

properties. The NOLHD uses the full range of allowable factor settings in order to initially examine the entire region of feasible operating conditions since there is no prior knowledge of how the process performs over varying operating conditions. Given the eight factors, the NOLHD EVOP has 33 design points. Table 6.17 shows the Phase I design for each design point.

In step 3, the Police Staffing Simulation is set up and run given the EVOP experimental design chosen. (Constitutes a complete cycle.)

In step 4, two replications are conducted and the data are collected. (Constitutes the phase.)

In step 5, a Gaussian process model is fit to the results given the three responses and the response prediction formula for each of these is estimated. See Appendix C for the response prediction formula for each response as a result of fitting the Gaussian process model. The Gaussian process model report can be found in Appendix D. The model report highlights the significant factors and interactions. It was found that all factors except for District 6 are significant as well as their interactions in determining the average response time and average occupied time. Districts 4, 5, 7 and 8 are the significant factors in determining cost, none of the interactions are significant in determining cost.

Table 6.17 Phase I NOLHD EVOP Design

Low Level # of Patrols	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**Low Level # of Patrols**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
High Level # of Patrols	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**High Level # of Patrols**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**
District / Design Point	1	2	3	4	5	6	7	8	**District / Design Point**	1	2	3	4	5	6	7	8
1	4	0	2	1	4	3	3	2	18	0	4	2	3	1	2	1	2
2	4	4	1	2	2	1	2	3	19	0	0	4	3	2	3	2	1
3	4	2	4	1	0	2	4	2	20	1	2	0	3	4	2	0	3
4	2	4	4	2	4	1	0	1	21	2	1	0	2	0	3	4	3
5	4	0	2	1	3	3	1	3	22	0	4	2	3	1	1	4	1
6	4	4	1	1	2	1	2	3	23	0	0	3	3	2	3	2	1
7	3	2	4	1	0	3	0	1	24	1	2	0	3	4	1	4	3
8	2	3	4	1	4	1	4	1	25	2	1	0	3	0	3	0	3
9	3	1	1	2	3	1	2	0	26	1	3	3	2	1	3	2	4
10	3	3	1	3	1	2	1	0	27	1	1	3	1	3	2	3	4
11	3	1	3	4	1	0	3	3	28	1	3	1	0	3	4	2	1
12	3	3	3	4	3	4	1	2	29	1	1	1	0	1	0	3	2
13	2	1	1	2	2	1	1	0	30	2	3	3	2	2	4	3	4
14	3	2	2	4	1	2	3	1	31	1	2	3	1	3	2	1	4
15	3	1	3	4	2	0	1	4	32	2	3	1	0	3	4	3	0
16	3	3	2	4	3	4	3	2	33	1	2	2	0	1	0	1	2
17	2	2	2	2	2	2	2	2									

Table 6.18 summarizes the significant factors by response but doesn't include the interactions for brevity. See Appendix D for the complete model report.

Table 6.18 Phase I Significant Factors by Response

Response/Factor	Avg Response Time	Avg % Time Occupied	Avg Total Cost
1	X	X	
2	X	X	
3	X	X	
4	X	X	X
5	X	X	X
6			
7	X	X	X
8	X	X	X

All of the response surfaces with respect to the response and their significant factor interactions were examined. The following Phase I RSMs for each response with each of the factor interactions are shown in Figure 6.24, Figure 6.27, and Figure 6.30:

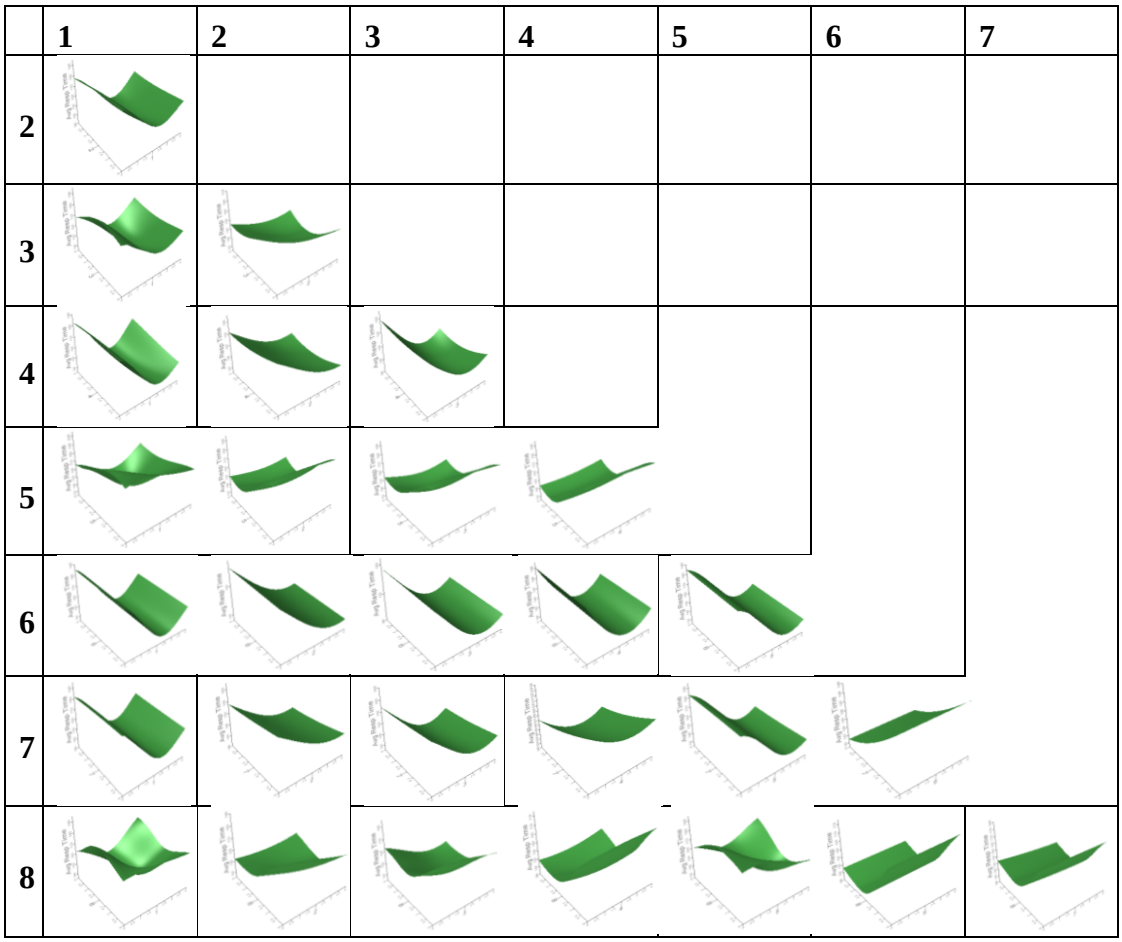


Figure 6.24 Phase I Average Response Time NOLHD RSM

Figure 6.25, Figure 6.28 and Figure 6.31, demonstrates the adequacy of the model fit to the data. For the average response time and the total time occupied the data falls further from the line as the average response time increases. Figure 6.31 gives us no information on the cost, therefore in subsequent analyses; the number of replications will be increased to five to obtain better fits and reduce variance.

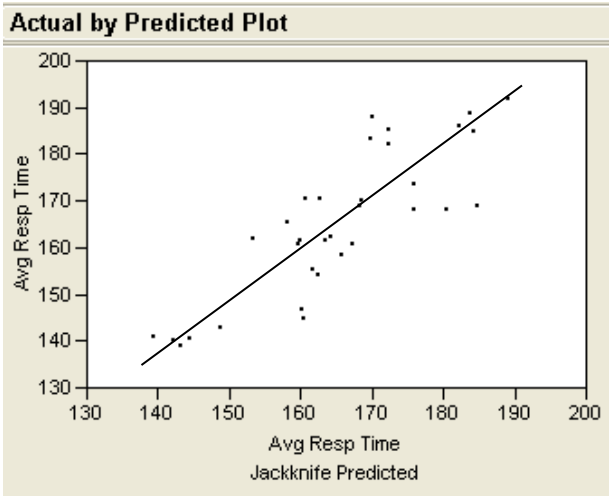


Figure 6.25 Phase I Average Response Time Actual by Predicted Plot

The prediction profiler for each response is also examined to help find the best factor settings to optimize the responses. Figure 6.26, Figure 6.29 and Figure 6.31 helps predict the best settings to achieve the optimal responses.

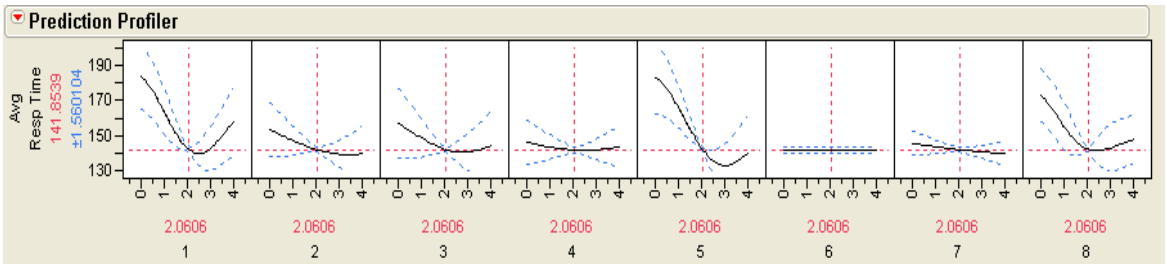


Figure 6.26 Phase I Average Response Time Prediction Profiler

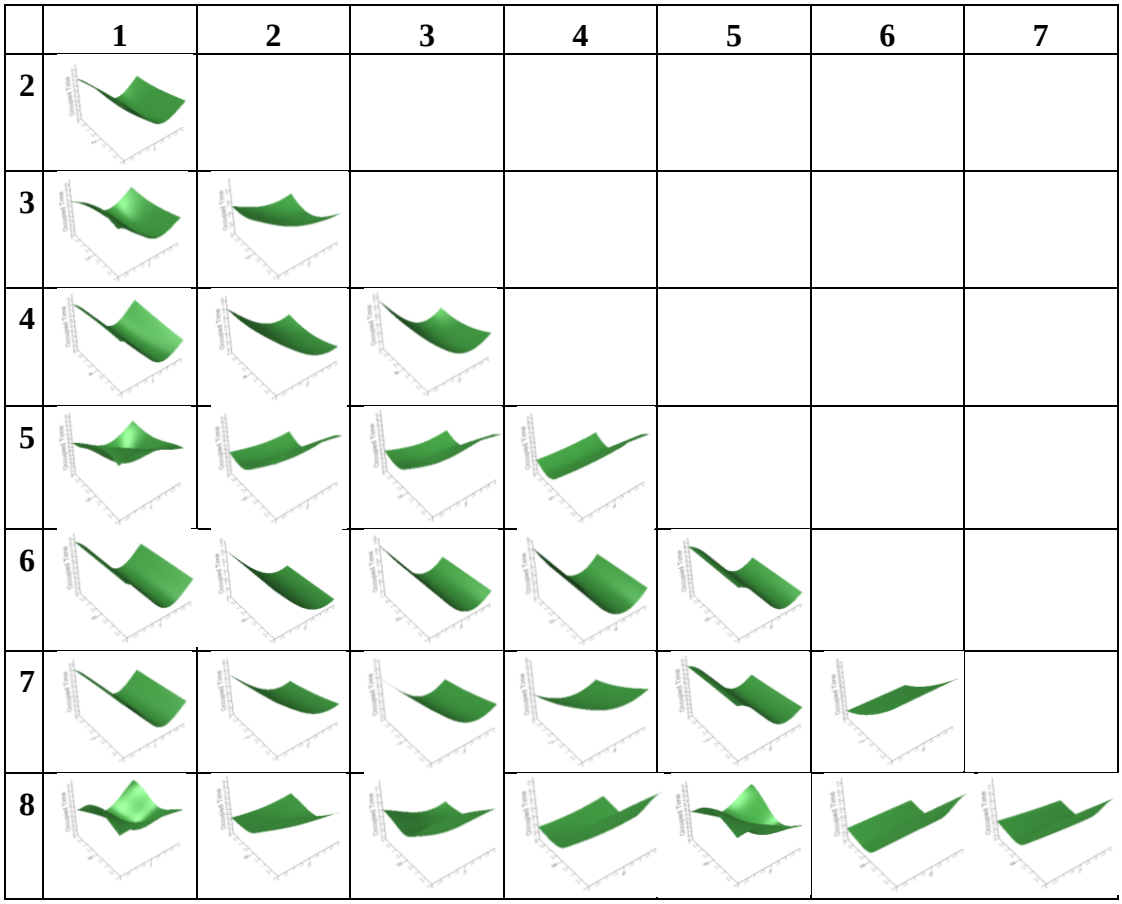


Figure 6.27 Phase I Percent Occupied Time NOLHD RSM

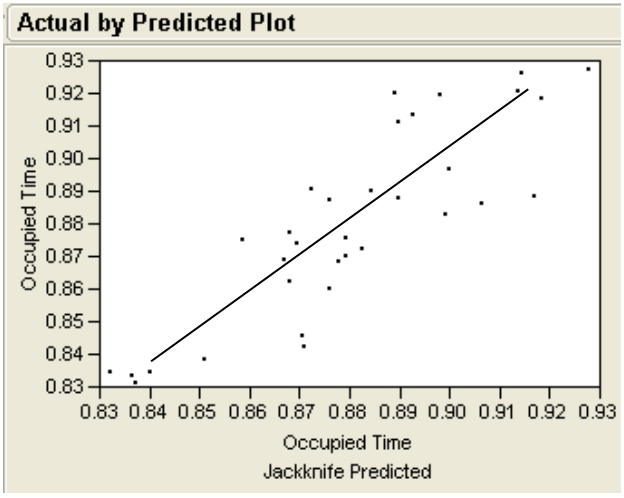


Figure 6.28 Phase I Average Percent Total Time Occupied Actual by Predicted Plot

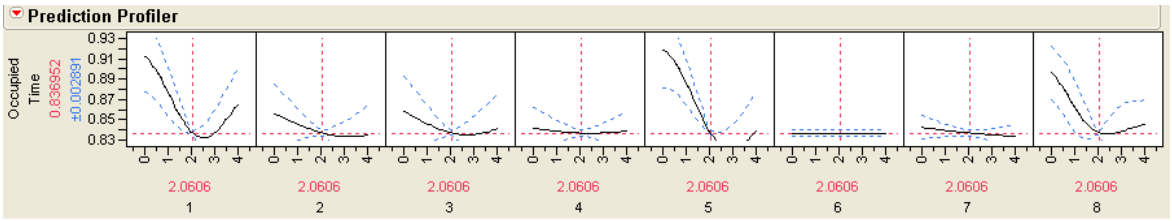


Figure 6.29 Phase I Average Percent Total Time Occupied Prediction Profiler

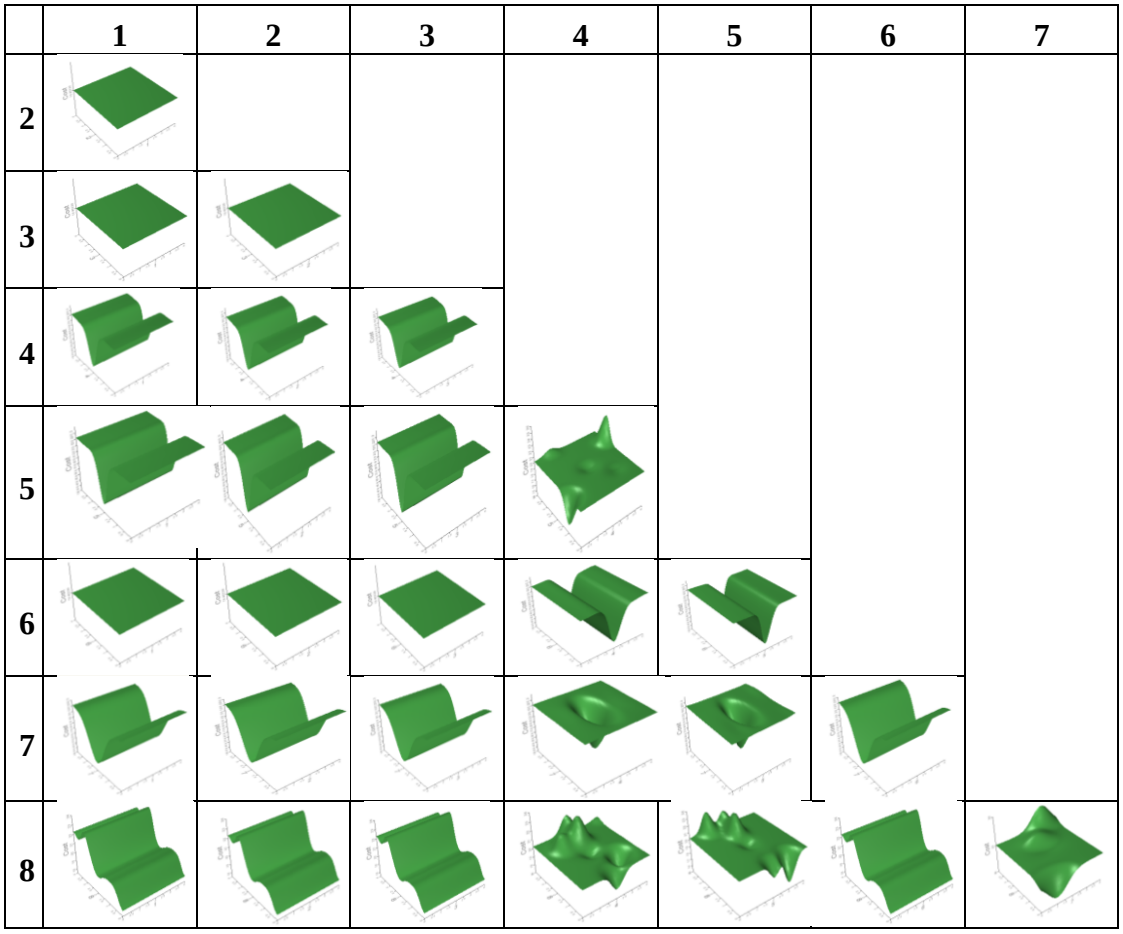


Figure 6.30 Phase I Average Total Cost NOLHD RSM

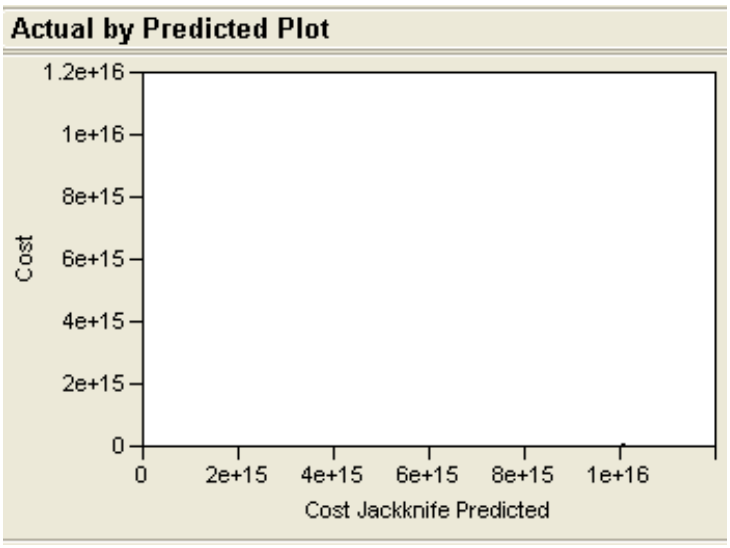


Figure 6.31 Phase I Average Total Cost Actual by Predicted Plot

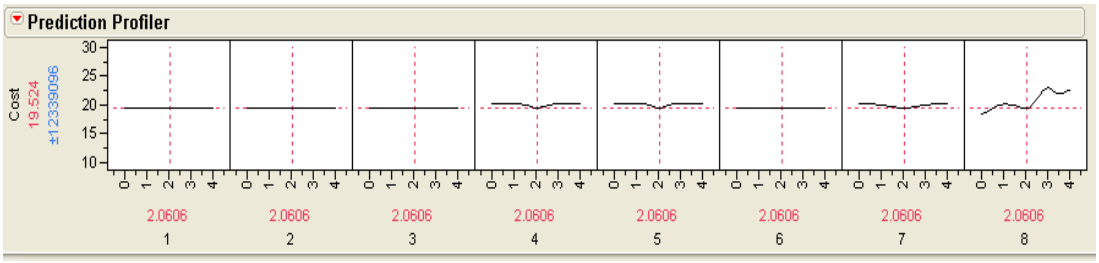


Figure 6.32 Phase I Average Total Cost Prediction Profiler

In step 6, optimize the response prediction formula using SA and multiple objective optimizations. The response prediction formulas are optimized using SA in order to maximize the responses and get the optimal solution and factor settings for this model. SA also gives us the bounds on the optimal responses for average % time occupied and average total cost in order to apply non-linear programming. Applying SA to the average response time response prediction formula indicates a minimum average response time when 2 – 3 patrols are assigned in each area. Applying SA to the average percent time occupied response prediction formula indicates a maximum average percent

time occupied is when the overall number of patrols is minimized. Applying SA to the average total cost response prediction formula indicates a minimum average total cost when the overall number of patrols is minimized especially in districts 4, 5, 7 and 8.

The SA results are shown in Table 6.19. Next, the three response prediction formulas, given the constraints on the responses, are solved simultaneously to find a solution that meets the three response objectives. This is done by prioritizing the average response time response as the main response to minimize and the other two responses are set to meet a goal. The goal is set based on the best results achieved thus far. The optimization problem and the constraints are as follows:

Minimize Average Response Time (6.1)

such that

$0 \leq \#\, patrols \leq 4$

Total # patrols ≤ 9

average total time occupied ≥ 0.85

cost ≤ 15

While the current policy uses 8 patrols, the goal of nine or fewer is used in the event that significant results for minimizing the response time can be achieved with nine cars over the eight cars while incurring a minimal additional cost.

Table 6.19 First Multi-Objective Optimization of Response Prediction Formulas

Optimal Patrol Policy (Obj)		Avg Response Time Predicted Value	Avg % Time Occupied Predicted Value	Avg Total Cost Predicted Value
SA Results	2 3 2 2 3 4 4 2 (Min response Time)	130.42	0.817	20.21
	1 1 1 0 0 0 0 1 (Max Percent Occupied Time)	190.19	0.923	20.18
	1 3 0 0 1 2 3 2 (Min Cos)	183.54	0.914	11.35
Multiple Objective Optimization	0 4 0 0 1 0 1 2	183.59	0.92	11.39
	4 1 0 0 1 0 1 2	192.42	0.93	20.21
	3 1 0 0 1 0 1 2	169.51	0.89	11.39

Step 7) From the above information, Treatment #2 is selected. Policies with more than nine cars are ruled out due to the expense of adding additional patrols (Nine patrols will be examined later). It is assumed that consideration for adding one more patrol for a total of nine cars could be beneficial if a drastically reduced response time over the eight car policy can be achieved in spite of the slightly higher cost. Strategy 3 1 0 0 1 0 1 2 is assigned to Treatment #2 as a result of its lower response time and relatively low cost than the other policies.

Step 8) The current staffing policy of one patrol in each district is assigned as Treatment #1.

In step 9, ten replications are conducted to collect the data to do the interim

analysis on the current factor settings, Treatment #1, and the new factor settings, Treatment #2. The two treatment results follow in Table 6.20.

Table 6.20 First Comparison of Treatment Results After 10 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 10 Reps Simulation Results	1 1 1 1 1 1 1 1	194.68 (109.52)	0.937 (0.017)	10.23 (0.061)
Treatment #2 10 Reps Simulation Results	3 1 0 0 1 0 1 2	217.56 (129.34)	0.978 (0.019)	10.37 (0.070)

In step 10, interim analysis is applied to the Treatment #1 and Treatment #2 data. The interim analysis indicates that further replications are needed to reach a conclusion about the average response time. It also indicates that the stopping boundary for futility is reached with respect to the % time occupied and the average cost responses as the lower stopping boundary is crossed. Each treatment is run for ten additional replications for a total of 20 replications (results in Table 6.21) and the interim analysis is repeated.

Table 6.21 First Comparison of Treatment Results After 20 Replications

Results	Police Staffing Policy	Average Response Time (STD)
Treatment #1 20 Reps Simulation Results	1 1 1 1 1 1 1 1	191.74 (108.43)
Treatment #2 20 Reps Simulation Results	3 1 0 0 1 0 1 2	218.46 (128.93)

After 20 replications the trial is stopped for futility with regards to the average response time as the lower stopping boundary is crossed. The results of the interim analyses are summarized in Table 6.22.

Table 6.22 First Interim Analysis Results

Analysis	Response	Test Statistic	Lower Stopping Boundary	Upper Stopping Boundary	Decision
1 (10 reps)	Average Response Time	-0.4269	-1.046	3.6321	Continue
	Average % Time Occupied	-5.085	-1.046	3.6321	Stop trial for futility
	Average Total Cost	-4.768	-1.046	3.6321	Stop trial for futility
2 (20 reps)	Average Response Time	-0.7093	-.4576	3.6084	Stop trial for futility

As described in section 6.1.1, the interim analysis results are interpreted as follows: stop the trial for efficacy if $\mathbf{Z}_k > \mathbf{u}_k(\alpha, ß)$, continue the trial if $\mathbf{l}_k(\alpha, ß) \leq \mathbf{Z}_k \leq \boldsymbol{\mu}_k(\alpha, ß)$ and stop the trial for futility if $\mathbf{Z}_k < \mathbf{l}_k(\alpha, ß)$. Figure 6.33 graphically depicts the interim analysis results, which show after the second analysis that the trial is stopped for futility as $\mathbf{Z}_k < \mathbf{l}_k(\alpha, ß)$. These results indicate a failure to reject the null hypothesis of no statistical difference between Treatment #1 and Treatment #2 for all three responses.

The trial comparing Treatment #1 and #2 is stopped and the response surface is analyzed to see if there are improvements that can be made elsewhere in the design region and if a phase II EVOP design can be selected.

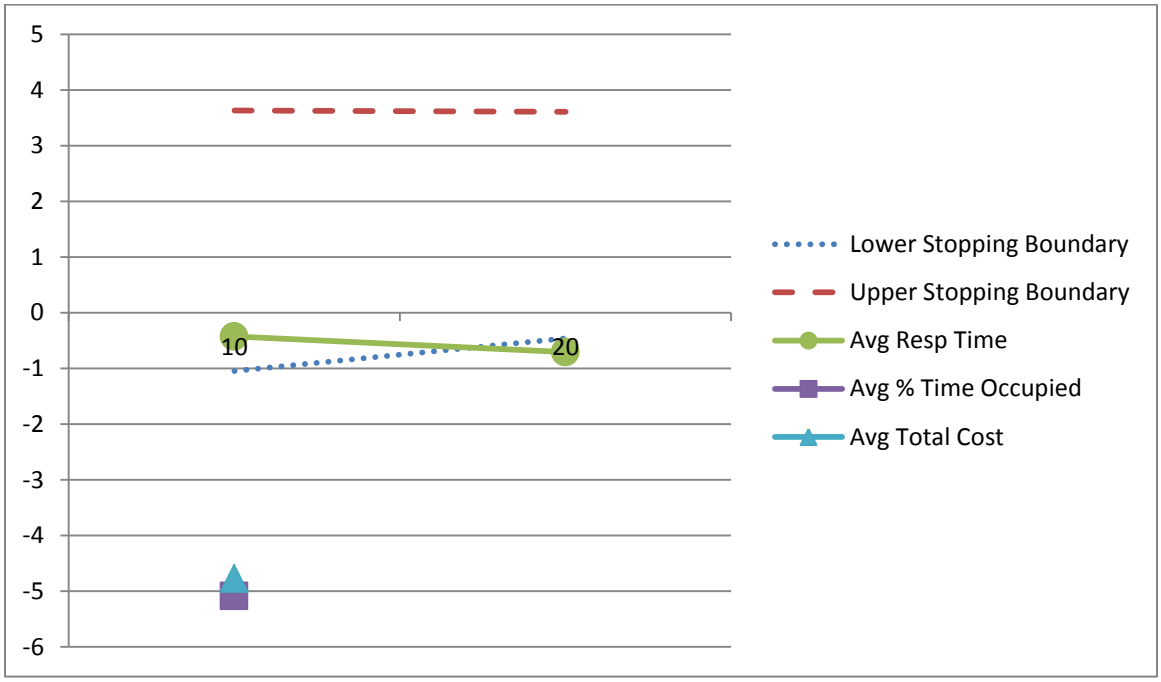


Figure 6.33 Police Staff Study Efficacy and Futility Monitoring Boundaries

For the phase II analysis, the factor levels are reduced between zero and three patrols per district and the process is repeated. The current condition of one car per district, Treatment #1, remains the benchmark for comparison since phase I failed to produce a better result. The process starts again beginning at step 3 and a new EVOP scheme is found and applied. The Police Staffing Simulation is set up and run given the EVOP experimental design shown in Table 6.23.

Table 6.23 Phase II NOLHD EVOP Design

Low Level # of Patrols	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**Low Level # of Patrols**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
High Level # of Patrols	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**High Level # of Patrols**	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**3**
District / Design Point	1	2	3	4	5	6	7	8	**District / Design Point**	1	2	3	4	5	6	7	8
1	3	0	1	1	3	2	2	2	18	0	3	2	2	0	1	1	1
2	3	3	0	1	1	1	1	3	19	0	0	3	2	2	2	2	0
3	3	1	3	0	0	2	3	1	20	0	2	0	3	3	1	0	2
4	2	3	3	1	3	0	0	1	21	1	0	0	2	0	3	3	2
5	3	0	1	1	2	2	0	2	22	0	3	2	2	1	1	3	1
6	3	3	1	1	1	1	2	2	23	0	0	2	2	2	2	1	1
7	2	1	3	1	0	2	0	1	24	1	2	0	2	3	1	3	2
8	2	2	3	1	3	1	3	1	25	1	1	0	2	0	2	0	2
9	2	1	1	2	2	1	2	0	26	1	2	2	1	1	2	1	3
10	2	2	1	2	1	2	1	0	27	1	1	2	1	2	1	2	3
11	2	1	2	3	1	0	2	2	28	1	2	1	0	2	3	1	1
12	2	2	2	3	2	3	0	2	29	1	1	1	0	1	0	3	1
13	2	0	1	2	2	0	1	0	30	1	3	2	1	1	3	2	3
14	3	2	1	3	0	2	2	0	31	0	1	2	0	3	1	1	3
15	2	1	3	3	1	0	1	3	32	1	2	0	0	2	3	2	0
16	2	2	2	3	2	3	2	2	33	1	1	1	0	1	0	1	1
17	2	2	2	2	2	2	2	2									

Five replications of the EVOP design are conducted. A new Gaussian process model is fit and new response prediction formulas are found for each response. See Appendix C for the response prediction formula for each response as a result of fitting the Gaussian process model. The Gaussian process model report can be found in Appendix D. It was found that for the average Response Time, district 2 and 4 are not significant.

For the average % time occupied, district 1, 3 and 6 are not significant. All districts are significant and contribute to the average total cost formula. Table 6.24 summarizes the significant factors by response but doesn't include the interactions for brevity.

Table 6.24 Phase II Significant Factors by Response

Response/Factor	Avg Response Time	Avg % Time Occupied	Avg Total Cost
1	X		X
2		X	X
3	X		X
4		X	X
5	X	X	X
6	X		X
7	X	X	X
8	X	X	X

All of the response surfaces with respect to the response and their significant factor interactions were examined. The RSMs for each response and factor interactions are demonstrated in Figure 6.34, Figure 6.37, and Figure 6.40:

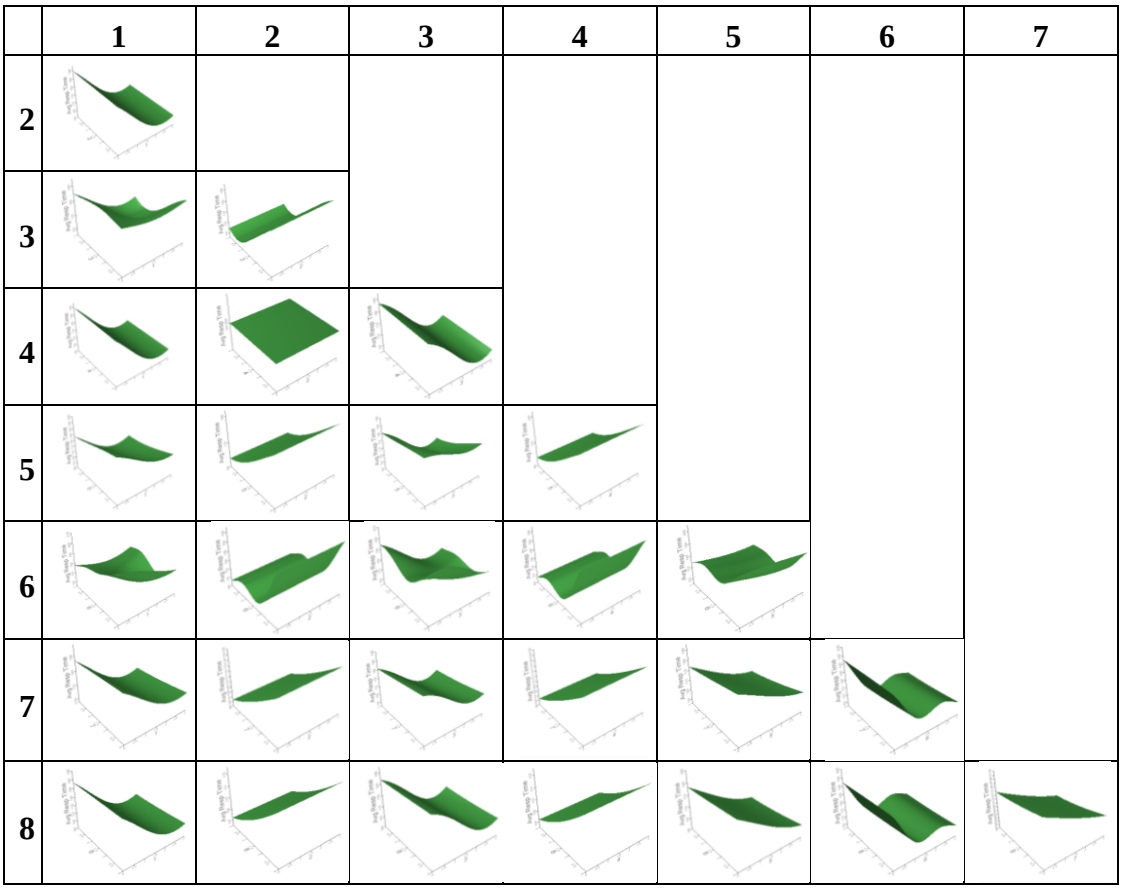


Figure 6.34 Phase II Average Response Time NOLHD RSM

Figure 6.35, Figure 6.38 and Figure 6.41 demonstrates the adequacy of the model fit to the data. As can be seen, the model appears to be a good fit for the data for each response. The prediction profiler for each response is shown in Figure 6.36, Figure 6.39 and Figure 6.42.

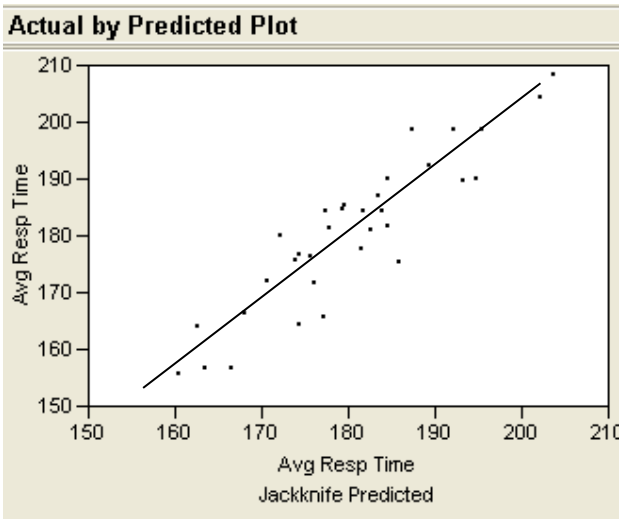


Figure 6.35 Phase II Average Response Time Actual by Predicted Plot

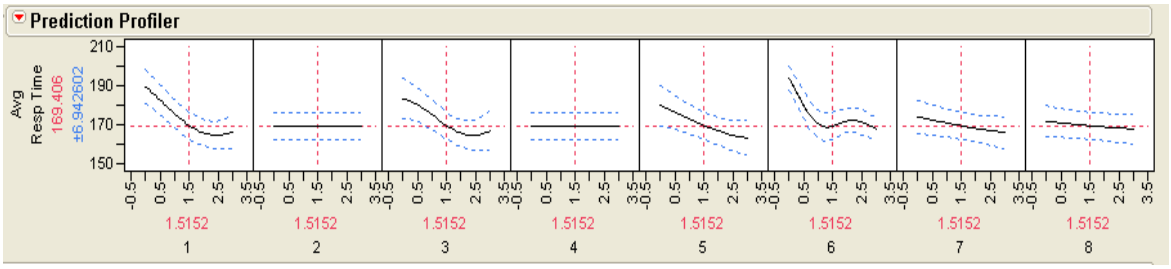


Figure 6.36 Phase II Average Response Time Prediction Profiler

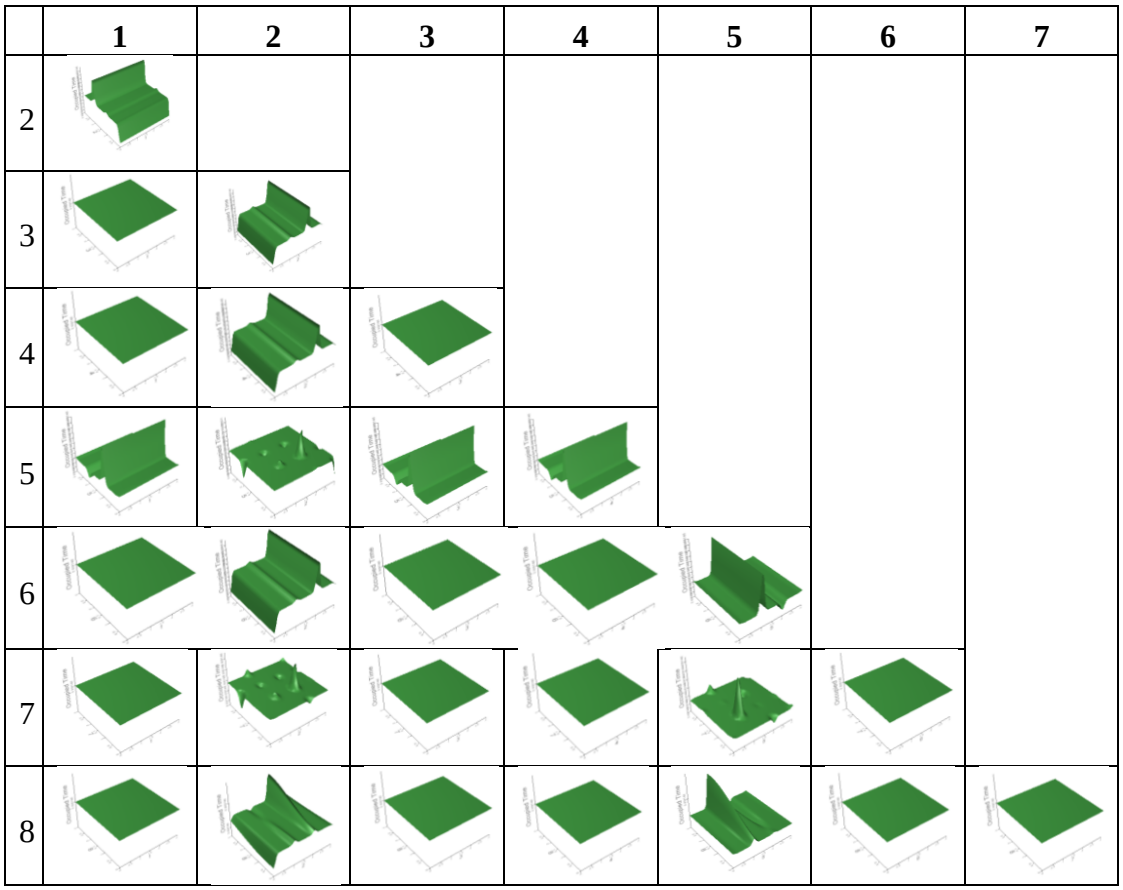


Figure 6.37 Phase II Average Percent Occupied Time NOLHD RSM

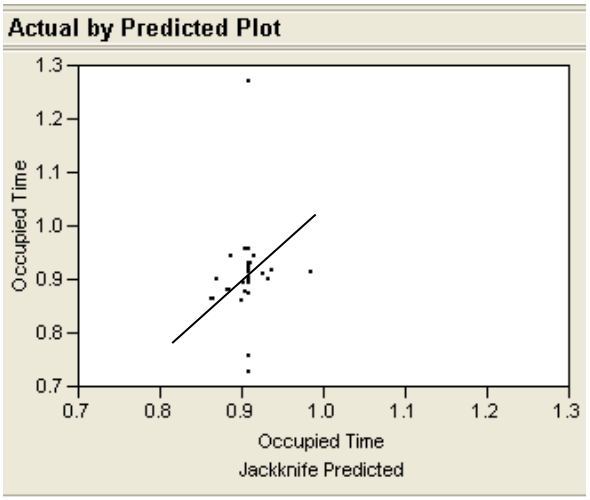


Figure 6.38 Phase II Average Percent Total Time Occupied Actual by Predicted Plot

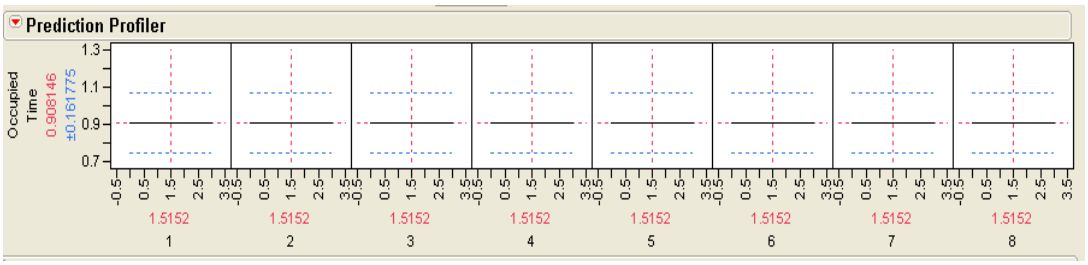


Figure 6.39 Phase II Average Percent Total Time Occupied Prediction Profiler

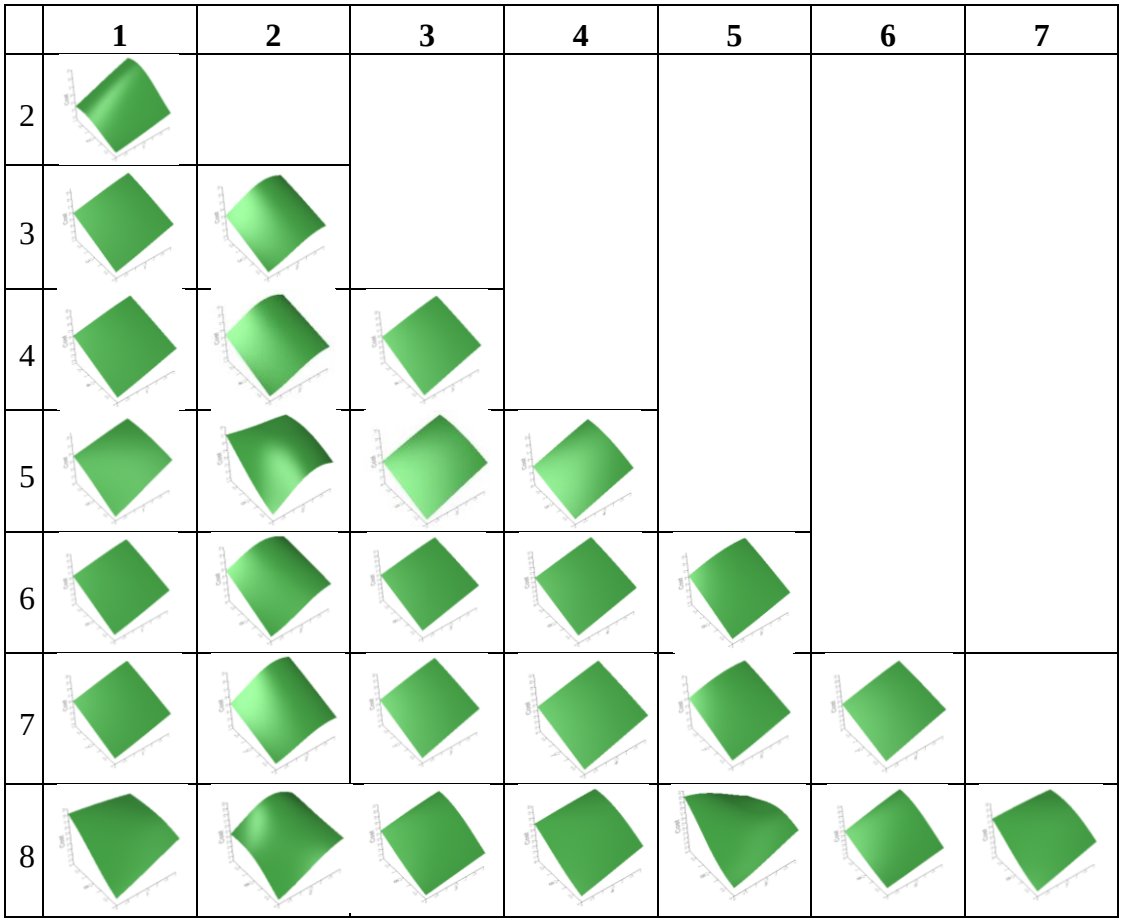


Figure 6.40 Phase II Average Total Cost NOLHD RSM

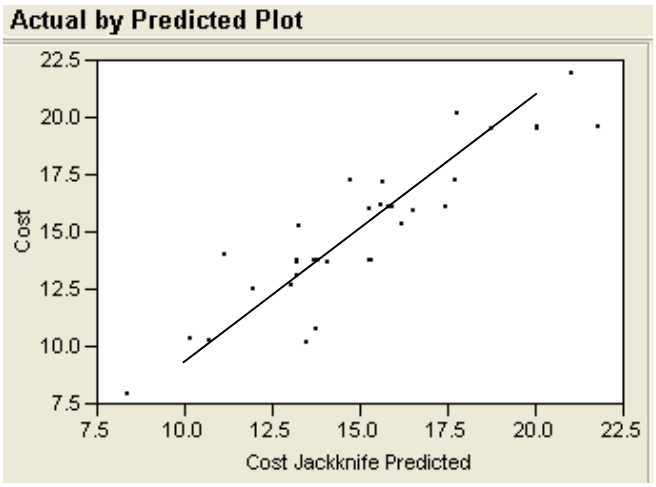


Figure 6.41 Phase II Average Total Cost Actual by Predicted Plot

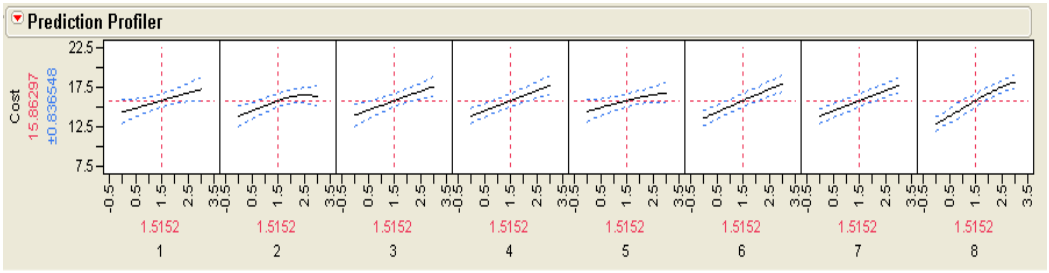


Figure 6.42 Phase II Average Total Cost Prediction Profiler

The response prediction formulas are optimized using SA and non-linear programming. The best results for each response are summarized in Table 6.25.

Table 6.25 Phase II Multi-Objective Optimization of Response Prediction Formulas

Optimal Patrol Policy by Obj		**Avg Response Time Predicted Value**	**Avg %Time Occupied Predicted Value**	**Avg Total Cost Predicted Value**
SA Results	2 1 2 0 3 1 3 3 (Min Avg Response Time)	152.74	0.908	14.19
	2 2 1 1 1 1 1 3 (Max % Time Occupied)	173.48	1.27	17.78
	0 1 0 0 0 0 0 0 (Min Total Cost)	204.03	0.897	4.03
Multiple Objective Optimization	2 0 2 02 1 1 0	162.32	0.943	9.65
	2 0 2 0 2 1 0 0	165.14	0.912	9.02
	2 0 2 0 1 1 0 0	170.21	0.908	7.67

The best strategy for Treatment #2 is 2 0 2 0 2 1 1 0. The current staffing policy of one patrol in each district remains as Treatment #1. Ten replications are conducted with the new factor settings and assigned as Treatment #2. The results follow in Table 6.26.

Table 6.26 Second Comparison of Treatment Results After 10 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 10 Reps Simulation Results	1 1 1 1 1 1 1 1	194.68 (109.52)	0.937 (0.017)	10.23 (0.061)
Treatment #2 10 Reps Simulation Results	2 0 2 0 2 1 1 0	199.56 (96.28)	0.944 (0.014)	10.25 (0.056)

After ten replications, the interim analysis results indicate that no conclusions can be drawn for any of the responses and the trial should continue to collect more data. Ten additional replications are run and the results are shown in Table 6.27.

Table 6.27 Second Comparison of Treatment Results After 20 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 20 Reps Simulation Results	1 1 1 1 1 1 1 1	191.74 (109.52)	0.931 (0.019)	10.21 (0.072)
Treatment #2 20 Reps Simulation Results	2 0 2 0 2 1 1 0	198.93 (96.52)	0.943 (0.016)	10.24 (0.066)

The interim analysis results indicate that the trial should continue to further evaluate the response time. The interim analysis results also indicate that the trial should be stopped for futility in further evaluating the average % time occupied and average total cost. Ten more replications are conducted for a total of 30 replications. The results for the average response time are shown in Table 6.28.

Table 6.28 Second Comparison of Treatment Results After 30 Replications

Results	Police Staffing Policy	Average Response Time (STD)
Treatment #1 30 Reps Simulation Results	1 1 1 1 1 1 1 1	192.31 (105.05)
Treatment #2 30 Reps Simulation Results	2 0 2 0 2 1 1 0	200.96 (96.70)

These results indicate that the trial should be stopped for futility as the null hypothesis cannot be rejected. There is no statistical difference between Treatment #1 and Treatment #2 for all three responses. The results of the interim analyses are summarized in Table 6.29.

Table 6.29 Second Interim Analysis Results

Analysis	Response	Test Statistic	Lower Stopping Boundary	Upper Stopping Boundary	Decision
1 (10 reps)	Average Response Time	-0.1058	-1.046	3.6321	Continue
	Average % Time Occupied	-1.005	-1.046	3.6321	Continue
	Average Total Cost	-0.7638	-1.046	3.6321	Continue
2 (20 reps)	Average Response Time	-0.2215	-0.4576	3.6084	Continue
	Average % Time Occupied	-2.160	-0.4576	3.6084	Stop trial for futility
	Average Total Cost	-1.374	-0.4576	3.6084	Stop trial for futility
3 (30 reps)	Average Response Time	-0.3318	0.0114	3.6921	Stop trial for futility

The trial comparing Treatment #1 and Treatment #2 is stopped and the response surface is analyzed to see if there are improvements that can be made elsewhere in the design region. Figure 6.43 graphically depicts these results.

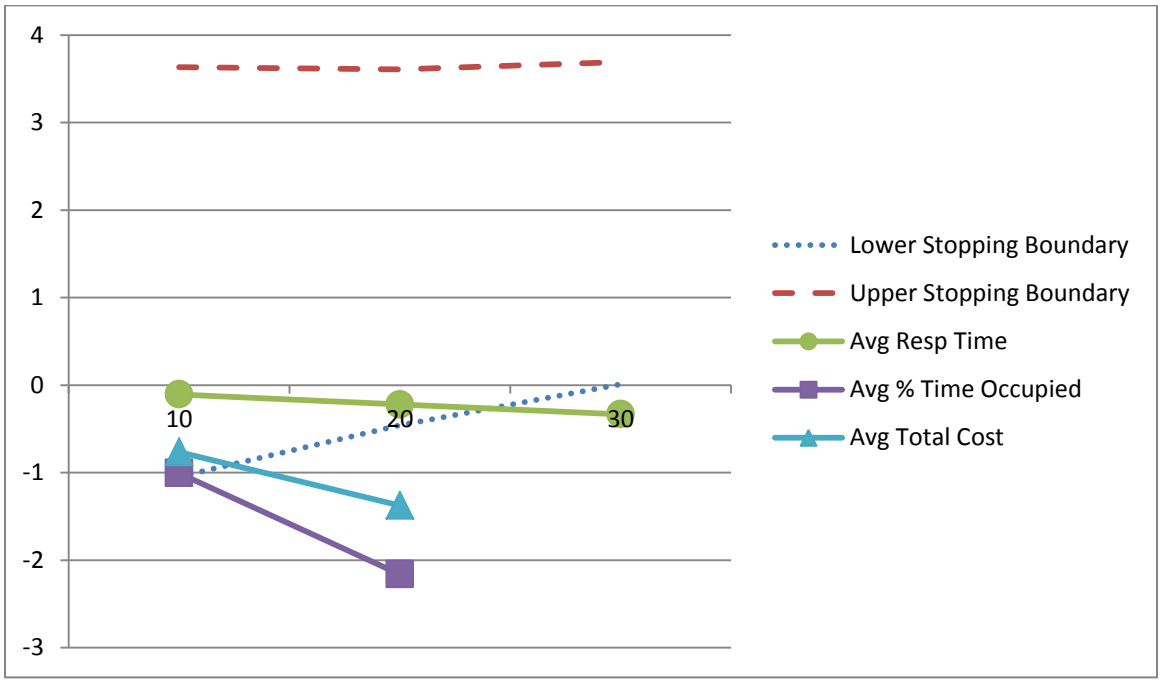


Figure 6.43 Second Efficacy and Futility Monitoring Boundaries

The Phase III EVOP design narrows the factors between zero and two patrols in each district and is shown in Table 6.30.

Table 6.30 Phase III NOLHD EVOP Design

Low Level # of Patrols	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**Low Level # of Patrols**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
High Level # of Patrols	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**High Level # of Patrols**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
District / Design Point	1	2	3	4	5	6	7	8	**District / Design Point**	1	2	3	4	5	6	7	8
1	2	0	1	0	2	1	1	1	18	0	2	1	2	0	1	1	1
2	2	2	0	1	1	0	1	2	19	0	0	2	1	1	2	1	0
3	2	1	2	0	0	1	2	1	20	0	1	0	2	2	1	0	1
4	1	2	2	1	2	0	0	0	21	1	0	0	1	0	2	2	2
5	2	0	1	0	1	1	0	1	22	0	2	1	2	1	1	2	1
6	2	2	1	1	1	0	1	2	23	0	0	1	1	1	2	1	1
7	1	1	2	1	0	1	0	1	24	1	1	0	2	2	1	2	1
8	1	1	2	1	2	1	2	0	25	1	1	0	1	0	2	0	2
9	1	1	0	1	1	1	1	0	26	1	2	2	1	1	1	1	2
10	2	1	1	1	0	1	0	0	27	1	1	1	1	2	1	2	2
11	1	0	2	2	1	0	1	1	28	1	2	1	0	1	2	1	1
12	2	1	1	2	2	2	0	1	29	0	1	1	0	1	0	2	1
13	1	0	0	1	1	0	1	0	30	1	2	2	1	1	2	1	2
14	2	1	1	2	0	1	2	0	31	0	1	1	0	2	1	0	2
15	1	0	2	2	1	0	1	2	32	1	2	0	0	1	2	1	0
16	2	1	1	2	2	2	2	1	33	0	1	1	0	0	0	1	1
17	1	1	1	1	1	1	1	1									

Five replications are conducted and the data are collected, a Gaussian process model is fit to the results given the three responses and the response prediction formula for each of these is estimated. (See Appendix C and D) In phase III all factors are found to be significant for all responses and are summed up on Table 6.31.

Table 6.31 Phase III Significant Factors by Response

Response/Factor	Avg Response Time	Avg % Time Occupied	Avg Total Cost
1	X	X	X
2	X	X	X
3	X	X	X
4	X	X	X
5	X	X	X
6	X	X	X
7	X	X	X
8	X	X	X

All of the RSMs with respect to the response and their significant factor interactions were examined. The following RSMs for each response are demonstrated in Figure 6.44, Figure 6.47 and Figure 6.50:

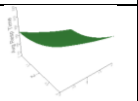
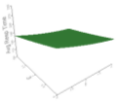
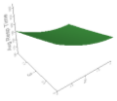
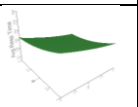
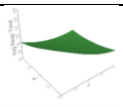
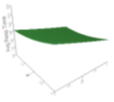
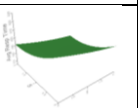
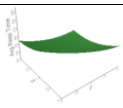
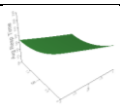
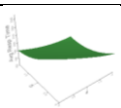
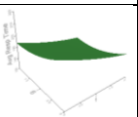
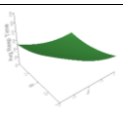
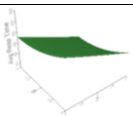
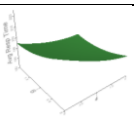
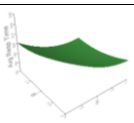
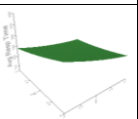
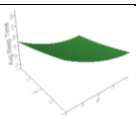
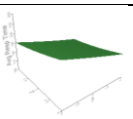
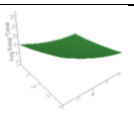
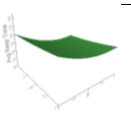
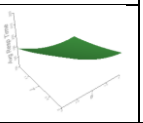
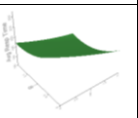
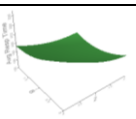
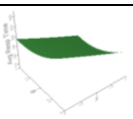
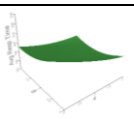
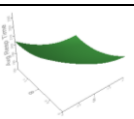
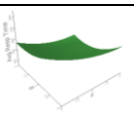
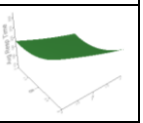
	1	2	3	4	5	6	7
2							
3							
4							
5							
6							
7							
8							

Figure 6.44 Phase III Average Response Time NOLHD RSM

Figure 6.45, Figure 6.48 and Figure 6.51 demonstrates the adequacy of the model fit to the data. As can be seen, the model appears to be a good fit for the data for each response. The prediction profiler for each response is shown in Figure 6.46, Figure 6.49 and Figure 6.52.

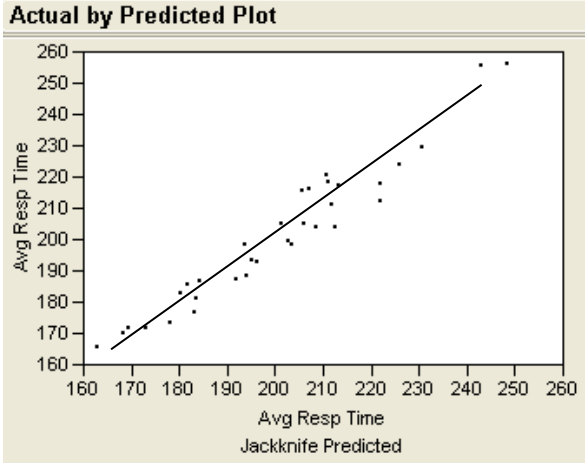


Figure 6.45 Phase III Average Response Time Actual by Predicted Plot

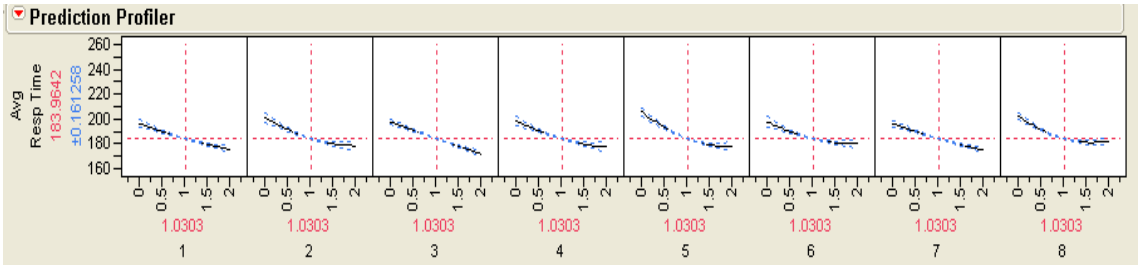


Figure 6.46 Phase III Average Response Time Prediction Profiler

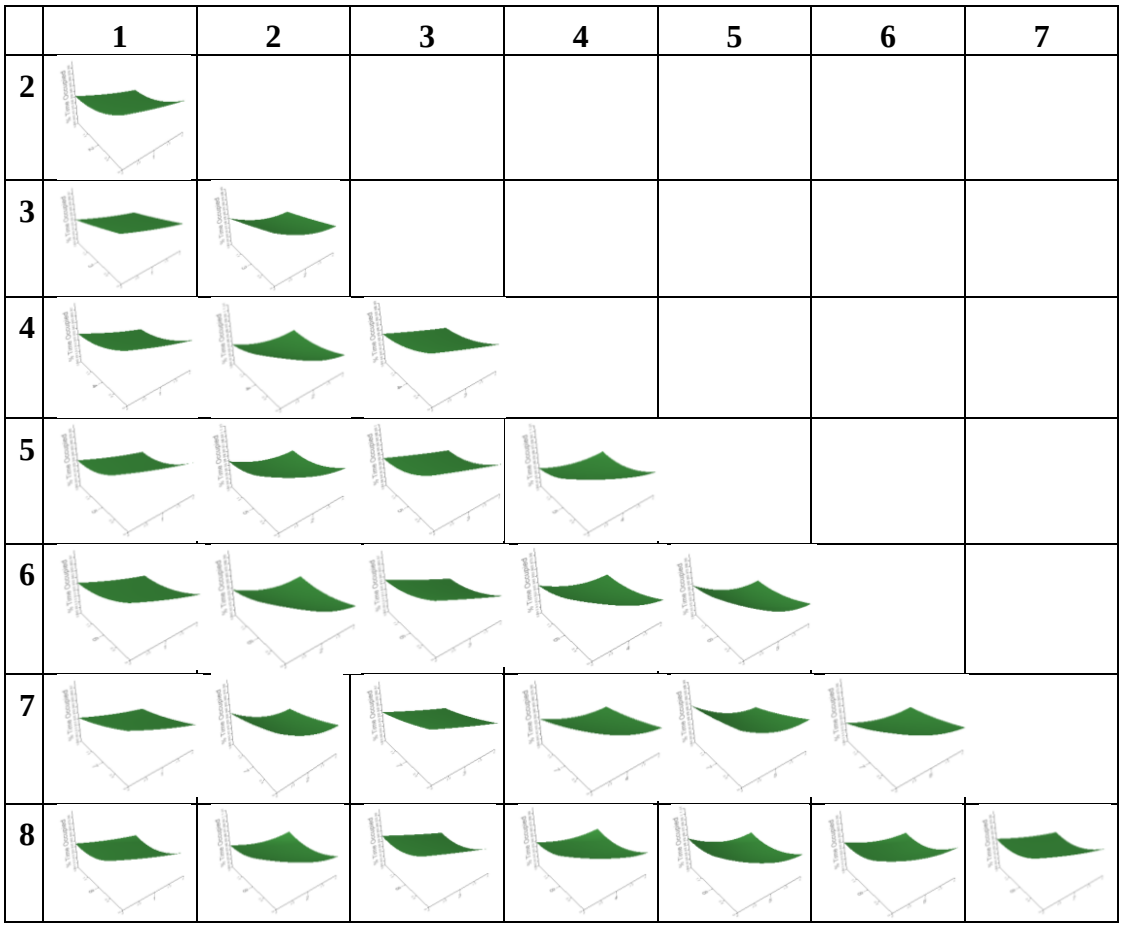


Figure 6.47 Phase III Average Percent Occupied Time NOLHD RSM

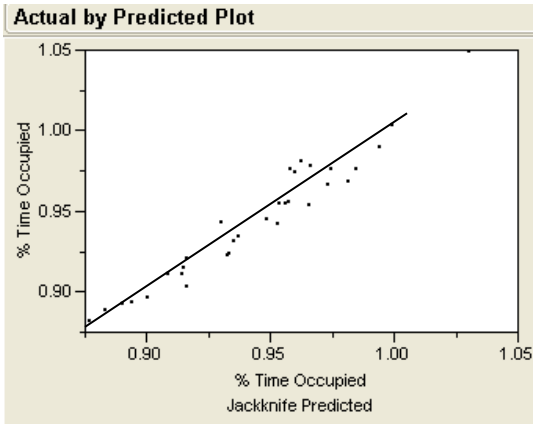


Figure 6.48 Phase III Average Percent Total Time Occupied Actual by Predicted Plot

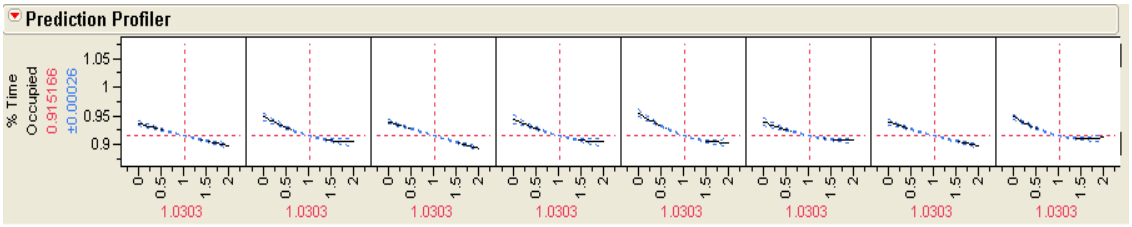


Figure 6.49 Phase III Average Percent Total Time Occupied Prediction Profiler

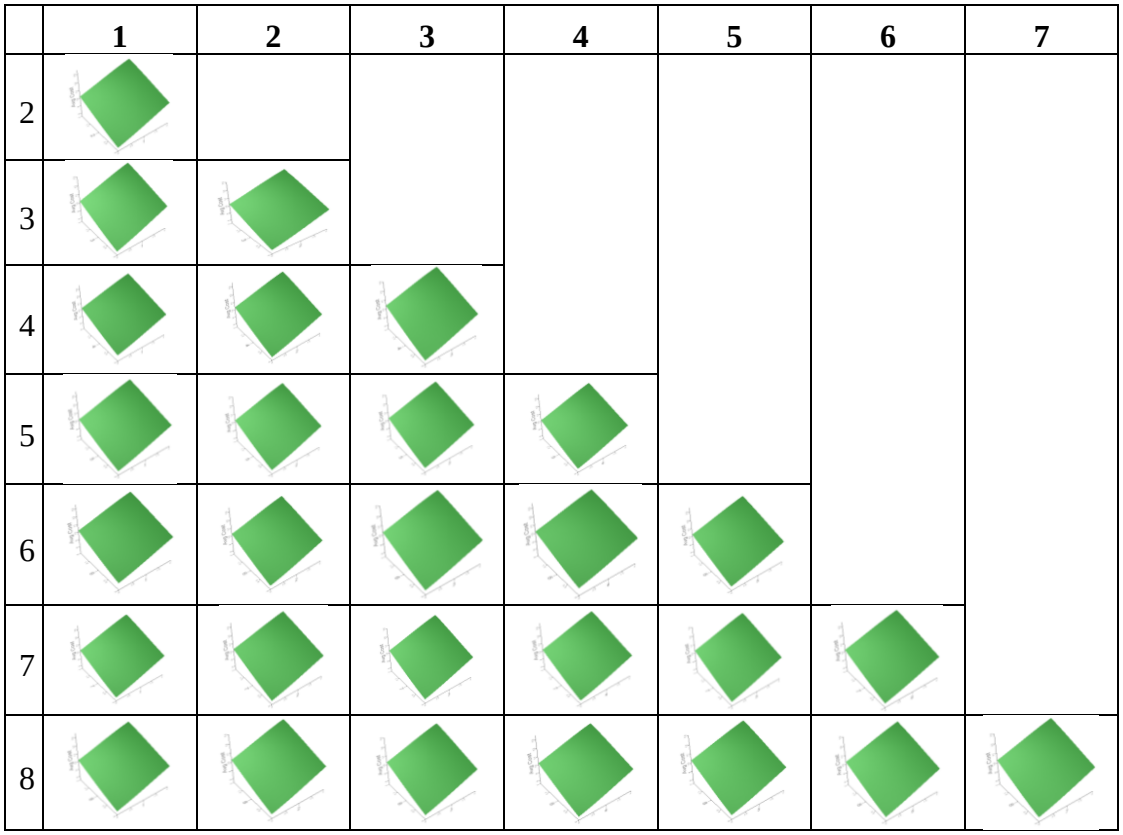


Figure 6.50 Phase III Average Total Cost NOLHD RSM

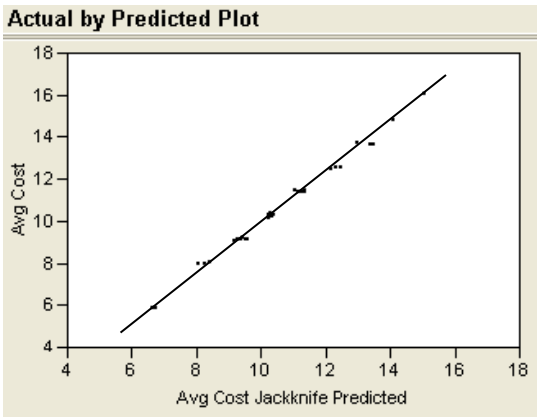


Figure 6.51 Phase III Average Total Cost Actual by Predicted Plot

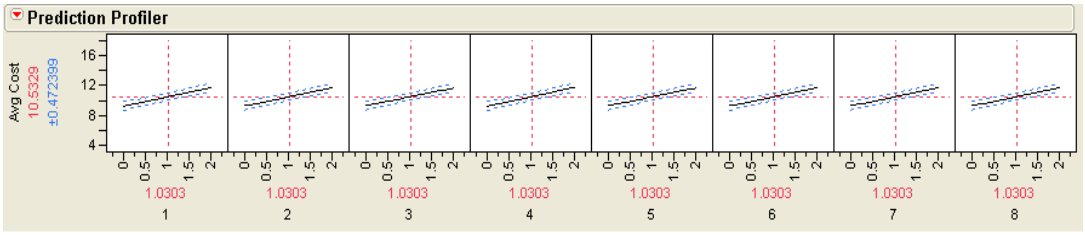


Figure 6.52 Phase III Average Total Cost Prediction Profiler

SA and multiple response optimization of the three response prediction formulas are shown in Table 6.32.

Table 6.32 Phase III Multi-Objective Optimization of Response Prediction Formulas

Optimal Patrol Policy by Obj		Avg Response Time Predicted Value	Avg % Time Occupied Predicted Value	Avg Total Cost Predicted Value
SA Results	1 1 2 2 1 1 0 1 (Min Avg Response Time)	176.79	0.90	11.34
	0 1 0 0 2 0 2 1 (Max Avg % Time Occupied)	208.31	0.96	8.28
	0 1 0 0 2 0 2 1 (Min Avg Total Cost)	208.31	0.96	8.28
Multiple Objective Optimization	0 1 2 0 2 1 2 1	173.51	0.90	11.05
	0 1 1 0 2 1 2 1	185.29	0.92	10.09
	0 1 2 0 2 0 2 1	184.68	0.92	10.07

The best policy is 0 1 2 0 2 0 2 1. Ten replications are conducted to collect the data to do the interim analysis on the current factor settings, Treatment #1, and the new factor settings, Treatment #2. Interim analysis indicates that more replications are needed for all three responses and the trial continues. The two treatment results follow in Table 6.33.

Table 6.33 Third Comparison of Treatment Results After 10 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 10 Reps Simulation Results	1 1 1 1 1 1 1 1	194.68 (109.52)	0.937 (0.017)	10.23 (0.061)
Treatment #2 10 Reps Simulation Results	0 1 2 0 2 0 2 1	190.61 (103.02	0.931 (0.014)	10.19 (0.056)

After ten additional replications, the interim analysis results again indicate that the trial should continue and no conclusions can be made for any of the three responses. The results of the interim analyses are summarized in Table 6.34.

Table 6.34 Third Comparison of Treatment Results After 20 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 20 Reps Simulation Results	1 1 1 1 1 1 1 1	191.74 (109.52)	0.931 (0.019)	10.21 (0.072)
Treatment #2 20 Reps Simulation Results	0 1 2 0 2 0 2 1	190.61 (103.02	0.931 (0.014)	10.19 (0.056)

After ten additional replications, the interim analysis results again indicate that the trial should be stopped for futility for the average total cost response. The trial should continue and no conclusions can be made for the average response time and the average % time occupied responses. The results of the interim analyses are summarized in Table 6.35.

Table 6.35 Third Comparison of Treatment Results After 30 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 30 Reps Simulation Results	1 1 1 1 1 1 1 1	192.31 (105.05)	0.932 (0.018)	10.16 (0.090)
Treatment #2 30 Reps Simulation Results	0 1 2 0 2 0 2 1	191.15 (101.56)	0.928 (0.16)	10.18 (0.071)

After an additional ten replications for a total of 40 replications the stopping boundary for futility is crossed for the average response time and the average % time occupied response. The results of the interim analyses are summarized in Table 6.36.

Table 6.36 Third Comparison of Treatment Results After 40 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)
Treatment #1 40 Reps Simulation Results	1 1 1 1 1 1 1 1	191.71 (109.92)	0.930 (0.012)
Treatment #2 40 Reps Simulation Results	0 1 2 0 2 0 2 1	191.71 (109.90)	0.930 (0.019)

Table 6.37 summarizes all the interim analysis results.

Table 6.37 Third Interim Analysis Results

Analysis	Response	Test Statistic	Lower Stopping Boundary	Upper Stopping Boundary	Decision
1 (10 reps)	Average Response Time	0.0856	-1.046	3.6321	Continue
	Average % Time Occupied	0.8615	-1.046	3.6321	Continue
	Average Total Cost	1.5275	-1.046	3.6321	Continue
2 (20 reps)	Average Response Time	0.0536	-0.4576	3.6084	Continue
	Average % Time Occupied	0.5542	-0.4576	3.6084	Continue
	Average Total Cost	0.9157	-0.4576	3.6084	Continue
3 (30 reps)	Average Response Time	0.0435	0.0114	3.6921	Continue
	Average % Time Occupied	0.9097	0.0114	3.6921	Continue
	Average Total Cost	-0.1213	0.0114	3.6921	Stop trial for futility
4 (40 reps)	Average Response Time	0.0000	0.3132	2.8605	Stop trial for futility
	Average % Time Occupied	0.0000	0.3132	2.8605	Stop trial for futility

The trial is stopped for futility as $\mathbf{Z_k} < \mathbf{l_k}(\alpha, ß)$. Figure 6.53 graphically depicts the interim analysis results which show the lower boundary being crossed by all three responses.

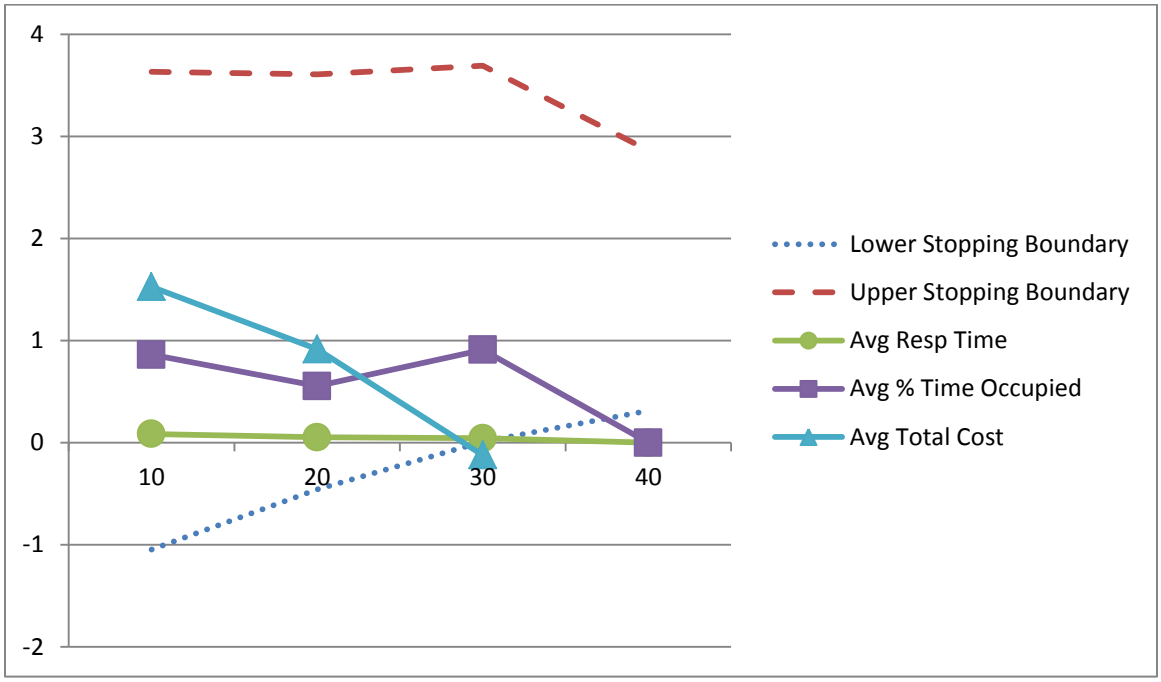


Figure 6.53 Efficacy and Futility Monitoring Boundaries

Thus far, no statistical difference has been found in any of the selected new policies that improve the responses over the current policy. Adding patrols improves the average response time and increases the average total cost. A policy with nine patrols is examined next to see if the improvement over response time is significant. The policy, previously found in Table 6.32, 0 1 2 0 2 1 2 1 patrol cars in each district is analyzed next. Interim analysis results after 10 replications found that the trial is stopped for efficacy for the average % time occupied and stopped for futility for the average total cost. The trial continues for the average response time response. Table 6.38 shows the results of the responses after ten replications.

Table 6.38 Nine Patrols Comparison of Treatment Results After 10 Replications

Results	Police Staffing Policy	Average Response Time (STD)	Average % Time Occupied (STD)	Average Total Cost (STD)
Treatment #1 10 Reps Simulation Results	1 1 1 1 1 1 1 1	194.68 (109.52)	0.937 (0.017)	10.23 (0.061)
Treatment #2 10 Reps Simulation Results	0 1 2 0 2 1 2 1	182.62 (96.96)	0.911 (0.014)	11.30 (0.053)

After 20 total replications no conclusions can be made about the average response time response and the trial continues (Table 6.39).

Table 6.39 Nine Patrols Comparison of Treatment Results After 20 Replications

Results	Police Staffing Policy	Average Response Time (STD)
Treatment #1 20 Reps Simulation Results	1 1 1 1 1 1 1 1	191.74 (109.52)
Treatment #2 20 Reps Simulation Results	0 1 2 0 2 1 2 1	184.39 (97.81)

After 30 total replications no conclusions can be made about the average response time response and the trial continues (Table 6.40).

Table 6.40 Nine Patrols Comparison of Treatment Results After 30 Replications

Results	Police Staffing Policy	Average Response Time (STD)
Treatment #1 30 Reps Simulation Results	1 1 1 1 1 1 1 1	192.31 (105.05)
Treatment #2 30 Reps Simulation Results	0 1 2 0 2 1 2 1	185.16 (99.86)

After 40 total replications the trial is stopped for futility as the lower stopping

boundary is crossed for the average response time response (Table 6.41).

Table 6.41 Nine Patrols Comparison of Treatment Results After 40 Replications

Results	Police Staffing Policy	Average Response Time (STD)
Treatment #1 40 Reps Simulation Results	1 1 1 1 1 1 1 1 1	191.71 (109.92)
Treatment #2 40 Reps Simulation Results	0 1 2 0 2 1 2 1	184.77 (99.86)

Table 6.42 summarizes the interim analysis results.

Table 6.42 Nine Patrols Interim Analysis Results

Analysis	Response	Test Statistic	Lower Stopping Boundary	Upper Stopping Boundary	Decision
1 (10 reps)	Average Response Time	0.2607	-1.046	3.6321	Continue
	Average % Time Occupied	3.7334	-1.046	3.6321	Stop trial for efficacy
	Average Total Cost	-41.87	-1.046	3.6321	Stop trial for futility
2 (20 reps)	Average Response Time	0.2239	-0.4576	3.6084	Continue
3 (30 reps)	Average Response Time	0.2702	0.0114	3.6921	Continue
4 (40 reps)	Average Response Time	0.2956	0.3132	2.8605	Stop trial for futility

Figure 6.54 graphically depicts the results of the interim analysis.

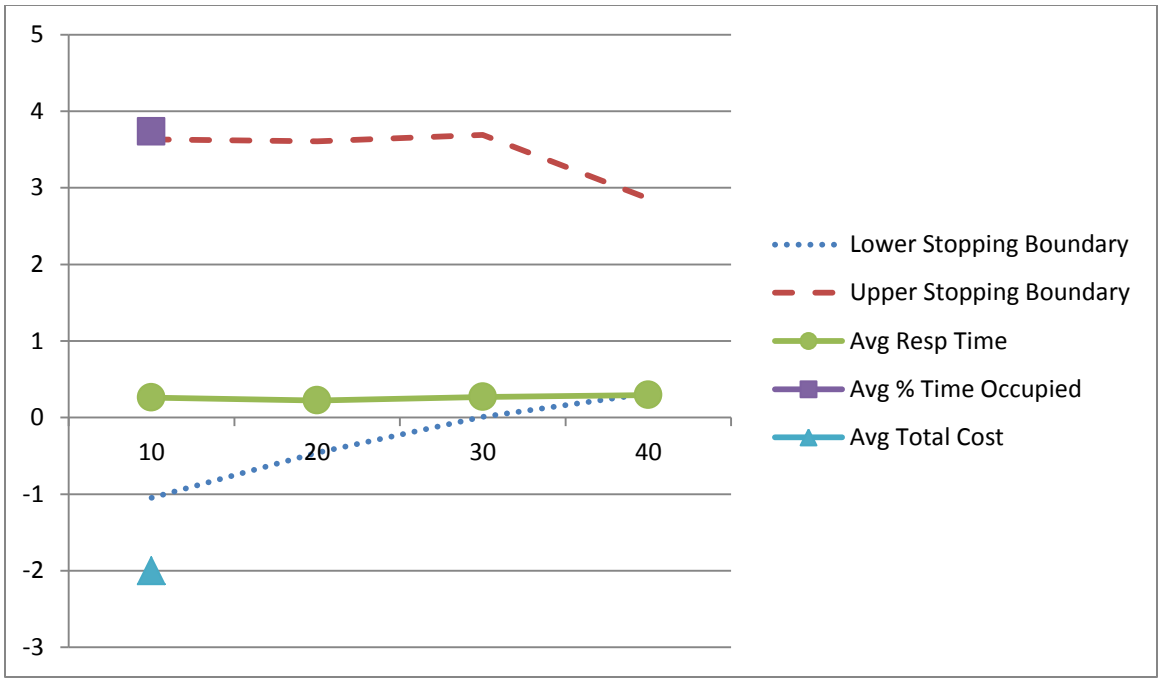


Figure 6.54 Efficacy and Futility Monitoring Boundaries

Further analysis indicates that in order for a policy to prove better than the current policy the response time would need to be around 115 time units. To achieve this result there would need to be a large number of patrols greater than ten and would more than likely prove to be too costly.

6.2.2 Results of the Police Staffing Simulation Study

The traditional 2^8 full factorial has 256 design points compared to 33 design points in the NOLHD. This is an 87% savings in the number of design points. The NOLHD EVOP scheme resulted in greater efficiency. The proposed methodology shows great potential in the application of efficient stochastic computer experiments with a large number of factors and multiple responses. The NOLHD ensures that the design is space-filling and therefore minimizes variance and maintains the assumption of normality.

The concept of EVOP is an ideal strategy for combating crime. As the patrols adjust to fight the crime, it is likely that the patterns in crime will change as well. By continually monitoring the crime, the patrols can be adjusted to change with the changing crime patterns. These adaptive patrols can deter crime as the patrols are unpredictable to the criminals. It is the intent that this ECEM process be continually applied as the crime adapts to the patrols and the patrols adapt to the new crime patterns. This is the original intent behind employing the EVOP process and therefore this problem is well suited for this methodology.

7 Conclusion

Section 7.1 summarizes the research that was presented here. Section 7.2 revisits the research questions and recaps how they were answered. Section 7.3 addresses the theoretical and applied research contributions.

7.1 Summary

This dissertation shows how a stochastic simulation, with many factors and multiple responses, can be solved in an efficient manner. The results of this research are an efficient computer experimentation methodology called the ECEM which leverages procedures used during the conduct of clinical trials such as the use of a monitoring plan and interim analysis, utilizes the EVOP concept while applying sophisticated experimental designs of OLHDs/NOLHDs, applies Gaussian process models to find the response prediction formulas and makes use of SA and multiple objective optimization on the response prediction formula in order to locate the local\global optimum.

Chapter 3 reviewed the literature in the following areas. First, experimental designs

and sampling techniques are reviewed and the current methods of gaining efficiency are addressed. Next, the EVOP concept is introduced and explained, optimization and how it can be employed during the experimental design process is shown and clinical trial analysis and procedures are reviewed.

Chapter 4 and 5 explains the ECEM developed for this research. Chapter 4 details the assumptions that go with each tool employed in the methodology while chapter 5 addresses the properties of the methodology which make it efficient. Chapter 6 applies the methodology to two areas. First the chemical mixing problem and second the police staffing simulation to demonstrate how this methodology can be applied to continuous time problems and solve them efficiently.

7.2 Research Questions Answered

The research questions posed are: 1) Will the statistical analysis tools employed during clinical trials and applied to stochastic simulation experiments improve the process? 2) How can complex, stochastic simulation computer experiments be conducted efficiently? 3) Can current EVOP and RSM be extended to apply more sophisticated DOE to gain efficiency?

This research has shown that developing a monitoring plan such as is done in the conduct of clinical trials, one can determine the accuracy desired and then plan the number of replications required to reach this accuracy. Given a budget, the experiment can also stay within this budget by planning the number of replications that are feasible given this budget and then reporting the results to a certain precision as a result of this constraint. This helps quantify the uncertainty given the stochastic nature of the experiment. Additionally, applying interim analysis during this sequential process allows

for the monitoring of the confidence interval and potentially terminating the study once the precision requirements are met.

When the simulation is complex and stochastic, OLHDs/NOLHDs greatly improve the efficiency without sacrificing results. These space-filling designs give greater fidelity in the response surface and therefore a better response prediction formula. The use of analyses typically found during the conduct of clinical trial also ensures that only necessary replications are completed to facilitate the analysis.

With today's computing power, advanced EVOP schemes can be utilized thus enabling efficiency through the use of these sophisticated designs. Where the practitioner relied on a statistician to manually compute the statistics of the factor analysis, this is now easily accomplished with today's computers and software.

RSM can be extended to include the use of a local\global optimizer such as SA and multiple response optimization in order to locate the local\global optimizer. Optimizing the response prediction formula using SA and multiple objective optimization may lead to finding optimal factor settings without actually conducting additional simulation runs. The application of Gaussian process modeling to a stochastic simulation is a viable modeling method for modeling the outputs of the simulation. This research demonstrated that these models are well suited for use given a stochastic computer simulation.

7.3 Contributions

This dissertation examined the design and analysis of large, complex stochastic computer simulations with multiple responses to develop a methodology to gain efficiency. It was demonstrated that efficiencies can be gained through the use of conducting experimental design and RSM with OLHD/NOLHD and Gaussian process

modeling while optimizing with the use of SA and multiple objective optimization. Additionally, statistical analysis applying futility and efficacy monitoring, interim analysis and application of error spending functions and stopping criteria found in clinical trial analysis was applied to large, complex, stochastic simulation model.

7.3.1 Theoretical Contributions

The ECEM shows several areas where efficiencies are gained. The theoretical contributions and hence the areas where efficiencies were gained are 1) the use of clinical trial procedures, 2) use of OLHDs and NOLHDs 3) Extension of RSM using optimization techniques on the response prediction equations and 4) use of sequential analysis and the EVOP process.

Clinical trial procedures employed are the development of a monitoring plan; the use of power analysis; stopping boundaries and error spending functions typically found in clinical trial; and the use of interim analysis to monitor these stopping boundaries. Beginning with a monitoring plan allows the experimenter to define the required precision and therefore determine how many replications are necessary to meet this precision, thus preserving the use of limited resources such as time and money. The use of clinical trial error spending functions and stopping boundaries helps accomplish the objectives laid out in the monitoring plan by preserves type I and II error while only using necessary replications. Additionally the experimenter can choose how the error is "spent" throughout the process thus maintaining the integrity of the experiment. These procedures help maintain the minimum sample size.

The use of OLHDs/NOLHDs helps maintain our assumptions of normality and minimum variance through their use of LHS. LHS has been shown to produce

distributions that are normal and have a smaller variance over SRS schemes as a result of the employment of stratification. This improves efficiency and helps minimize the number of samples needed.

The selection of the experimental design as OLHDs/NOLHDs ensures good coverage of the design space thus leading to the inclusion of the global optimal within the experimental design. This is important as the methodology relies on the use of Gaussian process modeling to produce the response prediction formulas in which the global optimum is sought. SA and multiple objective optimization are applied to these response prediction formulas that result in the selection of the treatment that will compete with the prior treatment to determine if one treatment results in better performance over the other. This process reduces the number of experimental runs as the prediction model is used rather than running additional experiments. The use of the response surfaces helps to define the starting location and focuses the global search thus also reducing the number of experimental runs.

The use of sequential analysis and EVOP helps gain much needed efficiency given the stochastic nature of the simulations and helps minimize sample size. Using EVOP, a sequential process, while employing today's sophisticated, efficient designs, is an improvement over traditional methods.

7.3.2 Applied Contributions

This research has applied contributions in many areas where stochastic simulations are utilized and events evolve over time. This research specifically demonstrates applied contributions to the areas of industrial engineering, the chemical mixing problem, and the assignment problem, such as the police staffing simulation.

This research applies the ECEM to a chemical process simulation and extends the use of EVOP to include the use of modern experimental designs while defining a procedure for ensuring that the experiment is efficient by making use of procedures found in the conduct of clinical trials.

Applying the ECEM to the police staff study demonstrated that patrols can be assigned to optimize output parameters and can help better understand the distribution of crime across a given area. This research has applications in related fields where in general a resource is assigned to an area to perform a function or service and it is desired that the resource is allocated as efficiently as possible. While the application in this research was to assign police patrols, the simulation and methodology used to evaluate the simulation could easily extend to the assignment of military patrols in a region, ambulance and fire-fighting services assignment, establishing hunting zones in wildlife preserves, meals on wheels and pizza delivery.

8 Future Work

This work concentrated on comparing the results from applying OLHD's\NOLHD's to full factorials. Since there were no interior points in the full factorial design, the response prediction formulas found applying the full factorials were significantly worse than the OLHD's\NOLHD's. The next step would be to see how the OLHD's\NOLHD's compare to designs with center points or interior points such as central composite and Box-Behnkin designs.

The research presented here touches on the use of clinical trial stopping boundaries and error spending functions. There are many different boundaries and error spending functions to explore and to determine how best to use each of these given different

problems. Other areas for future work includes the application of a Bayesian approach to futility and efficacy monitoring and predictive power analysis; application of adaptive trial designs; and examination of the error rates with respect to interim monitoring and analysis.

Having demonstrated this methodology on the chemical mixing problem and employment of police patrols, future work should apply this methodology to the other application areas discussed in Chapter 7 as well as in other disciplines to show that the ECEM is not only efficient but also versatile across a wide range of problems.

9 References

Asmussen, S. and Glynn, P.W. (2007), Stochastic Simulation: Algorithms and Analysis. Springer, New York.

Barton, R.R. and Ivey J.S. (1996), Nelder-Mead simplex modifications for simulation optimization. Management Science, 42:954-973.

Box, G. and Draper, N. (1969), Evolutionary Operations: A Statistical Method for Process Improvement. John Wiley & Sons, Inc., New York.

Box, G. and Draper, N. (1987) Empirical Model-Building and Response Surfaces. John Wiley and Sons, New York.

Burington, B. and Emerson, S. (2004), Constrained boundary monitoring for group sequential clinical trials. University of Washington Biostatistics Working Paper Series, Paper 199.

Butler N.A, "Optimal and Orthogonal Latin Hypercube Designs for Computer Experiments", Biometrika, Volume 88, Number 3, 1 October 2001, pp. 847-857.

Caris, A. and Janssens, G. (2005) "Selection of Control Variates for Variance Reduction in a Multiresponse Simulation with a Small Number of Replications". Proceedings of the European Multidisciplinary Society for Modelling and Simulation Technology, University of Porto, Porto, Portugal.

Chantarawong,T., Rungrattanaubol, J., and Na-udom, A., "Enhancement of enhanced stochastic evolutionary algorithm for computer simulated experiment", Research Paper, Naresuan University, http://suanpalm3.kmutnb.ac.th/journal/pdf/vol12/ch11.pdf, accessed March 2012.

Cioppa, T. (2002), Efficient Nearly Orthogonal and Space-Filling Experimental Designs for High-Dimensional Complex Models. Dissertation, Naval Post Graduate School, Monterey, CA.

Cohen, J. (1988) Statistical Power Analysis for the Behavioral Sciences. L. Erlbaum Associates, Hillsadale, N.J.

Cook, T. and DeMets, D. (2007), Introduction to Statistical Methods for Clinical Trials. Boca Raton: Chapman and Hall\CRC.

Crino, S. (2006), Combining Multivariate Adaptive Regression Splines with a Response Surface Methodology for Simulation-Based Design Optimization. Dissertation,University of Virginia, Charlottesville, VA.

Dattalo, P. (2008) Determining Sampling Size: Balancing Power, Precision and Practicality. Oxford University Press, Oxford.

Devore, J., (2008) Probability and Statistics for Engineering and the Sciences. Thomson, Brooks Cole.

Dmitrienko, A., Molenberghs, G., Chuang-Stein, C. and Offen, W. (2007), Analysis of Clinical Trials using SAS: A Practical Guide, Second Edition. SAS Institute Inc., Cary,NC.

Everitt, B. (2004), Statistical Aspects of the Design and Analysis of Clinical Trials, Revised Edition. Singapore: Imperial College Press.

Fang, K., Li, R. and Sudjianto, A. (2006), Design and Modeling for Computer Experiments. Chapman and Hall/CRC, Boca Raton.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Friedman, J, (1991), Multivariate adaptive regression splines. The Annals of Statistics, 19(1):1-141.

Gehan, E. and Lemak, N. (1994), Statistics in Medical Research. Plenum Publishing Company, New York.

Ghani, S. (1995), Performance of global optimization algorithm EVOP for non-linear non-differentiable constrained objective functions. Proceedings of the IEEE International Conference on Evolutionary Computing, The University of Western Australia, Perth, Western Australia, pages 320 - 325.

Ghosh, B.K. (1970) Sequential Tests of Statistical Hypotheses, Reading, Addison-Wesley Publishing Company.

Giunta, A., Wojtkiewicz, S., and Eldred, M. (2003). Overview of modern design of methods for computational simulations. 41st American Institute of Aeronautics and Astronautics Aerospace Sciences Meeting and Exhibit.

Graham, M. (2004), Improving critical drug development decision making through clinical trial simulation. Institute for International Research IIR (BV), PTI International.

Hagino, A., Hamada, C., Yoshimura, I., Ohashi, Y., Sakamoto, J. and Nakazato, H. (2004), Statistical comparison of random allocation methods in cancer clinical trials. Controlled Clinical Trials, 25(6):572–584.

Hill, R. (1998), OPER 683 Response Surface Methodology, Chemical Mixture Problem, AFIT, Wright Patterson Air Force Base, Ohio.

Holmes, D. and Mergen, A.E. (2006), Identifying need for new factors in EVOP. Quality Engineering, 18(2):219-224.

Hoshmand, Reza. (2006), Design of Experiments for Agriculture and the Natural Sciences, Chapman and Hall, CRC Press, Bacon Raton, FL.

Humphrey, D.G. and Wilson, J.R. (2000), A revised simplex search procedure for stochastic simulation response surface optimization. INFORMS J. Comput, 12:272-283.

Husslage, Bruce, Gijs Rennen, Edwin R. vanDam, "Space-Filling Latin Hypercube Designs for Computer Experiments", Discussion Paper No. 2006-18, Tillberg University, Mar 2006.

Jennison, C. and Turnbull, B. (2006), Adaptive and non-adaptive group sequential tests. Biometrika, 93(1):1-21.

Jin, R, Chen , W. and Sudjanto, A. (2005), "An efficient algorithm for construct optimal design of computer experiment," *Journal of statistical Planning and Inference*, Vol. 134, pp. 268-287.

JMP, Version 7. SAS Institute Inc., Cary, NC, 1989-2007.

Jones, B. and Johnson, R. T. (2009), Design and analysis for the Gaussian process model. Qual. Reliab. Eng. Int., 25: 515–524. doi: 10.1002/qre.1044.

Karrison, T., Huo, D. and Chappell, R. (2003), A group sequential, response-adaptive design for randomized clinical trials. Controlled Clinical Trials, 24:506–522.

Kleijnen, J.P.C., Design of Experiments, SciTopics, available at http://www.scitopics.com/Design_of_experiments.html, 28 August 2008.

Kocijan, J., Murray-Smith, R., Rasmussen, C.E. and Girard, A., Gaussian Process Model Based Predictive Control, Proceedings of the 2004 American Control Conference, Boston, Massachusetts, June 30, 2004 – July 2, 2004.

Lai, T.L., "Sequential Analysis: Some Classical Problems and New Challenges", Statistica Sinica 11, 303 – 403, 2001.

Law, A. and Kelton, W.D. (2000), Simulation Modeling and Analysis, Third Edition. McGraw-Hill Companies, Inc. Boston.

Lee, Eva K. "Optimization with Multiple Objectives". NCI-NSF Workshop on Operations Research in Radiation Therapy, Washington, D.C., February 7 – 9, 2002.

Malleson, N. RepastCity – A Demo Virtual City, http://portal.ncess.ac.uk/access/wiki/site/mass/repastcity.html.

Mehta, Cyrus R. Design and Interim Monitoring of Flexible Clinical Trials Featuring the East Software, 15th International DIA Workshop, Dublin, April 19, 2004.

Michalewicz, Z. and D.B. Fogel, How to Solve It: Modern Heuristics, Springer-Verlag, Berlin, 2000. (Chapters 1 – 7)

Molga, M. and Smutnicki, C. (2005) "Test functions for optimization needs", http://www.zsd.ict.pwr.wroc.pl/files/docs/functions.pdf, accessed 19 March 2012.

Murphy, K., Myors, B. and Wolach, A. (2009), Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Tests, Third Edition. Taylor and Francis Group, New York.

Myers, R.H. and Montgomery, D.C. (2002), Response Surface Methodology: Process and Product Optimization Using Designed Experiments, Second Edition. John Wiley & Sons, Inc., New York.

Mukhopadhyay, N. and de Silva, B. (2009) Sequential Methods and Their Application. Chapman and Hall, Boca Raton.

Owen, A.B. "A central limit theorem for Latin hypercube sampling". Journal of the Royal Statistical Society, B54, pp 541-551, 1992.

Ripley, B.D. (1987), Stochastic Simulation. John Wiley & Sons, Inc., New York.

Saab, Y. G. and Rao, Y. B., "Combinatorial optimization by stochastic evolution," *IEEE Transaction on Computeraided Design*, Vol.10, pp. 525–535, 1991.

Samoilenko, S. and Osei-Bryson, K. "Determining sources of relative inefficiency in heterogeneous samples: Methodology using Cluster Analysis, DEA and Neural Networks", European Journal of Operational Research, Vol. 206, pp 479 – 487, 2010.

Sanchez, S. M. 2005. NOLH designs spreadsheet. Available online via http://diana.cs.nps.navy.mil/SeedLab/ [accessed 08/16/2007]

Santner, T.J., Williams, B.J. and Notz, W.I. (2003), The Design and Analysis of Computer Experiments. Springer-Verlag, New York.

SAS Institute Inc. 2007. *JMP Design of Experiments*. Cary, NC: SAS Institute Inc.

Schamburg, J. B. (2004), A Modified Response Surface Methodology for Knowledge Discovery. Dissertation, University of Virginia, Charlottesville, VA.

Shang, T., Wan, Y., Fromherz, M. and Crawford, L. (2001), Toward adaptive cooperation between global and local solvers for continuous constraint problems.Workshop on Cooperative Solvers in Constraint Programming, CP'01.

Simpson, T. W., Mauery, T. M., Korte, J. J., and Mistree, F. (1998), Comparison of response surface and kriging models for multidisciplinary design optimization. Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis & Optimization, St. Louis, MO, AIAA Paper 98-4755.

Spendley, W., G.R. Hext and F.R. Himsworth, "Sequential Application of Simplex Designs in Optimisation and Evolutionary Operations", *Technometrics*, 4, pp. 441-461, (1962).

Stein, M.L., (1987) "Large Sample Properties of Simulations using Latin hypercube Sampling". *Technometrics* 29, 143-151.

Steinberg, David M. and Dennis K. J. Lin, "A Construction Method For Orthogonal Latin Hypercube Designs" Biometrika, Volume 93, Number 2, June 2006, pp. 279-288(10).

Stern, H.S. (1996), "Neural Networks in applied statistics", *Technometrics*, 38 (3), pp. 205-214.

Tang, B. (1993). "Orthogonal Array-Based Latin Hypercubes". *Journal of the American Statistical Association*, 88, 1392–1397.

Wald, A. (1947), Sequential Analysis. John Wiley and Sons, Inc. London.

Wang, G. (2003), Adaptive response surface method using inherited Latin hypercube design points. Journal of Mechanical Design, 125:210-220.

Whitehead, J. (1999), A unified theory for sequential clinical trials. Statistical Medicine, 18:2271-2286 [Medline].

Ye, K.Q. (1998), Orthogonal column Latin hypercubes and their application in computer experiments. Journal of the American Statistical Association, 93:1430–1439.

Appendix A: Chemical Mixing Experiment Response Prediction Formul

Equation A.1 Response Prediction Formula for Yield given Treatment #1

Yield=(-54.97)+(-2367.07*Exp(-(0.00011*(Temp-278)^2+0.0000040*(Conc-500)^2))+1066.85*Exp(-(0.00011*(Temp-216)^2+0.0000040*(Conc-200)^2))+-2587.57*Exp(-(0.00011*(Temp-231)^2+0.0000040*(Conc-275)^2))+3880.11*Exp(-(0.00011*(Temp-247)^2+0.0000040*(Conc-350)^2))+-448.45*Exp(-(0.00011*(Temp-388)^2+0.0000040*(Conc-475)^2))+-554.20*Exp(-(0.00011*(Temp-450)^2+0.0000040*(Conc-225)^2))+1039.14*Exp(-(0.00011*(Temp-356)^2+0.0000040*(Conc-175)^2))+1296.57*Exp(-(0.00011*(Temp-341)^2+0.0000040*(Conc-450)^2))+-3383.61*Exp(-(0.00011*(Temp-325)^2+0.0000040*(Conc-300)^2))+-1042.76*Exp(-(0.00011*(Temp-372)^2+0.0000040*(Conc-100)^2))+-408.65*Exp(-(0.00011*(Temp-434)^2+0.0000040*(Conc-400)^2))+1238.66*Exp(-(0.00011*(Temp-419)^2+0.0000040*(Conc-325)^2))+594.92*Exp(-(0.00011*(Temp-403)^2+0.0000040*(Conc-250)^2))+-1177.73*Exp(-(0.00011*(Temp-263)^2+0.0000040*(Conc-125)^2))+-563.03*Exp(-(0.00011*(Temp-200)^2+0.0000040*(Conc-375)^2))+1877.32*Exp(-(0.00011*(Temp-294)^2+0.0000040*(Conc-425)^2))+1539.47*Exp(-(0.00011*(Temp-309)^2+0.0000040*(Conc-150)^2)))

Equation A.2 Response Prediction Formula for Viscosity given Treatment #1

Viscosity=92.82+(1039.82033640314*Exp(-(0.00012*(Temp-278)^2+0.000011*(FeedRate-525)^2+0.0027*(Time-2)^2))+-1028.19*Exp(-(0.00012*(Temp-216)^2+0.000011*(FeedRate-550)^2+0.0027*(Time-1)^2))+-1240.40*Exp(-(0.00012*(Temp-231)^2+0.000011*(FeedRate-225)^2+0.0027*(Time-4)^2))+2089.03*Exp(-(0.00012*(Temp-247)^2+0.000011*(FeedRate-325)^2+0.0027*(Time-4)^2))+512.99*Exp(-(0.00012*(Temp-388)^2+0.000011*(FeedRate-375)^2+0.0027*(Time-3)^2))+0.57*Exp(-(0.00012*(Temp-450)^2+0.000011*(FeedRate-350)^2+0.0027*(Time-1)^2))+30.21*Exp(-(0.00012*(Temp-356)^2+0.000011*(FeedRate-600)^2+0.0027*(Time-5)^2))+-1585.17*Exp(-(0.00012*(Temp-341)^2+0.000011*(FeedRate-500)^2+0.0027*(Time-5)^2))+3255.16*Exp(-(0.00012*(Temp-325)^2+0.000011*(FeedRate-400)^2+0.0027*(Time-

4)^2))+161.02*Exp(-(0.00012*(Temp-372)^2+0.000011*(FeedRate-275)^2+0.0027*(Time-5)^2))+-666.46*Exp(-(0.00012*(Temp-434)^2+0.000011*(FeedRate-250)^2+0.0027*(Time-6)^2))+-899.27*Exp(-(0.00012*(Temp-419)^2+0.000011*(FeedRate-575)^2+0.0027*(Time-3)^2))+1229.26*Exp(-(0.00012*(Temp-403)^2+0.000011*(FeedRate-475)^2+0.0026*(Time-3)^2))+-2424.05*Exp(-(0.00012*(Temp-263)^2+0.000011*(FeedRate-425)^2+0.0027*(Time-4)^2))+1419.53*Exp(-(0.00012*(Temp-200)^2+0.000011*(FeedRate-450)^2+0.0027*(Time-6)^2))+1136.33*Exp(-(0.00012*(Temp-294)^2+0.000011*(FeedRate-200)^2+0.0027*(Time-2)^2))+-3030.35*Exp(-(0.00012*(Temp-309)^2+0.000011*(FeedRate-300)^2+0.0027*(Time-2)^2)))

Equation A.3 Response Prediction Formula for Molecular Weight given Treatment #1

Molecularweight=342.05+(-2612.89*Exp(-(0.000000026*(Temp-278)^2+0.000033*(Conc-500)^2+0.00000034*(FeedRate-525)^2+0.0000082*(Pressure-163)^2))+-1246.907*Exp(-(0.000000026*(Temp-216)^2+0.000034*(Conc-200)^2+0.00000034*(FeedRate-550)^2+0.0000082*(Pressure-174)^2))+4445.48*Exp(-(0.000000026*(Temp-231)^2+0.000034*(Conc-275)^2+0.00000034*(FeedRate-225)^2+0.0000082*(Pressure-155)^2))+4867.78*Exp(-(0.000000026*(Temp-247)^2+0.000034*(Conc-350)^2+0.00000034*(FeedRate-325)^2+0.0000082*(Pressure-200)^2))+4488.40*Exp(-(0.000000026*(Temp-388)^2+0.000034*(Conc-475)^2+0.00000034*(FeedRate-375)^2+0.0000082*(Pressure-148)^2))+-376.034*Exp(-(0.000000026*(Temp-450)^2+0.000034*(Conc-225)^2+0.00000034*(FeedRate-350)^2+0.0000082*(Pressure-189)^2))+-369.41*Exp(-(0.000000026*(Temp-356)^2+0.000034*(Conc-175)^2+0.00000034*(FeedRate-600)^2+0.0000082*(Pressure-159)^2))+-570.048*Exp(-(0.000000026*(Temp-341)^2+0.000034*(Conc-450)^2+0.00000034*(FeedRate-500)^2+0.0000082*(Pressure-196)^2))+-9906.25*Exp(-(0.000000026*(Temp-325)^2+0.000034*(:Conc-300)^2+0.00000034*(FeedRate-400)^2+0.0000082*(Pressure-170)^2))+-2632.26*Exp(-(0.000000026*(Temp-372)^2+0.000034*(Conc-100)^2+0.00000034*(FeedRate-275)^2+0.0000082*(Pressure-178)^2))+-5651.40*Exp(-(0.000000026*(Temp-434)^2+0.000034*(Conc-400)^2+0.00000034*(FeedRate-250)^2+0.0000082*(Pressure-166)^2))+2634.30*Exp(-(0.000000026*(Temp-419)^2+0.000034*(Conc-325)^2+0.00000034*(FeedRate-575)^2+0.0000082*(Pressure-185)^2))+1889.66*Exp(-(0.000000026*(Temp-

403)^2+0.000034*(Conc-250)^2+0.00000034*(FeedRate-475)^2+0.0000082*(Pressure-140)^2))+3479.56*Exp(-(0.000000026*(:Temp-263)^2+0.000034*(Conc-125)^2+0.00000034*(FeedRate-425)^2+0.0000082*(Pressure-193)^2))+1138.21*Exp(-(0.000000026*(Temp-200)^2+0.000034*(Conc-375)^2+ 0.00000034*(FeedRate-450)^2+0.0000082*(Pressure-151)^2))+797.31*Exp(-(0.000000026*(Temp-294)^2+0.000034*(Conc-425)^2+0.00000034*(FeedRate-200)^2+0.0000082*(Pressure-181)^2))+-375.50*Exp(-(0.000000026*(Temp-309)^2+0.000034*(Conc-150)^2+0.00000034*(FeedRate-300)^2+0.0000082*(Pressure-144)^2)))

Equation A.4 Response Prediction Formula for Yield given Treatment #2

Yield=231.019+(101.53*Exp(-(0.0024*(Temp-398)^2+0.000015*(Conc-500)^2+0.000033*(Pressure-163)^2))+85.82*Exp(-(0.0024*(Temp-380)^2+0.000015*(Conc-319)^2+0.000033*(Pressure-174)^2))+-98.11*Exp(-(0.0024*(Temp-384)^2+0.000015*(Conc-364)^2+0.000033*(Pressure-155)^2))+192.70*Exp(-(0.0024*(Temp-389)^2+0.000015*(Conc-410)^2+0.000033*(Pressure-200)^2))+-103.19*Exp(-(0.0024*(Temp-431)^2+0.000015*(Conc-485)^2+0.000033*(Pressure-148)^2))+-129.77*Exp(-(0.0024*(Temp-450)^2+0.000015*(Conc-334)^2+0.000033*(Pressure-189)^2))+-419.55*Exp(-(0.0024*(Temp-422)^2+0.000015*(Conc-304)^2+0.000033*(Pressure-159)^2))+-290.63*Exp(-(0.0024*(Temp-417)^2+0.000015*(Conc-470)^2+0.000033*(Pressure-196)^2))+-10.73*Exp(-(0.0024*(Temp-413)^2+0.000015*(Conc-380)^2+0.000033*(Pressure-170)^2))+101.95*Exp(-(0.0024*(Temp-427)^2+0.000015*(Conc-259)^2+0.000033*(Pressure-178)^2))+359.53*Exp(-(0.0024*(Temp-418)^2+0.000015*(Conc-440)^2+0.000033*(Pressure-166)^2))+275.11*Exp(-(0.0024*(Temp-419)^2+0.000015*(Conc-395)^2+0.000033*(Pressure-185)^2))+163.60*Exp(-(0.0024*(Temp-436)^2+0.000015*(Conc-349)^2+0.000033*(Pressure-140)^2))+-100.94*Exp(-(0.0024*(Temp-394)^2+0.000015*(Conc-274)^2+0.000033*(Pressure-193)^2))+-81.90*Exp(-(0.0024*(Temp-375)^2+0.000015*(Conc-425)^2+0.000033*(Pressure-151)^2))+-194.35*Exp(-(0.0024*(Temp-403)^2+0.000015*(Conc-455)^2+0.000033*(Pressure-181)^2))+148.93*Exp(-(0.0024*(Temp-408)^2+0.000015*(Conc-289)^2+0.000033*(Pressure-144)^2)))

Equation A.5 Response Prediction Formula for Viscosity given Treatment #2

(-922.44)+(115588.30*Exp(-(0.000040*(Temp-398)^2+0.0000034*(FeedRate-525)^2+0.000047*(Time-2)^2))+-22988.19*Exp(-(0.000040*(Temp-380)^2+0.0000034*(FeedRate-550)^2+0.000047*(Time-1)^2))+-986.79*Exp(-(0.000040*(Temp-384)^2+0.0000034*(FeedRate-225)^2+0.000047*(Time-4)^2))+52185.83*Exp(-(0.000040*(Temp-389)^2+0.0000034*(FeedRate-325)^2+0.000047*(Time-4)^2))+-38384.33*Exp(-(0.000040*(Temp-431)^2+0.0000034*(FeedRate-375)^2+0.000047*(Time-3)^2))+4165.95*Exp(-(0.000040*(Temp-450)^2+0.0000034*(FeedRate-350)^2+0.000047*(Time-1)^2))+1544.97*Exp(-(0.000040*(Temp-422)^2+0.0000034*(FeedRate-600)^2+0.000047*(Time-5)^2))+-2887.89*Exp(-(0.000040*(Temp-417)^2+0.0000034*(FeedRate-500)^2+0.000047*(Time-5)^2))+-25098.64*Exp(-(0.000040*(Temp-413)^2+0.0000034*(FeedRate-400)^2+0.000047*(Time-4)^2))+-12152.59*Exp(-(0.000040*(Temp-427)^2+0.0000034*(FeedRate-275)^2+0.000047*(Time-5)^2))+42727.88*Exp(-(0.000040*(Temp-418)^2+0.0000034*(FeedRate-250)^2+0.000047*(Time-6)^2))+-53182.25*Exp(-(0.000040*(Temp-419)^2+0.0000034*(FeedRate-575)^2+0.000047*(Time-3)^2))+37425.37*Exp(-(0.000040*(Temp-436)^2+0.0000034*(FeedRate-475)^2+0.000047*(Time-3)^2))+-60380.42*Exp(-(0.000040*(Temp-394)^2+0.0000034*(FeedRate-425)^2+0.000047*(Time-4)^2))+-17342.29*Exp(-(0.000040*(Temp-375)^2+0.0000034*(FeedRate-450)^2+0.000047*(Time-6)^2))+-36979.32*Exp(-(0.000040*(Temp-403)^2+0.0000034*(FeedRate-200)^2+0.000047*(Time-2)^2))+16744.41*Exp(-(0.000040*(Temp-408)^2+0.0000034*(FeedRate-300)^2+0.000047*(Time-2)^2)))

Equation A.6 Response Prediction Formula for Molecular Weight given Treatment #2

467.24+(-6.37*Exp(-(0.00013*(Concentration-500)^2+0.000010*(FeedRate-525)^2+0.0614495575950812*(Time-2)^2))+-10.76*Exp(-(0.00013*(Concentration-319)^2+0.000010*(FeedRate-550)^2+0.06*(Time-1)^2))+-21.45*Exp(-(0.00013*(Concentration-364)^2+0.000010*(FeedRate-225)^2+0.06*(Time-4)^2))+34.83*Exp(-(0.00013*(Concentration-410)^2+0.000010*(FeedRate-325)^2+0.061*(Time-4)^2))+23.68*Exp(-(0.00012843400127843 7*(Concentration-485)^2+0.000010*(FeedRate-375)^2+0.06*(Time-3)^2))+-24.25*Exp(-

(0.00013*(Concentration-334)^2+0.000010*(FeedRate-350)^2+0.06*(Time-1)^2))+28.14*Exp(-(0.00013*(Concentration-304)^2+0.000010*(FeedRate-600)^2+0.06*(Time-5)^2))+-16.02*Exp(-(0.00013*(Concentration-470)^2+0.000010*(FeedRate-500)^2+0.06*(Time-5)^2))+-7.62*Exp(-(0.00013*(Concentration-380)^2+0.000010*(FeedRate-400)^2+0.06*(Time-4)^2))+3.07*Exp(-(0.00013*(Concentration-259)^2+0.000010*(FeedRate-275)^2+0.06*(Time-5)^2))+43.99*Exp(-(0.00013*(Concentration-440)^2+0.000010*(FeedRate-250)^2+0.06*(Time-6)^2))+9.04*Exp(-(0.00013*(Concentration-395)^2+0.000010*(FeedRate-575)^2+0.06*(Time-3)^2))+63.37*Exp(-(0.00013*(Concentration-349)^2+0.000010*(FeedRate-475)^2+0.06*(Time-3)^2))+-98.96*Exp(-(0.00013*(Concentration-274)^2+0.000010*(FeedRate-425)^2+0.06*(Time-4)^2))+-39.25*Exp(-(0.00013*(Concentration-425)^2+0.000010*(FeedRate-450)^2+0.06*(Time-6)^2))+-27.06*Exp(-(0.00013*(Concentration-455)^2+0.000010*(FeedRate-200)^2+0.06*(Time-2)^2))+45.61*Exp(-(0.00013*(Concentration-289)^2+0.000010*(FeedRate-300)^2+0.061*(Time-2)^2)))

Appendix B: Chemical Mixing Experiment Gaussian Process Model Reports

Table B.1 Model Report Phase I Gaussian Process Model – Yield

Model Report

Column	Theta	Total Sensitivity	Main Effect	Temp Interaction	Concentration Interaction	Feed Rate Interaction	Pressure Interaction	Time Interaction
Temp	0.0001121	0.856917	0.8493205	.	0.0075965	0	0	0
Concentration	4.0046e-6	0.1506795	0.143083	0.0075965	.	0	0	0
Feed Rate	0	0	0	0	0	.	0	0
Pressure	0	0	0	0	0	0	.	0
Time	0	0	0	0	0	0	0	.

Mu	Sigma
-54.96873	21450.419

-2*LogLikelihood
168.49809

Table B.2 Model Report Phase I Gaussian Process Model – Viscosity

Model Report

Column	Theta	Total Sensitivity	Main Effect	Temp Interaction	Concentration Interaction	Feed Rate Interaction	Pressure Interaction	Time Interaction
Temp	0.0001203	0.9034598	0.8702451	.	0	0.0293453	0	0.0038694
Concentration	0	0	0	0	.	0	0	0
Feed Rate	1.125e-5	0.1238165	0.0943799	0.0293453	0	.	0	9.1355e-5
Pressure	0	0	0	0	0	0	.	0
Time	0.0026918	0.0057744	0.0018137	0.0038694	0	9.1355e-5	0	.

Mu	Sigma
92.816169	91331.394

-2*LogLikelihood
212.35094

Table B.3 Model Report Phase I Gaussian Process Model – Molecular Weight

Model Report

Column	Theta	Total Sensitivity	Main Effect	Temp Interaction	Concentration Interaction	Feed Rate Interaction	Pressure Interaction	Time Interaction
Temp	2.5571e-8	0.0006506	0.0006073	.	4.3011e-5	2.666e-7	7.9664e-9	0
Concentration	3.3561e-5	0.9903398	0.9848428	4.3011e-5	.	0.0035235	0.0019305	0
Feed Rate	3.3923e-7	0.0061869	0.0026629	2.666e-7	0.0035235	.	2.5252e-7	0
Pressure	8.2425e-6	0.0083194	0.0063886	7.9664e-9	0.0019305	2.5252e-7	.	0
Time	0	0	0	0	0	0	0	.

Mu	Sigma
342.05018	25109.086

-2*LogLikelihood
160.41615

Table B.4 Model Report Phase II Gaussian Process Model – Yield

Model Report

Column	Theta	Total Sensitivity	Main Effect	Temp Interaction	Concentration Interaction	Feed Rate Interaction	Pressure Interaction	Time Interaction
Temp	0.0024376	0.9716536	0.9468932	.	0.0149168	0	0.0098436	0
Concentration	1.4637e-5	0.0404558	0.0254999	0.0149168	.	0	3.9112e-5	0
Feed Rate	0	0	0	0	0	.	0	0
Pressure	0.0000335	0.0123668	0.0024841	0.0098436	3.9112e-5	0	.	0
Time	0	0	0	0	0	0	0	.

Mu	Sigma
231.0186	1647.6468

-2*LogLikelihood
146.04908

Table B.5 Model Report Phase II Gaussian Process Model – Viscosity

Model Report

Column	Theta	Total Sensitivity	Main Effect	Temp Interaction	Concentration Interaction	Feed Rate Interaction	Pressure Interaction	Time Interaction
Temp	3.9955e-5	0.6659999	0.6651242	.	0	0.0008717	0	4.0384e-6
Concentration	0	0	0	0	.	0	0	0
Feed Rate	3.3861e-6	0.3338907	0.3327592	0.0008717	0	.	0	0.0002597
Pressure	0	0	0	0	0	0	.	0
Time	4.7023e-5	0.0012448	0.000981	4.0384e-6	0	0.0002597	0	.

Mu	Sigma	Nugget
-922.4411	573708.44	0.0001

-2*LogLikelihood

174.56656

Nugget parameter set to avoid singular variance matrix.

Table B.6 Model Report Phase II Gaussian Process Model – Molecular Weight

Model Report

Column	Theta	Total Sensitivity	Main Effect	Temp Interaction	Concentration Interaction	Feed Rate Interaction	Pressure Interaction	Time Interaction
Temp	0	0	0	.	0	0	0	0
Concentration	0.0001284	0.9452716	0.8326057	0	.	0.0722919	0	0.040374
Feed Rate	0.0000101	0.0859931	0.0016662	0	0.0722919	.	0	0.012035
Pressure	0	0	0	0	0	0	.	0
Time	0.0614496	0.0699324	0.0175234	0	0.040374	0.012035	0	.

Mu	Sigma
467.236	378.95285

-2*LogLikelihood

133.85142

Appendix C: Police Staffing Study Response Prediction Formulas

Equation C.1 Phase I Response Prediction Formula for Average Response Time

-(167.31+(-3.15*exp(-(0.25*(dist(1)-4)^2+0.045*dist(2)^2+0.11*(dist(3)-
2)^2+0.03*(dist(4)-1)^2+0.23*(dist(5)-4)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-
2)^2))+135.33*exp(-(0.25*(dist(1)-4)^2+0.045*(dist(2)-4)^2+0.11*(dist(3)-
1)^2+0.03*(dist(4)-2)^2+0.23*(dist(5)-2)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-3)^2))+-
11.43*exp(-(0.25*(dist(1)-4)^2+0.045*(dist(2)-2)^2+0.11*(dist(3)-4)^2+0.031*(dist(4)-
1)^2+0.23*dist(5)^2+0.0075*(dist(7)-4)^2+0.32*(dist(8)-2)^2))+-12.66*exp(-
(0.25*(dist(1)-2)^2+0.04*(dist(2)-4)^2+0.116*(dist(3)-4)^2+0.03*(dist(4)-
2)^2+0.23*(dist(5)-4)^2+0.0075*dist(7)^2+0.32*(dist(8)-1)^2))+17.49*exp(-
(0.25*(dist(1)-4)^2+0.045*dist(2)^2+0.11*(dist(3)-2)^2+0.031*(dist(4)-
1)^2+0.23*(dist(5)-3)^2+0.0075*(dist(7)-1)^2+0.32*(dist(8)-3)^2))+-132.26*exp(-
(0.25*(dist(1)-4)^2+0.04*(dist(2)-4)^2+0.11*(dist(3)-1)^2+0.03*(dist(4)-
1)^2+0.23*(dist(5)-2)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-3)^2))+27.43*exp(-
(0.25*(dist(1)-3)^2+0.045*(dist(2)-2)^2+0.11*(dist(3)-4)^2+0.03*(dist(4)-
1)^2+0.23*dist(5)^2+0.0075*dist(7)^2+0.32*(dist(8)-1)^2))+-6.63*exp(-(0.25*(dist(1)-
2)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-4)^2+0.03*(dist(4)-1)^2+0.23*(dist(5)-
4)^2+0.0075*(dist(7)-4)^2+0.32*(dist(8)-1)^2))+3.18*exp(-(0.25*(dist(1)-
3)^2+0.045*(dist(2)-1)^2+0.11*(dist(3)-1)^2+0.031*(dist(4)-2)^2+0.23*(dist(5)-
3)^2+0.0075*(dist(7)-2)^2+0.32*dist(8)^2))+-4.54*exp(-(0.25*(dist(1)-
3)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-1)^2+0.030*(dist(4)-3)^2+0.23*(dist(5)-
1)^2+0.0075*(dist(7)-1)^2+0.32*dist(8)^2))+3.17*exp(-(0.247455342005525*(dist(1)-
3)^2+0.045*(dist(2)-1)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-4)^2+0.23*(dist(5)-
1)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-3)^2))+7.79*exp(-(0.25*(dist(1)-
3)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-4)^2+0.23*(dist(5)-
3)^2+0.0075*(dist(7)-1)^2+0.32*(dist(8)-2)^2))+23.81*exp(-(0.25*(dist(1)-
2)^2+0.045*(dist(2)-1)^2+0.11*(dist(3)-1)^2+0.031*(dist(4)-2)^2+0.23*(dist(5)-
2)^2+0.0075*(dist(7)-1)^2+0.32*dist(8)^2))+0.98*exp(-(0.25*(dist(1)-
3)^2+0.045*(dist(2)-2)^2+0.11*(dist(3)-2)^2+0.031*(dist(4)-4)^2+0.23*(dist(5)-
1)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-1)^2))+2.060*exp(-(0.25*(dist(1)-
3)^2+0.045*(dist(2)-1)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-4)^2+0.23*(dist(5)-
2)^2+0.0075*(dist(7)-1)^2+0.32*(dist(8)-4)^2))+-24.68*exp(-(0.25*(dist(1)-
3)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-2)^2+0.031*(dist(4)-4)^2+0.23*(dist(5)-
3)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-2)^2))+-43.69*exp(-(0.25*(dist(1)-

2)^2+0.045*(dist(2)-2)^2+0.11*(dist(3)-2)^2+0.031*(dist(4)-2)^2+0.23*(dist(5)-
2)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-2)^2))+30.20*exp(-
(0.25*dist(1)^2+0.045*(dist(2)-4)^2+0.11*(dist(3)-2)^2+0.031*(dist(4)-
3)^2+0.23*(dist(5)-1)^2+0.0075*(dist(7)-1)^2+0.32*(dist(8)-2)^2))+-69.020*exp(-
(0.25*dist(1)^2+0.045*dist(2)^2+0.11*(dist(3)-4)^2+0.031*(dist(4)-3)^2+0.23*(dist(5)-
2)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-1)^2))+35.78*exp(-(0.25*(dist(1)-
1)^2+0.045*(dist(2)-2)^2+0.11*dist(3)^2+0.031*(dist(4)-3)^2+0.23*(dist(5)-
4)^2+0.0075*dist(7)^2+0.32*(dist(8)-3)^2))+16.11*exp(-(0.25*(dist(1)-
2)^2+0.045*(dist(2)-1)^2+0.11*dist(3)^2+0.031*(dist(4)-
2)^2+0.23*dist(5)^2+0.0075*(dist(7)-4)^2+0.32*(dist(8)-3)^2))+-33.050*exp(-
(0.25*dist(1)^2+0.045*(dist(2)-4)^2+0.11*(dist(3)-2)^2+0.031*(dist(4)-
3)^2+0.23*(dist(5)-1)^2+0.0075*(dist(7)-4)^2+0.32*(dist(8)-1)^2))+81.65*exp(-
(0.25*dist(1)^2+0.045*dist(2)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-3)^2+0.23*(dist(5)-
2)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-1)^2))+-32.14*exp(-(0.25*(dist(1)-
1)^2+0.045*(dist(2)-2)^2+0.11*dist(3)^2+0.031*(dist(4)-3)^2+0.23*(dist(5)-
4)^2+0.0075*(dist(7)-4)^2+0.32*(dist(8)-3)^2))+2.28*exp(-(0.25*(dist(1)-
2)^2+0.045*(dist(2)-1)^2+0.11*dist(3)^2+0.031*(dist(4)-
3)^2+0.23*dist(5)^2+0.0075*(dist(7)^2+0.32*(dist(8)-3)^2))+-24.76*exp(-
(0.25*(dist(1)-1)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-
2)^2+0.23*(dist(5)-1)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-4)^2))+-61.17*exp(-
(0.25*(dist(1)-1)^2+0.045*(dist(2)-1)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-
1)^2+0.23*(dist(5)-3)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-4)^2))+-12.11*exp(-
(0.25*(dist(1)-1)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-
1)^2+0.031*dist(4)^2+0.23*(dist(5)-3)^2+0.0075*(dist(7)-2)^2+0.32*(dist(8)-
1)^2))+10.80*exp(-(0.25*(dist(1)-1)^2+0.045*(dist(2)-1)^2+0.11*(dist(3)-
1)^2+0.031*dist(4)^2+0.23*(dist(5)-1)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-2)^2))+-
17.89*exp(-(0.25*(dist(1)-2)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-
2)^2+0.23*(dist(5)-2)^2+0.0075*(dist(7)-3)^2+0.32*(dist(8)-4)^2))+70.87*exp(-
(0.25*(dist(1)-1)^2+0.045*(dist(2)-2)^2+0.11*(dist(3)-3)^2+0.031*(dist(4)-
1)^2+0.23*(dist(5)-3)^2+0.0075*(dist(7)-1)^2+0.32*(dist(8)-4)^2))+-4.73*exp(-
(0.25*(dist(1)-2)^2+0.045*(dist(2)-3)^2+0.11*(dist(3)-
1)^2+0.031*dist(4)^2+0.23*(dist(5)-3)^2+0.0075*(dist(7)-
3)^2+0.32*dist(8)^2))+25.01*exp(-(0.25*(dist(1)-1)^2+0.045*(dist(2)-
2)^2+0.11*(dist(3)-2)^2+0.031*dist(4)^2+0.23*(dist(5)-1)^2+0.0075*(dist(7)-
1)^2+0.32*(dist(8)-2)^2)))))

Equation C.2 Phase I Response Prediction Formula for Average Percent Time Occupied

0.88+(-0.0057*exp(-(0.25*(dist(1)-4)^2+0.046*dist(2)^2+0.088*(dist(3)-2)^2+0.020*(dist(4)-1)^2+0.27*(dist(5)-4)^2+0.0055*(dist(7)-3)^2+0.34*(dist(8)-2)^2))+0.267*exp(-(0.25*(dist(1)-4)^2+0.046*(dist(2)-4)^2+0.088*(dist(3)-1)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-3)^2))+(-0.025)*exp(-(0.25*(dist(1)-4)^2+0.046*(dist(2)-2)^2+0.088*(dist(3)-4)^2+0.020*(dist(4)-1)^2+0.27*dist(5)^2+0.0055*(dist(7)-4)^2+0.34*(dist(8)-2)^2))+(-0.023)*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-4)^2+0.088*(dist(3)-4)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-4)^2+0.0055*dist(7)^2+0.34*(dist(8)-1)^2))+0.032*exp(-(0.25*(dist(1)-4)^2+0.046*dist(2)^2+0.088*(dist(3)-2)^2+0.019*(dist(4)-1)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-1)^2+0.34*(dist(8)-3)^2))+(-0.26)*exp(-(0.25*(dist(1)-4)^2+0.046*(dist(2)-4)^2+0.088*(dist(3)-1)^2+0.019*(dist(4)-1)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-3)^2))+0.050*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-2)^2+0.088*(dist(3)-4)^2+0.019*(dist(4)-1)^2+0.27*dist(5)^2+0.0055*dist(7)^2+0.34*(dist(8)-1)^2))+(-0.012)*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-4)^2+0.020*(dist(4)-1)^2+0.27*(dist(5)-4)^2+0.0055*(dist(7)-4)^2+0.34*(dist(8)-1)^2))+0.012*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-1)^2+0.088*(dist(3)-1)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-2)^2+0.34*dist(8)^2))+(-0.0058)*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-1)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-1)^2+0.34*dist(8)^2))+0.0083*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-1)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-4)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-3)^2+0.34*(dist(8)-3)^2))+0.015*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-4)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-1)^2+0.34*(dist(8)-2)^2))+0.048*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-1)^2+0.088*(dist(3)-1)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-1)^2+0.34*dist(8)^2))+(-0.0043)*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-2)^2+0.088*(dist(3)-2)^2+0.019*(dist(4)-4)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-3)^2+0.34*(dist(8)-1)^2))+0.0045*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-1)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-4)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-1)^2+0.34*(dist(8)-4)^2))+(-0.039)*exp(-(0.25*(dist(1)-3)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-2)^2+0.019*(dist(4)-4)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-3)^2+0.34*(dist(8)-2)^2))+(-0.083)*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-

2)^2+0.088*(dist(3)-2)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-2)^2))+0.051*exp(-(0.25*dist(1)^2+0.046*(dist(2)-4)^2+0.088*(dist(3)-2)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-1)^2+0.34*(dist(8)-2)^2))+(-0.14)*exp(-(0.25*dist(1)^2+0.046*dist(2)^2+0.088*(dist(3)-4)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-1)^2))+0.069*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-2)^2+0.088*dist(3)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-4)^2+0.0055*dist(7)^2+0.34*(dist(8)-3)^2))+0.037*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-1)^2+0.088*dist(3)^2+0.019*(dist(4)-2)^2+0.27*dist(5)^2+0.0055*(dist(7)-4)^2+0.34*(dist(8)-3)^2))+(-0.059)*exp(-(0.25*dist(1)^2+0.046*(dist(2)-4)^2+0.088*(dist(3)-2)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-4)^2+0.34*(dist(8)-1)^2))+0.16*exp(-(0.25*dist(1)^2+0.046*dist(2)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-1)^2))+(-0.064)*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-2)^2+0.088*dist(3)^2+0.019*(dist(4)-3)^2+0.27*(dist(5)-4)^2+0.0055*(dist(7)-4)^2+0.34*(dist(8)-3)^2))+0.0012*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-1)^2+0.088*dist(3)^2+0.019*(dist(4)-3)^2+0.27*dist(5)^2+0.0055*dist(7)^2+0.34*(dist(8)-3)^2))+(-0.044)*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-4)^2))+(-0.13)*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-1)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-1)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-3)^2+0.34*(dist(8)-4)^2))+(-0.011)*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-1)^2+0.019*dist(4)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-2)^2+0.34*(dist(8)-1)^2))+(-0.00092)*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-1)^2+0.088*(dist(3)-1)^2+0.019*dist(4)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-3)^2+0.345*(dist(8)-2)^2))+(-0.040)*exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-2)^2+0.27*(dist(5)-2)^2+0.0055*(dist(7)-3)^2+0.34*(dist(8)-4)^2))+0.15*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-2)^2+0.088*(dist(3)-3)^2+0.019*(dist(4)-1)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-1)^2+0.34*(dist(8)-4)^2))+(-0.023)exp(-(0.25*(dist(1)-2)^2+0.046*(dist(2)-3)^2+0.088*(dist(3)-1)^2+0.019*dist(4)^2+0.27*(dist(5)-3)^2+0.0055*(dist(7)-3)^2+0.34*dist(8)^2))+0.066*exp(-(0.25*(dist(1)-1)^2+0.046*(dist(2)-2)^2+0.088*(dist(3)-2)^2+0.019*dist(4)^2+0.27*(dist(5)-1)^2+0.0055*(dist(7)-1)^2+0.34*(dist(8)-2)^2)))

Equation C.3 Phase I Response Prediction Formula for Average Total Cost

-(20.21+(2.95*exp(-(4.38*(dist(4)-1)^2+6.7*(dist(5)-4)^2+1.31*(dist(7)-3)^2+4.23*(dist(8)-2)^2))+2.87*exp(-(4.38*(dist(4)-2)^2+6.7*(dist(5)-2)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-3)^2))+2.91*exp(-(4.388*(dist(4)-1)^2+6.7*dist(5)^2+1.31*(dist(7)-4)^2+4.233*(dist(8)-2)^2))+1.631*exp(-(4.38*(dist(4)-2)^2+6.7*(dist(5)-4)^2+1.31*dist(7)^2+4.23*(dist(8)-1)^2))+0.60*exp(-(4.38*(dist(4)-1)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-1)^2+4.233*(dist(8)-3)^2))+1.68*exp(-(4.38*(dist(4)-1)^2+6.79*(dist(5)-2)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-3)^2))+(-2.92)*exp(-(4.38*(dist(4)-1)^2+6.79*dist(5)^2+1.31*dist(7)^2+4.23*(dist(8)-1)^2))+3.99*exp(-(4.38*(dist(4)-1)^2+6.79*(dist(5)-4)^2+1.31*(dist(7)-4)^2+4.23*(dist(8)-1)^2))+(-4.17)*exp(-(4.387*(dist(4)-2)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-2)^2+4.23*dist(8)^2))+(-2.95)*exp(-(4.388*(dist(4)-3)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-1)^2+4.23*dist(8)^2))+1.71*exp(-(4.38*(dist(4)-4)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-3)^2+4.233*(dist(8)-3)^2))+7.57*exp(-(4.38*(dist(4)-4)^2+6.79*(dist(5)-3)^2+1.313*(dist(7)-1)^2+4.23*(dist(8)-2)^2))+(-7.67)*exp(-(4.38*(dist(4)-2)^2+6.79*(dist(5)-2)^2+1.31924103818723*(dist(7)-1)^2+4.23*dist(8)^2))+1.82*exp(-(4.3878*(dist(4)-4)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-3)^2+4.23*(dist(8)-1)^2))+1.67*exp(-(4.387*(dist(4)-4)^2+6.79*(dist(5)-2)^2+1.31*(dist(7)-1)^2+4.23*(dist(8)-4)^2))+8.76*exp(-(4.387*(dist(4)-4)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-3)^2+4.23*(dist(8)-2)^2))+(-0.803111671064)*exp(-(4.38*(dist(4)-2)^2+6.79*(dist(5)-2)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-2)^2))+(-1.68)*exp(-(4.38*(dist(4)-3)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-1)^2+4.23*(dist(8)-2)^2))+(-1.17451931854076)*exp(-(4.38*(dist(4)-3)^2+6.79*(dist(5)-2)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-1)^2))+(-1.81)*exp(-(4.38*(dist(4)-3)^2+6.79*(dist(5)-4)^2+1.31*dist(7)^2+4.23*(dist(8)-3)^2))+(-1.66)*exp(-(4.387*(dist(4)-2)^2+6.79*dist(5)^2+1.31*(dist(7)-4)^2+4.23*(dist(8)-3)^2))+(-0.60)*exp(-(4.38*(dist(4)-3)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-4)^2+4.23*(dist(8)-1)^2))+(-1.17)*exp(-(4.38*(dist(4)-3)^2+6.79*(dist(5)-2)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-1)^2))+1.74*exp(-(4.387*(dist(4)-3)^2+6.79*(dist(5)-4)^2+1.31*(dist(7)-4)^2+4.23*(dist(8)-3)^2))+(-5.24)*exp(-(4.38*(dist(4)-3)^2+6.79*dist(5)^2+1.31*dist(7)^2+4.23*(dist(8)-3)^2))+2.79*exp(-(4.38*(dist(4)-2)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-4)^2))+1.71*exp(-(4.38706943003598*(dist(4)-1)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-3)^2+4.23*(dist(8)-4)^2))+(-1.88)*exp(-

(4.387*dist(4)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-2)^2+4.23*(dist(8)-1)^2))+(-8.814)*exp(-(4.38*dist(4)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-3)^2+4.23*(dist(8)-2)^2))+7.58*exp(-(4.38*(dist(4)-2)^2+6.79*(dist(5)-2)^2+1.31924103818723*(dist(7)-3)^2+4.23*(dist(8)-4)^2))+0.56*exp(-(4.38*(dist(4)-1)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-1)^2+4.23*(dist(8)-4)^2))+(-0.69)*exp(-(4.38*dist(4)^2+6.79*(dist(5)-3)^2+1.31*(dist(7)-3)^2+4.23*dist(8)^2))+(-8.77797460735803)*exp(-(4.38*dist(4)^2+6.79*(dist(5)-1)^2+1.31*(dist(7)-1)^2+4.23*(dist(8)-2)^2))))

Equation C.4 Phase II Response Prediction Formula for Average Response Time

-(186.21+(2.21*exp(-(0.158*(dist(1)-3)^2+0.25*(dist(3)-1)^2+0.05*(dist(5)-3)^2+0.56*(dist(6)-2)^2+0.01*(dist(7)-2)^2+0.0096*(dist(8)-2)^2))+14.013*exp(-(0.158*(dist(1)-3)^2+0.25*dist(3)^2+0.058*(dist(5)-1)^2+0.561*(dist(6)-1)^2+0.01*(dist(7)-1)^2+0.0096*(dist(8)-3)^2))+4.34923840641689*exp(-(0.15*(dist(1)-3)^2+0.25*(dist(3)-3)^2+0.05*dist(5)^2+0.56*(dist(6)-2)^2+0.013*(dist(7)-3)^2+0.00963*(dist(8)-1)^2))+(-8.91)*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-3)^2+0.05*(dist(5)-3)^2+0.56*dist(6)^2+0.01*dist(7)^2+0.0096*(dist(8)-1)^2))+20.80*exp(-(0.155*(dist(1)-3)^2+0.256*(dist(3)-1)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-2)^2+0.01*dist(7)^2+0.00963*(dist(8)-2)^2))+(-36.97)*exp(-(0.15*(dist(1)-3)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-1)^2+0.561*(dist(6)-1)^2+0.013*(dist(7)-2)^2+0.0096*(dist(8)-2)^2))+8.58*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-3)^2+0.053*dist(5)^2+0.56*(dist(6)-2)^2+0.013*dist(7)^2+0.0096*(dist(8)-1)^2))+(-8.40)*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-3)^2+0.05*(dist(5)-3)^2+0.56*(dist(6)-1)^2+0.01*(dist(7)-3)^2+0.0096*(dist(8)-1)^2))+4.798*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-1)^2+0.013*(dist(7)-2)^2+0.0096*dist(8)^2))+10.62*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-1)^2+0.56*(dist(6)-2)^2+0.0131*(dist(7)-1)^2+0.0096*dist(8)^2))+(-9.268)*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-2)^2+0.053*(dist(5)-1)^2+0.56*dist(6)^2+0.013*(dist(7)-2)^2+0.0096*(dist(8)-2)^2))+4.69*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-2)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-3)^2+0.013*dist(7)^2+0.0096*(dist(8)-2)^2))+17.98*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-2)^2+0.56*dist(6)^2+0.013*(dist(7)-1)^2+0.0096*dist(8)^2))+8.90*exp(-(0.155*(dist(1)-3)^2+0.25*(dist(3)-1)^2+0.053*dist(5)^2+0.56*(dist(6)-2)^2+0.013*(dist(7)-

2)^2+0.0096*dist(8)^2))+4.30*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-
3)^2+0.053*(dist(5)-1)^2+0.56*dist(6)^2+0.013*(dist(7)-1)^2+0.0096*(dist(8)-3)^2))+(-
30.394)*exp(-(0.15*(dist(1)-2)^2+0.25*(dist(3)-2)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-
3)^2+0.013*(dist(7)-2)^2+0.0096*(dist(8)-2)^2))+(-39.94)*exp(-(0.15*(dist(1)-
2)^2+0.25*(dist(3)-2)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-2)^2+0.013*(dist(7)-
2)^2+0.0096*(dist(8)-2)^2))+9.72*exp(-(0.15*dist(1)^2+0.25*(dist (3)-
2)^2+0.053*dist(5)^2+0.56*(dist(6)-1)^2+0.013*(dist(7)-1)^2+0.0096*(dist(8)-1)^2))+(-
3.08)*exp(-(0.15*dist(1)^2+0.25*(dist(3)-3)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-
2)^2+0.013*(dist(7)-2)^2+0.00963*dist(8)^2))+10.37*exp(-
(0.15*dist(1)^2+0.25*dist(3)^2+0.053*(dist(5)-3)^2+0.56*(dist(6)-
1)^2+0.013*dist(7)^2+0.0096*(dist(8)-2)^2))+(-7.33)*exp(-(0.155*(dist(1)-
1)^2+0.25*dist(3)^2+0.053*dist(5)^2+0.56*(dist(6)-3)^2+0.013*(dist(7)-
3)^2+0.0096*(dist(8)-2)^2))+11.35*exp(-(0.15*dist(1)^2+0.25*(dist(3)-
2)^2+0.053*(dist(5)-1)^2+0.56*(dist(6)-1)^2+0.013*(dist(7)-3)^2+0.0096*(dist(8)-
1)^2))+1.23*exp(-(0.15*dist(1)^2+0.25*(dist(3)-2)^2+0.05*(dist(5)-2)^2+0.56*(dist(6)-
2)^2+0.013*(dist(7)-1)^2+0.0096*(dist(8)-1)^2))+(-25.93)*exp(-(0.15*(dist(1)-
1)^2+0.25*dist(3)^2+0.053*(dist(5)-3)^2+0.56*(dist(6)-1)^2+0.01*(dist(7)-
3)^2+0.0096*(dist(8)-2)^2))+(-8.21)*exp(-(0.15*(dist(1)-
1)^2+0.25*dist(3)^2+0.053*dist(5)^2+0.56*(dist(6)-
2)^2+0.013*dist(7)^2+0.0096*(dist(8)-2)^2))+31.77*exp(-(0.15*(dist(1)-
1)^2+0.25*(dist(3)-2)^2+0.05*(dist(5)-1)^2+0.56*(dist(6)-2)^2+0.013*(dist(7)-
1)^2+0.0096*(dist(8)-3)^2))+(-42.28)*exp(-(0.15*(dist(1)-1)^2+0.25*(dist(3)-
2)^2+0.05*(dist(5)-2)^2+0.56*(dist(6)-1)^2+0.013*(dist(7)-2)^2+0.0096*(dist(8)-
3)^2))+(-13.10)*exp(-(0.15*(dist(1)-1)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-
2)^2+0.56*(dist(6)-3)^2+0.013*(dist(7)-1)^2+0.0096*(dist(8)-1)^2))+8.30*exp(-
(0.15*(dist(1)-1)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-
1)^2+0.56*dist(6)^2+0.013*(dist(7)-3)^2+0.0096*(dist(8)-1)^2))+(-4.72)*exp(-
(0.15*(dist(1)-1)^2+0.25*(dist(3)-2)^2+0.053*(dist(5)-1)^2+0.56*(dist(6)-
3)^2+0.013*(dist(7)-2)^2+0.0096*(dist(8)-3)^2))+14.51*exp(-
(0.15*dist(1)^2+0.25*(dist(3)-2)^2+0.053*(dist(5)-3)^2+0.56*(dist(6)-
1)^2+0.013*(dist(7)-1)^2+0.0096*(dist(8)-3)^2))+30.1863420257917*exp(-
(0.15*(dist(1)-1)^2+0.25*dist(3)^2+0.053*(dist(5)-2)^2+0.56*(dist(6)-
3)^2+0.013*(dist(7)-2)^2+0.0096*dist(8)^2))+19.82*exp(-(0.15*(dist(1)-
1)^2+0.25*(dist(3)-1)^2+0.053*(dist(5)-1)^2+0.56*dist(6)^2+0.013*(dist(7)-
1)^2+0.0096*(dist(8)-1)^2))))

Equation C.5 Phase II Response Prediction Formula for Average Percent Time Occupied

0.90+(-0.0076*exp(-(73.86*dist(2)^2+0.00015*(dist(4)-1)^2+73.059*(dist(5)-3)^2+32.29*(dist(7)-2)^2+0.17*(dist(8)-2)^2))+0.0058*exp(-(73.86*(dist(2)-3)^2+0.00015*(dist(4)-1)^2+73.059*(dist(5)-1)^2+32.29*(dist(7)-1)^2+0.17*(dist(8)-3)^2))+(-0.0068)*exp(-(73.86*(dist(2)-1)^2+0.00015*dist(4)^2+73.05*dist(5)^2+32.29*(dist(7)-3)^2+0.17*(dist(8)-1)^2))+(-0.016)*exp(-(73.86*(dist(2)-3)^2+0.000154*(dist(4)-1)^2+73.05*(dist(5)-3)^2+32.29*dist(7)^2+0.17*(dist(8)-1)^2))+0.0081*exp(-(73.86*dist(2)^2+0.00015*(dist(4)-1)^2+73.059*(dist(5)-2)^2+32.299*dist(7)^2+0.17*(dist(8)-2)^2))+(-0.020)*exp(-(73.866*(dist(2)-3)^2+0.00015*(dist(4)-1)^2+73.05*(dist(5)-1)^2+32.29*(dist(7)-2)^2+0.17*(dist(8)-2)^2))+(-0.051)*exp(-(73.86*(dist(2)-1)^2+0.00015*(dist(4)-1)^2+73.059*dist(5)^2+32.29*dist(7)^2+0.17*(dist(8)-1)^2))+(-0.13)*exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-1)^2+73.05*(dist(5)-3)^2+32.29*(dist(7)-3)^2+0.17*(dist(8)-1)^2))+(-0.00953771987572999)*exp(-(73.8663234706366*(dist(2)-1)^2+0.00015*(dist(4)-2)^2+73.05*(dist(5)-2)^2+32.29*(dist(7)-2)^2+0.17*dist(8)^2))+(-0.07)*exp(-(73.8663234706366*(dist(2)-2)^2+0.00015*(dist(4)-2)^2+73.059*(dist(5)-1)^2+32.29*(dist(7)-1)^2+0.17*dist(8)^2))+0.0044*exp(-(73.86*(dist(2)-1)^2+0.00015*(dist(4)-3)^2+73.059*(dist(5)-1)^2+32.29*(dist(7)-2)^2+0.17*(dist(8)-2)^2))+(-0.034)*exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-3)^2+73.059*(dist(5)-2)^2+32.29*dist(7)^2+0.17*(dist(8)-2)^2))+0.093*exp(-(73.86*dist(2)^2+0.00015*(dist(4)-2)^2+73.059*(dist(5)-2)^2+32.29*(dist(7)-1)^2+0.17*dist(8)^2))+0.012exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-3)^2+73.059*dist(5)^2+32.29*(dist(7)-2)^2+0.17*dist(8)^2))+(-0.04)*exp(-(73.86*(dist(2)-1)^2+0.00015dist(4)-3)^2+73.059*(dist(5)-1)^2+32.299*(dist(7)-1)^2+0.17*(dist(8)-3)^2))+(-4.64)*exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-3)^2+73.05*(dist(5)-2)^2+32.29*(dist(7)-2)^2+0.17*(dist(8)-2)^2))+4.567*exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-2)^2+73.059dist(5)-2)^2+32.29*(dist(7)-2)^2+0.17*(dist(8)-2)^2))+(-0.15)*exp(-(73.86*(dist(2)-3)^2+0.00015*(dist(4)-2)^2+73.059*dist(5)^2+32.29*(dist(7)-1)^2+0.17*(dist(8)-1)^2))+(-0.18)*exp(-(73.86*dist(2)^2+0.00015*(dist(4)-2)^2+73.05*(dist(5)-2)^2+32.29*(dist(7)-2)^2+0.17*dist(8)^2))+0.01*exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-3)^2+73.05*(dist(5)-3)^2+32.29*dist(7)^2+0.17*(dist(8)-2)^2))+0.019*exp(-

(73.86*dist(2)^2+0.00015*(dist(4)-2)^2+73.05*dist(5)^2+32.29*(dist(7)-3)^2+0.17*(dist(8)-2)^2))+0.021*exp(-(73.86*(dist(2)-3)^2+0.00015*(dist(4)-2)^2+73.058*(dist(5)-1)^2+32.29*(dist(7)-3)^2+0.17*(dist(8)-1)^2))+(-0.069)*exp(-(73.86*dist(2)^2+0.00015*(dist(4)-2)^2+73.059*(dist(5)-2)^2+32.29*(dist(7)-1)^2+0.17*(dist(8)-1)^2))+0.10*exp(-(73.8663234706366*(dist(2)-2)^2+0.00015*(dist(4)-2)^2+73.05*(dist(5)-3)^2+32.29*(dist(7)-3)^2+0.17*(dist(8)-2)^2))+0.06*exp(-(73.86*(dist(2)-1)^2+0.00015*(dist(4)-2)^2+73.059*dist(5)^2+32.29*dist(7)^2+0.17*(dist(8)-2)^2))+0.37*exp(-(73.86*(dist(2)-2)^2+0.00015*(dist(4)-1)^2+73.059*(dist(5)-1)^2+32.29*(dist(7)-1)^2+0.17*(dist(8)-3)^2))+(-0.029)*exp(-(73.86*(dist(2)-1)^2+0.00015*(dist(4)-1)^2+73.05*(dist(5)-2)^2+32.29*(dist(7)-2)^2+0.172004439961235*(dist(8)-3)^2))+(-0.0048)*exp(-(73.86*(dist(2)-2)^2+0.00015*dist(4)^2+73.059*(dist(5)-2)^2+32.29*(dist(7)-1)^2+0.172004439961235*(dist(8)-1)^2))+0.047*exp(-(73.86*(dist(2)-1)^2+0.00015*dist(4)^2+73.05*(dist(5)-1)^2+32.29*(dist(7)-3)^2+0.17*(dist(8)-1)^2))+(-0.011)*exp(-(73.86*(dist(2)-3)^2+0.00015*(dist(4)-1)^2+73.059*(dist(5)-1)^2+32.29*(dist(7)-2)^2+0.172004439961235*(dist(8)-3)^2))+0.0092*exp(-(73.86*(dist(2)-1)^2+0.00015*dist(4)^2+73.059*(dist(5)-3)^2+32.29*(dist(7)-1)^2+0.17*(dist(8)-3)^2))+0.073*exp(-(73.86*(dist(2)-2)^2+0.00015*dist(4)^2+73.059*(dist(5)-2)^2+32.29*(dist(7)-2)^2+0.17*dist(8)^2))+0.072*exp(-(73.86*(dist(2)-1)^2+0.00015*dist(4)^2+73.059*(dist(5)-1)^2+32.29*(dist(7)-1)^2+0.17*(dist(8)-1)^2)))

Equation C.6 Phase II Response Prediction Formula for Average Total Cost

-(13.24+(30.26*exp(-(0.018*(dist(1)-3)^2+0.11*dist(2)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-3)^2+0.011*(dist(6)-2)^2+0.0090*(dist(7)-2)^2+0.16*(dist(8)-2)^2))+(-12.45)*exp(-(0.018*(dist(1)-3)^2+0.11*(dist(2)-3)^2+0.010*dist(3)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-1)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-3)^2))+2.59423691511974*exp(-(0.018*(dist(1)-3)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-3)^2+0.0085*dist(4)^2+0.062*dist(5)^2+0.011*(dist(6)-2)^2+0.0091*(dist(7)-3)^2+0.16*(dist(8)-1)^2))+5.38*exp(-(0.018*(dist(1)-2)^2+0.11*(dist(2)-3)^2+0.010*(dist(3)-3)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-3)^2+0.011*dist(6)^2+0.0090*dist(7)^2+0.16*(dist(8)-1)^2))+(-25.10)*exp(-(0.018*(dist(1)-3)^2+0.11*dist(2)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-2)^2+0.0118490768935915*(dist(6)-

2)^2+0.0090*dist(7)^2+0.16*(dist(8)-2)^2))+23.16*exp(-(0.018*(dist(1)-
3)^2+0.11*(dist(2)-3)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-
1)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-2)^2+0.16*(dist(8)-2)^2))+12.27*exp(-
(0.018*(dist(1)-2)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-3)^2+0.0085*(dist(4)-
1)^2+0.062*dist(5)^2+0.011*(dist(6)-2)^2+0.0090*dist(7)^2+0.16*(dist(8)-
1)^2))+11.39*exp(-(0.018*(dist(1)-2)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-
3)^2+0.0085*(dist(4)-1)^2+0.0622*(dist(5)-3)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-
3)^2+0.16*(dist(8)-1)^2))+14.38*exp(-(0.018*(dist(1)-2)^2+0.11*(dist(2)-
1)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-2)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-
1)^2+0.0090*(dist(7)-2)^2+0.16*dist(8)^2))+1.11*exp(-(0.018*(dist(1)-
2)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-2)^2+0.062*(dist(5)-
1)^2+0.011*(dist(6)-2)^2+0.0090*(dist(7)-1)^2+0.165*dist(8)^2))+7.39*exp(-
(0.018*(dist(1)-2)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-2)^2+0.0085*(dist(4)-
3)^2+0.0622*(dist(5)-1)^2+0.011*dist(6)^2+0.0090*(dist(7)-2)^2+0.16*(dist(8)-
2)^2))+(-33.54)*exp(-(0.018*(dist(1)-2)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-
2)^2+0.0085*(dist(4)-3)^2+0.062*(dist(5)-2)^2+0.0118490768935915*(dist(6)-
3)^2+0.0090*dist(7)^2+0.16*(dist(8)-2)^2))+(-6.15)*exp(-(0.018*(dist(1)-
2)^2+0.11*dist(2)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-2)^2+0.062*(dist(5)-
2)^2+0.011*dist(6)^2+0.0090*(dist(7)-1)^2+0.16*dist(8)^2))+7.55*exp(-(0.018*(dist(1)-
3)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-1)^2+0.0085*(dist(4)-
3)^2+0.062*dist(5)^2+0.011*(dist(6)-2)^2+0.0090*(dist(7)-2)^2+0.16*dist(8)^2))+(-
7.22)*exp(-(0.018*(dist(1)-2)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-
3)^2+0.0085*(dist(4)-3)^2+0.062*(dist(5)-1)^2+0.011*dist(6)^2+0.0090*(dist(7)-
1)^2+0.165*(dist(8)-3)^2))+94.80*exp(-(0.018*(dist(1)-2)^2+0.11*(dist(2)-
2)^2+0.010*(dist(3)-2)^2+0.0085*(dist(4)-3)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-
3)^2+0.0090*(dist(7)-2)^2+0.16*(dist(8)-2)^2))+(-84.90)*exp(-(0.018*(dist(1)-
2)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-2)^2+0.0085*(dist(4)-2)^2+0.062*(dist(5)-
2)^2+0.011*(dist(6)-2)^2+0.0090*(dist(7)-2)^2+0.16*(dist(8)-2)^2))+(-40.082)*exp(-
(0.018*dist(1)^2+0.11*(dist(2)-3)^2+0.010*(dist(3)-2)^2+0.0085*(dist(4)-
2)^2+0.062*dist(5)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-1)^2))+(-
27.76)*exp(-(0.018*dist(1)^2+0.11*dist(2)^2+0.010*(dist(3)-
3)^2+0.0085381830964748*(dist(4)-2)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-
2)^2+0.0090*(dist(7)-2)^2+0.165*dist(8)^2))+9.959*exp(-
(0.018*dist(1)^2+0.11*(dist(2)-2)^2+0.010*dist(3)^2+0.00853*(dist(4)-
3)^2+0.062*(dist(5)-3)^2+0.011*(dist(6)-1)^2+0.0090*dist(7)^2+0.16*(dist(8)-2)^2))+(-
10.32)*exp(-(0.018*(dist(1)-1)^2+0.11*dist(2)^2+0.010*dist(3)^2+0.0085*(dist(4)-

2)^2+0.0622055731237769*dist(5)^2+0.0118490768935915*(dist(6)-3)^2+0.0090*(dist(7)-3)^2+0.16*(dist(8)-2)^2))+30.144940129768*exp(-(0.018*dist(1)^2+0.11*(dist(2)-3)^2+0.01*(dist(3)-2)^2+0.0085*(dist(4)-2)^2+0.062*(dist(5)-1)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-3)^2+0.16*(dist(8)-1)^2))+35.57*exp(-(0.018*dist(1)^2+0.11*dist(2)^2+0.01*(dist(3)-2)^2+0.0085381830964748*(dist(4)-2)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-2)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-1)^2))+(-21.84418476183)*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-2)^2+0.010*dist(3)^2+0.0085*(dist(4)-2)^2+0.06*(dist(5)-3)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-3)^2+0.16*(dist(8)-2)^2))+(-1.83)*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-1)^2+0.010*dist(3)^2+0.0085*(dist(4)-2)^2+0.062*dist(5)^2+0.011*(dist(6)-2)^2+0.0090*dist(7)^2+0.16*(dist(8)-2)^2))+76.97*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-2)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-1)^2+0.011*(dist(6)-2)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-3)^2))+(-24.70)*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-2)^2+0.0085*(dist(4)-1)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-1)^2+0.0090*(dist(7)-2)^2+0.165702986391966*(dist(8)-3)^2))+(-8.31)*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-2)^2+0.010*(dist(3)-1)^2+0.0085*dist(4)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-3)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-1)^2))+10.22*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-1)^2+0.0085*dist(4)^2+0.062*(dist(5)-1)^2+0.011*dist(6)^2+0.0090*(dist(7)-3)^2+0.16*(dist(8)-1)^2))+(-28.5205512161643)*exp(-(0.018*(dist(1)-1)^2+0.11(dist(2)-3)^2+0.010*(dist(3)-2)^2+0.008*(dist(4)-1)^2+0.06*(dist(5)-1)^2+0.011*(dist(6)-3)^2+0.0090*(dist(7)-2)^2+0.16*(dist(8)-3)^2))+(-0.81)*exp(-(0.018*dist(1)^2+0.11*(dist(2)-1)^2+0.010*(dist(3)-2)^2+0.0085*dist(4)^2+0.062*(dist(5)-3)^2+0.0118490768935915*(dist(6)-1)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-3)^2))+(-5.94)*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-2)^2+0.010*dist(3)^2+0.0085*dist(4)^2+0.062*(dist(5)-2)^2+0.011*(dist(6)-3)^2+0.0090*(dist(7)-2)^2+0.165702986391966*dist(8)^2))+(-33.65)*exp(-(0.018*(dist(1)-1)^2+0.11*(dist(2)-1)^2+0.0103143784510041*(dist(3)-1)^2+0.0085*dist(4)^2+0.062*(dist(5)-1)^2+0.011*dist(6)^2+0.0090*(dist(7)-1)^2+0.16*(dist(8)-1)^2))))

Equation C.7 Phase III Response Prediction Formula for Average Response Time

-(326.80+(777.45*exp(-(0.010*(dist(1)-2)^2+0.03*dist(2)^2+0.0049*(dist(3)-1)^2+0.02*dist(4)^2+0.03*(dist(5)-2)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-

1)^2+0.036*(dist(8)-1)^2))+(-2472.59)*exp(-(0.010*(dist(1)-2)^2+0.035*(dist(2)-
2)^2+0.0049*dist(3)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-
1)^2+0.024*dist(6)^2+0.0107*(dist(7)-1)^2+0.036*(dist(8)-2)^2))+(-620.56)*exp(-
(0.010*(dist(1)-2)^2+0.03*(dist(2)-1)^2+0.004*(dist(3)-
2)^2+0.025*dist(4)^2+0.03*dist(5)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-
2)^2+0.036*(dist(8)-1)^2))+(-135.54)*exp(-(0.010*(dist(1)-1)^2+0.0355*(dist(2)-
2)^2+0.0049*(dist(3)-2)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-
2)^2+0.024*dist(6)^2+0.010*dist(7)^2+0.036*dist(8)^2))+(-322.859595254887)*exp(-
(0.010*(dist(1)-2)^2+0.03*dist(2)^2+0.0049*(dist(3)-
1)^2+0.025*dist(4)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-
1)^2+0.010*dist(7)^2+0.036*(dist(8)-1)^2))+2825.79*exp(-(0.010*(dist(1)-
2)^2+0.035*(dist(2)-2)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-
1)^2+0.024*dist(6)^2+0.010*(dist(7)-1)^2+0.036*(dist(8)-2)^2))+(-333.73)*exp(-
(0.010*(dist(1)-1)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-2)^2+0.025*(dist(4)-
1)^2+0.032*dist(5)^2+0.0240*(dist(6)-1)^2+0.010*dist(7)^2+0.036*(dist(8)-1)^2))+(-
743.758)*exp(-(0.010*(dist(1)-1)^2+0.03*(dist(2)-1)^2+0.0049*(dist(3)-
2)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-2)^2+0.0240451012885587*(dist(6)-
1)^2+0.010*(dist(7)-2)^2+0.036*dist(8)^2))+(-3460.35484340777)*exp(-
(0.010*(dist(1)-1)^2+0.03*(dist(2)-1)^2+0.0049*dist(3)^2+0.025*(dist(4)-
1)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-1)^2+0.036*dist(8)^2))+(-
327.68)*exp(-(0.010*(dist(1)-2)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-
1)^2+0.025*(dist(4)-1)^2+0.032*dist(5)^2+0.024*(dist(6)-
1)^2+0.010*dist(7)^2+0.036*dist(8)^2))+(-451.74)*exp(-(0.010*(dist(1)-
1)^2+0.035*dist(2)^2+0.0049*(dist(3)-2)^2+0.025*(dist(4)-2)^2+0.032*(dist(5)-
1)^2+0.0240*dist(6)^2+0.010*(dist(7)-1)^2+0.0360*(dist(8)-1)^2))+(-789.790)*exp(-
(0.010*(dist(1)-2)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-1)^2+0.0259*(dist(4)-
2)^2+0.032*(dist(5)-2)^2+0.0240*(dist(6)-2)^2+0.010*dist(7)^2+0.0360*(dist(8)-
1)^2))+1532.14*exp(-(0.010*(dist(1)-1)^2+0.035*dist(2)^2+0.0049*dist
(3)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-1)^2+0.024*dist(6)^2+0.010*(dist(7)-
1)^2+0.036*dist(8)^2))+562.13*exp(-(0.010*(dist(1)-2)^2+0.035*(dist(2)-
1)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-
2)^2+0.0324943350254703*dist(5)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-
2)^2+0.036*dist(8)^2))+(-778.90)*exp(-(0.010*(dist(1)-
1)^2+0.03*dist(2)^2+0.0049*(dist(3)-2)^2+0.02*(dist(4)-2)^2+0.032*(dist(5)-
1)^2+0.0240*dist(6)^2+0.010*(dist(7)-1)^2+0.036*(dist(8)-2)^2))+552.37*exp(-
(0.010*(dist(1)-2)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-

2)^2+0.0324943350254703*(dist(5)-2)^2+0.024*(dist(6)-2)^2+0.010*(dist(7)-
2)^2+0.036*(dist(8)-1)^2))+2952.57*exp(-(0.010*(dist(1)-1)^2+0.035*(dist(2)-
1)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-
1)^2+0.010*(dist(7)-1)^2+0.036*(dist(8)-1)^2))+(-1153.45)*exp(-
(0.010*dist(1)^2+0.035*(dist(2)-2)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-
2)^2+0.032*dist(5)^2+0.024*(dist(6)-1)^2+0.0107588794338961*(dist(7)-
1)^2+0.036*(dist(8)-1)^2))+1287.91*exp(-
(0.0109660143774632*dist(1)^2+0.035*dist(2)^2+0.0049*(dist(3)-2)^2+0.025*(dist(4)-
1)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-2)^2+0.010*(dist(7)-
1)^2+0.036*dist(8)^2))+1473.71*exp(-(0.010*dist(1)^2+0.035*(dist(2)-
1)^2+0.0049*dist(3)^2+0.025*(dist(4)-2)^2+0.03*(dist(5)-2)^2+0.024*(dist(6)-
1)^2+0.010*dist(7)^2+0.036*(dist(8)-1)^2))+551.86*exp(-(0.010*(dist(1)-
1)^2+0.035*dist(2)^2+0.0049*dist(3)^2+0.025*(dist(4)-
1)^2+0.032*dist(5)^2+0.024*(dist(6)-2)^2+0.010*(dist(7)-2)^2+0.036*(dist(8)-
2)^2))+1180.67*exp(-(0.010*dist(1)^2+0.035*(dist(2)-2)^2+0.00492*(dist(3)-
1)^2+0.025*(dist(4)-2)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-
2)^2+0.036*(dist(8)-1)^2))+(-2251.17)*exp(-
(0.010*dist(1)^2+0.035*dist(2)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-
1)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-2)^2+0.010*(dist(7)-
1)^2+0.0360617360074524*(dist(8)-1)^2))+(-775.85)*exp(-(0.010*(dist(1)-
1)^2+0.035*(dist(2)-1)^2+0.0049*dist(3)^2+0.025*(dist(4)-2)^2+0.032*(dist(5)-
2)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-2)^2+0.036*(dist(8)-1)^2))+421.35*exp(-
(0.010*(dist(1)-1)^2+0.035*(dist(2)-1)^2+0.0049*dist(3)^2+0.025*(dist(4)-
1)^2+0.0324943350254703*dist(5)^2+0.024*(dist(6)-
2)^2+0.010*dist(7)^2+0.03*(dist(8)-2)^2))+(-1116.81)*exp(-(0.010*(dist(1)-
1)^2+0.035*(dist(2)-2)^2+0.0049*(dist(3)-2)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-
1)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-1)^2+0.036*(dist(8)-2)^2))+386.077*exp(-
(0.010*(dist(1)-1)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-1)^2+0.025*(dist(4)-
1)^2+0.03*(dist(5)-2)^2+0.024*(dist(6)-1)^2+0.010*(dist(7)-2)^2+0.036*(dist(8)-
2)^2))+(-2151.67)*exp(-(0.010*(dist(1)-1)^2+0.03*(dist(2)-2)^2+0.0049*(dist(3)-
1)^2+0.025*dist(4)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-2)^2+0.010*(dist(7)-
1)^2+0.0360617360074524*(dist(8)-1)^2))+(-1620.42)*exp(-
(0.010*dist(1)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-
1)^2+0.025*dist(4)^2+0.032*(dist(5)-1)^2+0.024*dist(6)^2+0.010*(dist(7)-
2)^2+0.036*(dist(8)-1)^2))+823.63*exp(-(0.010*(dist(1)-1)^2+0.035*(dist(2)-
2)^2+0.0049*(dist(3)-2)^2+0.025*(dist(4)-1)^2+0.032*(dist(5)-1)^2+0.024*(dist(6)-

2)^2+0.010*(dist(7)-1)^2+0.036*(dist(8)-2)^2))+(-46.48)*exp(-
(0.010*dist(1)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-
1)^2+0.025*dist(4)^2+0.032*(dist(5)-2)^2+0.024*(dist(6)-
1)^2+0.010*dist(7)^2+0.036*(dist(8)-2)^2))+2331.62*exp(-(0.010*(dist(1)-
1)^2+0.035*(dist(2)-2)^2+0.0049*dist(3)^2+0.025*dist(4)^2+0.032*(dist(5)-
1)^2+0.024*(dist(6)-2)^2+0.010*(dist(7)-1)^2+0.036*dist(8)^2))+1894.10*exp(-
(0.010*dist(1)^2+0.035*(dist(2)-1)^2+0.0049*(dist(3)-
1)^2+0.025*dist(4)^2+0.032*dist(5)^2+0.024*dist(6)^2+0.010*(dist(7)-
1)^2+0.036*(dist(8)-1)^2))))

Equation C.8 Phase III Response Prediction Formula for Average Percent Time Occupied

1.23+(0.30*exp(-(0.0063*(dist(1)-2)^2+0.027*dist(2)^2+0.0034*(dist(3)-
1)^2+0.024*dist(4)^2+0.025*(dist(5)-2)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-
1)^2+0.031*(dist(8)-1)^2))+(-4.66)*exp(-(0.00634274625236588*(dist(1)-
2)^2+0.027*(dist(2)-2)^2+0.0034*dist(3)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-
1)^2+0.018*dist(6)^2+0.010*(dist(7)-1)^2+0.031*(dist(8)-2)^2))+(-2.13)*exp(-
(0.0063*(dist(1)-2)^2+0.027*(dist(2)-1)^2+0.0034*(dist(3)-
2)^2+0.024*dist(4)^2+0.025*dist(5)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-
2)^2+0.031*(dist(8)-1)^2))+(-0.26)*exp(-(0.0063*(dist(1)-1)^2+0.027*(dist(2)-
2)^2+0.0034*(dist(3)-2)^2+0.02*(dist(4)-1)^2+0.025*(dist(5)-
2)^2+0.018*dist(6)^2+0.010*dist(7)^2+0.031*dist(8)^2))+1.61*exp(-(0.0063*(dist(1)-
2)^2+0.027065168178968 7*dist(2)^2+0.0034*(dist(3)-
1)^2+0.024*dist(4)^2+0.025*(dist(5)-1)^2+0.018*(dist(6)-
1)^2+0.010*dist(7)^2+0.031*(dist(8)-1)^2))+4.79*exp(-(0.0063*(dist(1)-
2)^2+0.027*(dist(2)-2)^2+0.0034*(dist(3)-1)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-
1)^2+0.018*dist(6)^2+0.010*(dist(7)-1)^2+0.031*(dist(8)-2)^2))+(-1.48)*exp(-
(0.0063*(dist(1)-1)^2+0.027*(dist(2)-1)^2+0.0034*(dist(3)-2)^2+0.024*(dist(4)-
1)^2+0.025*dist(5)^2+0.018*(dist(6)-1)^2+0.010*dist(7)^2+0.031*(dist(8)-1)^2))+(-
0.97)*exp(-(0.0063*(dist(1)-1)^2+0.027*(dist(2)-1)^2+0.0034*(dist(3)-
2)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-2)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-
2)^2+0.031*dist(8)^2))+(-11.04)*exp(-(0.00634*(dist(1)-1)^2+0.027*(dist(2)-
1)^2+0.0034*dist(3)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-1)^2+0.018*(dist(6)-
1)^2+0.010*(dist(7)-1)^2+0.031*dist(8)^2))+(-1.53)*exp(-(0.00634*(dist(1)-
2)^2+0.027*(dist(2)-1)^2+0.00345210627022864*(dist(3)-1)^2+0.024*(dist(4)-

1)^2+0.025*dist(5)^2+0.018*(dist(6)-1)^2+0.010*dist(7)^2+0.031*dist(8)^2))+(-
1.231)*exp(-(0.0063*(dist(1)-1)^2+0.027*dist(2)^2+0.0034*(dist(3)-
2)^2+0.024*(dist(4)-2)^2+0.025*(dist(5)-1)^2+0.018*dist(6)^2+0.010*(dist(7)-
1)^2+0.031*(dist(8)-1)^2))+(-2.57)*exp(-(0.0063*(dist(1)-2)^2+0.027*(dist(2)-
1)^2+0.0034*(dist(3)-1)^2+0.024*(dist(4)-2)^2+0.025*(dist(5)-2)^2+0.0183*(dist(6)-
2)^2+0.010*dist(7)^2+0.031*(dist(8)-1)^2))+3.98*exp(-(0.0063*(dist(1)-
1)^2+0.027*dist(2)^2+0.0034*dist(3)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-
1)^2+0.0189*dist(6)^2+0.0103*(dist(7)-1)^2+0.031*dist(8)^2))+2.42*exp(-
(0.0063*(dist(1)-2)^2+0.0270*(dist(2)-1)^2+0.0034*(dist(3)-1)^2+0.024*(dist(4)-
2)^2+0.025*dist(5)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-2)^2+0.031*dist(8)^2))+(-
2.41)*exp(-(0.00634*(dist(1)-1)^2+0.027*dist(2)^2+0.00345*(dist(3)-
2)^2+0.024*(dist(4)-2)^2+0.025*(dist(5)-1)^2+0.018*dist(6)^2+0.010*(dist(7)-
1)^2+0.031*(dist(8)-2)^2))+1.70*exp(-(0.0063*(dist(1)-2)^2+0.027*(dist(2)-
1)^2+0.0034*(dist(3)-1)^2+0.024*(dist(4)-2)^2+0.025*(dist(5)-2)^2+0.018*(dist(6)-
2)^2+0.010*(dist(7)-2)^2+0.031*(dist(8)-1)^2))+8.025*exp(-(0.0063*(dist(1)-
1)^2+0.027*(dist(2)-1)^2+0.00345*(dist(3)-1)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-
1)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-1)^2+0.031*(dist(8)-1)^2))+(-0.71)*exp(-
(0.00634274625236588*dist(1)^2+0.0270*(dist(2)-2)^2+0.0034*(dist(3)-
1)^2+0.0246*(dist(4)-2)^2+0.025*dist(5)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-
1)^2+0.031*(dist(8)-1)^2))+3.0398*exp(-
(0.0063*dist(1)^2+0.027*dist(2)^2+0.0034*(dist(3)-2)^2+0.024*(dist(4)-
1)^2+0.025*(dist(5)-1)^2+0.018*(dist(6)-2)^2+0.010*(dist(7)-
1)^2+0.031*dist(8)^2))+4.89*exp(-(0.0063*dist(1)^2+0.027*(dist(2)-
1)^2+0.00345*dist(3)^2+0.024*(dist(4)-2)^2+0.025*(dist(5)-2)^2+0.018*(dist(6)-
1)^2+0.010*dist(7)^2+0.031*(dist(8)-1)^2))+1.42*exp(-(0.0063*(dist(1)-
1)^2+0.027*dist(2)^2+0.0034*dist(3)^2+0.024*(dist(4)-
1)^2+0.025*dist(5)^2+0.018*(dist(6)-2)^2+0.010*(dist(7)-2)^2+0.031*(dist(8)-2)^2))+(-
0.261)*exp(-(0.0063*dist(1)^2+0.027*(dist(2)-2)^2+0.0034*(dist(3)-
1)^2+0.024*(dist(4)-2)^2+0.0256974880147803*(dist(5)-1)^2+0.018*(dist(6)-
1)^2+0.010*(dist(7)-2)^2+0.031*(dist(8)-1)^2))+(-5.45)*exp(-
(0.0063*dist(1)^2+0.027*dist(2)^2+0.0034*(dist(3)-1)^2+0.024*(dist(4)-
1)^2+0.025*(dist(5)-1)^2+0.018*(dist(6)-2)^2+0.01*(dist(7)-1)^2+0.031*(dist(8)-
1)^2))+(-1.66)*exp(-(0.0063*(dist(1)-1)^2+0.027*(dist(2)-
1)^2+0.0034*dist(3)^2+0.024*(dist(4)-2)^2+0.025*(dist(5)-2)^2+0.018*(dist(6)-
1)^2+0.010*(dist(7)-2)^2+0.031*(dist(8)-1)^2))+0.32*exp(-(0.0063*(dist(1)-
1)^2+0.027*(dist(2)-1)^2+0.00345*dist(3)^2+0.0246764144934028*(dist(4)-

1)^2+0.0256974880147803*dist(5)^2+0.018*(dist(6)-
2)^2+0.010*dist(7)^2+0.031*(dist(8)-2)^2))+(-1.055)*exp(-(0.0063*(dist(1)-
1)^2+0.027*(dist(2)-2)^2+0.0034*(dist(3)-2)^2+0.024*(dist(4)-1)^2+0.025*(dist(5)-
1)^2+0.018*(dist(6)-1)^2+0.010*(dist(7)-1)^2+0.031*(dist(8)-2)^2))+1.69*exp(-
(0.0063*(dist(1)-1)^2+0.027*(dist(2)-1)^2+0.0034*(dist(3)-1)^2+0.02*(dist(4)-
1)^2+0.025*(dist(5)-2)^2+0.01*(dist(6)-1)^2+0.010*(dist(7)-2)^2+0.031*(dist(8)-
2)^2))+(-5.73)*exp(-(0.00634*(dist(1)-1)^2+0.027*(dist(2)-2)^2+0.0034*(dist(3)-
1)^2+0.024*dist(4)^2+0.025*(dist(5)-1)^2+0.018*(dist(6)-2)^2+0.010*(dist(7)-
1)^2+0.031*(dist(8)-1)^2))+(-3.20)*exp(-(0.0063*dist(1)^2+0.027*(dist(2)-
1)^2+0.0034*(dist(3)-1)^2+0.024*dist(4)^2+0.025*(dist(5)-
1)^2+0.018*dist(6)^2+0.010*(dist(7)-2)^2+0.031*(dist(8)-1)^2))+2.12*exp(-
(0.0063*(dist(1)-1)^2+0.027*(dist(2)-2)^2+0.0034*(dist(3)-2)^2+0.024*(dist(4)-
1)^2+0.025*(dist(5)-1)^2+0.0189*(dist(6)-2)^2+0.010*(dist(7)-1)^2+0.031*(dist(8)-
2)^2))+(-1.086)*exp(-(0.0063*dist(1)^2+0.027*(dist(2)-1)^2+0.0034*(dist(3)-
1)^2+0.024*dist(4)^2+0.025*(dist(5)-2)^2+0.018*(dist(6)-
1)^2+0.010*dist(7)^2+0.031*(dist(8)-2)^2))+6.64*exp(-(0.0063*(dist(1)-
1)^2+0.027*(dist(2)-2)^2+0.0034*dist(3)^2+0.024*dist(4)^2+0.025*(dist(5)-
1)^2+0.018*(dist(6)-2)^2+0.010*(dist(7)-1)^2+0.031*dist(8)^2))+4.50*exp(-
(0.0063*dist(1)^2+0.0270*(dist(2)-1)^2+0.0034*(dist(3)-
1)^2+0.024*dist(4)^2+0.025*dist(5)^2+0.018*dist(6)^2+0.010*(dist(7)-
1)^2+0.031*(dist(8)-1)^2)))

Equation C.9 Phase III Response Prediction Formula for Average Total Cost

-(10.56+(0.273*exp(-(0.0560*(dist(1)-2)^2+0.058*dist(2)^2+0.05*(dist(3)-
1)^2+0.057*dist(4)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-1)^2+0.05*(dist(7)-
1)^2+0.053*(dist(8)-1)^2))+0.25*exp(-(0.056*(dist(1)-2)^2+0.058*(dist(2)-
2)^2+0.052*dist(3)^2+0.057*(dist(4)-1)^2+0.051*(dist(5)-
1)^2+0.0498*dist(6)^2+0.051*(dist(7)-1)^2+0.053*(dist(8)-2)^2))+1.40*exp(-
(0.056*(dist(1)-2)^2+0.058*(dist(2)-1)^2+0.052*(dist(3)-
2)^2+0.0573*dist(4)^2+0.05*dist(5)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-
2)^2+0.053*(dist(8)-1)^2))+(-0.2067)*exp(-(0.056*(dist(1)-1)^2+0.058*(dist(2)-
2)^2+0.052*(dist(3)-2)^2+0.057*(dist(4)-1)^2+0.051*(dist(5)-
2)^2+0.049*dist(6)^2+0.051*dist(7)^2+0.053*dist(8)^2))+(-1.576)*exp(-(0.056*(dist(1)-
2)^2+0.0586*dist(2)^2+0.052*(dist(3)-1)^2+0.057*dist(4)^2+0.051*(dist(5)-
1)^2+0.049*(dist(6)-1)^2+0.051*dist(7)^2+0.0532843606093131*(dist(8)-

1)^2))+1.087*exp(-(0.0560*(dist(1)-2)^2+0.058*(dist(2)-2)^2+0.0529*(dist(3)-
1)^2+0.057*(dist(4)-1)^2+0.051*(dist(5)-1)^2+0.049*dist(6)^2+0.051*(dist(7)-
1)^2+0.053*(dist(8)-2)^2))+(-0.94)*exp(-(0.056*(dist(1)-1)^2+0.058*(dist(2)-
1)^2+0.052*(dist(3)-2)^2+0.057*(dist(4)-1)^2+0.05*dist(5)^2+0.049*(dist(6)-
1)^2+0.051*dist(7)^2+0.053*(dist(8)-1)^2))+1.34*exp(-(0.056*(dist(1)-
1)^2+0.058*(dist(2)-1)^2+0.052*(dist(3)-2)^2+0.0573961760206857*(dist(4)-
1)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-2)^2+0.053*dist(8)^2))+(-
0.71)*exp(-(0.056*(dist(1)-1)^2+0.058*(dist(2)-1)^2+0.052*dist(3)^2+0.057*(dist(4)-
1)^2+0.051*(dist(5)-1)^2+0.049*(dist(6)-1)^2+0.05*(dist(7)-1)^2+0.053*dist(8)^2))+(-
1.97)*exp(-(0.056*(dist(1)-2)^2+0.058*(dist(2)-1)^2+0.052*(dist(3)-
1)^2+0.057*(dist(4)-1)^2+0.051*dist(5)^2+0.049*(dist(6)-
1)^2+0.051*dist(7)^2+0.0532*dist(8)^2))+(-0.53)*exp(-(0.056*(dist(1)-
1)^2+0.058*dist(2)^2+0.052*(dist(3)-2)^2+0.0573961760206857*(dist(4)-
2)^2+0.051*(dist(5)-1)^2+0.049*dist(6)^2+0.051*(dist(7)-1)^2+0.053*(dist(8)-
1)^2))+2.73*exp(-(0.056*(dist(1)-2)^2+0.058*(dist(2)-1)^2+0.0529*(dist(3)-
1)^2+0.057*(dist(4)-2)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-
2)^2+0.051*dist(7)^2+0.053*(dist(8)-1)^2))+(-3.95)*exp(-(0.056*(dist(1)-
1)^2+0.058*dist(2)^2+0.052*dist(3)^2+0.057*(dist(4)-1)^2+0.051*(dist(5)-
1)^2+0.049*dist(6)^2+0.051*(dist(7)-1)^2+0.053*dist(8)^2))+0.400*exp(-
(0.056*(dist(1)-2)^2+0.058*(dist(2)-1)^2+0.052*(dist(3)-1)^2+0.057*(dist(4)-
2)^2+0.051*dist(5)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-
2)^2+0.053*dist(8)^2))+1.076*exp(-(0.056*(dist(1)-
1)^2+0.058*dist(2)^2+0.052*(dist(3)-2)^2+0.0573961760206857*(dist(4)-
2)^2+0.051*(dist(5)-1)^2+0.049*dist(6)^2+0.051*(dist(7)-1)^2+0.053*(dist(8)-
2)^2))+4.62*exp(-(0.056*(dist(1)-2)^2+0.0586531559842805*(dist(2)-
1)^2+0.052*(dist(3)-1)^2+0.057*(dist(4)-2)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-
2)^2+0.051*(dist(7)-2)^2+0.053*(dist(8)-1)^2))+(-0.77)*exp(-(0.056*(dist(1)-
1)^2+0.058*(dist(2)-1)^2+0.0529924836217211*(dist(3)-1)^2+0.057*(dist(4)-
1)^2+0.051*(dist(5)-1)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-1)^2+0.053*(dist(8)-
1)^2))+(-0.46)*exp(-(0.056*dist(1)^2+0.058*(dist(2)-2)^2+0.052*(dist(3)-
1)^2+0.057*(dist(4)-2)^2+0.051*dist(5)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-
1)^2+0.053*(dist(8)-1)^2))+(-0.99)*exp(-
(0.056*dist(1)^2+0.058*dist(2)^2+0.052*(dist(3)-2)^2+0.057*(dist(4)-
1)^2+0.051*(dist(5)-1)^2+0.049*(dist(6)-2)^2+0.051*(dist(7)-1)^2+0.053*dist(8)^2))+(-
1.561)*exp(-(0.056*dist(1)^2+0.058*(dist(2)-1)^2+0.052*dist(3)^2+0.057*(dist(4)-
2)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-1)^2+0.051*dist(7)^2+0.0533*(dist(8)-

```
1)^2))+0.14*exp(-(0.056*(dist(1)-1)^2+0.058*dist(2)^2+0.052*dist(3)^2+0.057*(dist(4)-
1)^2+0.051*dist(5)^2+0.049*(dist(6)-2)^2+0.051*(dist(7)-2)^2+0.053*(dist(8)-
2)^2))+1.63*exp(-(0.056*dist(1)^2+0.058*(dist(2)-2)^2+0.052*(dist(3)-
1)^2+0.057*(dist(4)-2)^2+0.051*(dist(5)-1)^2+0.0498927610443326*(dist(6)-
1)^2+0.051*(dist(7)-2)^2+0.053*(dist(8)-1)^2))+(-0.49)*exp(-
(0.056*dist(1)^2+0.058*dist(2)^2+0.052*(dist(3)-1)^2+0.057*(dist(4)-
1)^2+0.051*(dist(5)-1)^2+0.049*(dist(6)-2)^2+0.051*(dist(7)-1)^2+0.053*(dist(8)-
1)^2))+0.40*exp(-(0.0560*(dist(1)-1)^2+0.058*(dist(2)-
1)^2+0.0529924836217211*dist(3)^2+0.057*(dist(4)-2)^2+0.051*(dist(5)-
2)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-2)^2+0.053*(dist(8)-1)^2))+(-1.364)*exp(-
(0.056*(dist(1)-1)^2+0.058*(dist(2)-1)^2+0.052*dist(3)^2+0.057*(dist(4)-
1)^2+0.051*dist(5)^2+0.049*(dist(6)-2)^2+0.051*dist(7)^2+0.053*(dist(8)-
2)^2))+1.266*exp(-(0.056*(dist(1)-1)^2+0.058*(dist(2)-2)^2+0.052*(dist(3)-
2)^2+0.057*(dist(4)-1)^2+0.051*(dist(5)-1)^2+0.049*(dist(6)-1)^2+0.051*(dist(7)-
1)^2+0.053*(dist(8)-2)^2))+1.30*exp(-(0.056*(dist(1)-1)^2+0.058*(dist(2)-
1)^2+0.052*(dist(3)-1)^2+0.057*(dist(4)-1)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-
1)^2+0.051*(dist(7)-2)^2+0.053*(dist(8)-2)^2))+0.56*exp(-(0.056*(dist(1)-
1)^2+0.058*(dist(2)-2)^2+0.052*(dist(3)-1)^2+0.057*dist(4)^2+0.0515*(dist(5)-
1)^2+0.049*(dist(6)-2)^2+0.051*(dist(7)-1)^2+0.0532*(dist(8)-1)^2))+(-1.46)*exp(-
(0.056*dist(1)^2+0.0585*(dist(2)-1)^2+0.052*(dist(3)-
1)^2+0.057*dist(4)^2+0.051*(dist(5)-1)^2+0.049*dist(6)^2+0.051*(dist(7)-
2)^2+0.053*(dist(8)-1)^2))+4.078*exp(-(0.056*(dist(1)-1)^2+0.0586*(dist(2)-
2)^2+0.052*(dist(3)-2)^2+0.0573*(dist(4)-1)^2+0.051*(dist(5)-1)^2+0.049*(dist(6)-
2)^2+0.051*(dist(7)-1)^2+0.05*(dist(8)-2)^2))+(-0.687)*exp(-
(0.0560*dist(1)^2+0.0585*(dist(2)-1)^2+0.052*(dist(3)-
1)^2+0.0573*dist(4)^2+0.051*(dist(5)-2)^2+0.049*(dist(6)-
1)^2+0.051*dist(7)^2+0.053*(dist(8)-2)^2))+(-1.087)*exp(-(0.056*(dist(1)-
1)^2+0.058*(dist(2)-2)^2+0.052*dist(3)^2+0.057*dist(4)^2+0.051*(dist(5)-
1)^2+0.049*(dist(6)-2)^2+0.051*(dist(7)-1)^2+0.053*dist(8)^2))+(-3.80)*exp(-
(0.056*dist(1)^2+0.058*(dist(2)-1)^2+0.0529924836217211*(dist(3)-
1)^2+0.057*dist(4)^2+0.051*dist(5)^2+0.049*dist(6)^2+0.051*(dist(7)-
1)^2+0.053*(dist(8)-1)^2))))
```

Appendix D: Police Staffing Study Gaussian Process Model Reports

Table D.1 Phase I Gaussian Process Model – Avg Response Time

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.2474553	0.2678075	0.1580786	.	0.0011563	0.0145371	0.0045131	0.0336623	0	0.0005543	0.0553058
2	0.0449372	0.060778	0.044368	0.0011563	.	0.0013614	0.0002153	0.0019047	0	2.9089e-5	0.0117432
3	0.1062285	0.088186	0.0619342	0.0145371	0.0013614	.	0.0007394	0.0032303	0	7.6516e-5	0.0063073
4	0.0308951	0.0084685	0.0001402	0.0045131	0.0002153	0.0007394	.	0.0014698	0	8.875e-6	0.0013819
5	0.2276556	0.5314223	0.3771221	0.0336623	0.0019047	0.0032303	0.0014698	.	0	0.0003496	0.1136834
6	0	0	0	0	0	0	0	0	.	0	0
7	0.0075445	0.0094404	0.0080075	0.0005543	2.9089e-5	7.6516e-5	8.875e-6	0.0003496	0	.	0.0004145
8	0.3159127	0.2432875	0.0544514	0.0553058	0.0117432	0.0063073	0.0013819	0.1136834	0	0.0004145	.

Mu	Sigma	Nugget
167.31207	259.70467	1e-12

-2*LogLikelihood
255.2918

Table D.2 Phase I Gaussian Process Model – Avg % Time Occupied

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.2464654	0.2637601	0.1487396	.	0.0010979	0.0093035	0.0026265	0.0424405	0	0.0003888	0.0591631
2	0.0462478	0.0540784	0.0364442	0.0010979	.	0.0010228	7.6226e-5	0.0019908	0	1.7644e-5	0.0134288
3	0.0876936	0.0578791	0.0412751	0.0093035	0.0010228	.	0.0003007	0.0026373	0	3.7743e-5	0.0033019
4	0.0193537	0.0047851	6.6337e-5	0.0026265	7.6226e-5	0.0003007	.	0.0009345	0	2.6422e-6	0.0007782
5	0.2652582	0.5770732	0.398546	0.0424405	0.0019908	0.0026373	0.0009345	.	0	0.0002414	0.1302827
6	0	0	0	0	0	0	0	0	.	0	0
7	0.0055265	0.0061764	0.0052391	0.0003888	1.7644e-5	3.7743e-5	2.6422e-6	0.0002414	0	.	0.0002491
8	0.3445177	0.2616691	0.0544653	0.0591631	0.0134288	0.0033019	0.0007782	0.1302827	0	0.0002491	.

Mu	Sigma	Nugget
0.8830547	0.0008664	1e-12

-2*LogLikelihood
-162.7704

Table D.3 Phase I Gaussian Process Model – Avg Total Cost

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0	0	0	.	0	0	0	0	0	0	0
2	0	0	0	0	.	0	0	0	0	0	0
3	0	0	0	0	0	.	0	0	0	0	0
4	4.3870694	0.2729381	0.056317	0	0	0	.	0.1015246	0	0.0076666	0.10743
5	6.7998753	0.2528568	0.0245656	0	0	0	0.1015246	.	0	0.0127627	0.114004
6	0	0	0	0	0	0	0	0	.	0	0
7	1.319241	0.045387	0.0113286	0	0	0	0.0076666	0.0127627	0	.	0.0136292
8	4.2367848	0.2671126	0.0320494	0	0	0	0.10743	0.114004	0	0.0136292	.

Mu	Sigma
20.214111	3.428e+14

-2*LogLikelihood
1161.3593

Table D.4 Phase II Model Report Gaussian Process Model – Avg Response Time

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.1552244	0.3102872	0.2315202	.	0	0.0302616	0	0.0011614	0.0470108	0.0002512	0.0000821
2	0	0	0	0	.	0	0	0	0	0	0
3	0.2566099	0.450578	0.3881008	0.0302616	0	.	0	0.0088272	0.0229428	0.0002472	0.0001984
4	0	0	0	0	0	0	.	0	0	0	0
5	0.0531772	0.0956186	0.0813823	0.0011614	0	0.0088272	0	.	0.0041719	6.5822e-5	9.9496e-6
6	0.5625357	0.2292305	0.154026	0.0470108	0	0.0229428	0	0.0041719	.	0.0008577	0.0002214
7	0.0138819	0.0190633	0.0176365	0.0002512	0	0.0002472	0	6.5822e-5	0.0008577	.	4.9874e-6
8	0.0096724	0.0058861	0.0053693	0.0000821	0	0.0001984	0	9.9496e-6	0.0002214	4.9874e-6	.

Mu	Sigma	Nugget
186.21791	164.5206	0.1

-2*LogLikelihood
240.40096

Nugget parameter set to avoid singular variance matrix.

Table D.5 Phase II Model Report Gaussian Process Model – Avg % Time Occupied

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0	0	0	.	0	0	0	0	0	0	0
2	73.866323	0.155982	0.0072396	0	.	0	2.2844e-7	0.0919258	0	0.0550339	0.0017824
3	0	0	0	0	0	.	0	0	0	0	0
4	0.0001542	6.0674e-7	2.6978e-8	0	2.2844e-7	0	.	2.1773e-7	0	1.3292e-7	6.696e-10
5	73.059701	0.1809208	0.0154292	0	0.0919258	0	2.1773e-7	.	0	0.0718546	0.0017109
6	0	0	0	0	0	0	0	0	.	0	0
7	32.299709	0.1348601	0.0068731	0	0.0550339	0	1.3292e-7	0.0718546	0	.	0.0010983
8	0.1720044	0.0047405	0.0001489	0	0.0017824	0	6.696e-10	0.0017109	0	0.0010983	.

Mu	Sigma	Nugget
0.9081463	0.0068128	1e-8

-2*LogLikelihood

-84.67838

Nugget parameter set to avoid singular variance matrix.

Table D.6 Phase II Model Report Gaussian Process Model – Avg Total Cost

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.0185656	0.0979944	0.0667447	.	0.0032838	0.0000189	2.1349e-5	0.0013886	2.1731e-5	0.0002357	0.0262797
2	0.1137389	0.1392856	0.1083888	0.0032838	.	0.0005706	0.0004679	0.0123459	0.0026345	0.001302	0.0102922
3	0.0103144	0.0861834	0.0832586	0.0000189	0.0005706	.	1.9622e-5	0.0001388	1.8256e-5	1.2024e-5	0.0021466
4	0.0085382	0.0954511	0.0924081	2.1349e-5	0.0004679	1.9622e-5	.	0.0004612	4.6277e-5	1.7589e-5	0.002009
5	0.0622056	0.155037	0.0430341	0.0013886	0.0123459	0.0001388	0.0004612	.	0.000171	0.0002304	0.0972671
6	0.0118491	0.129143	0.1215786	2.1731e-5	0.0026345	1.8256e-5	4.6277e-5	0.000171	.	1.144e-5	0.0046613
7	0.0090763	0.1018547	0.0940511	0.0002357	0.001302	1.2024e-5	1.7589e-5	0.0002304	1.144e-5	.	0.0059945
8	0.165703	0.3592655	0.2106151	0.0262797	0.0102922	0.0021466	0.002009	0.0972671	0.0046613	0.0059945	.

Mu	Sigma	Nugget
13.249123	31.276031	0.0001

-2*LogLikelihood

143.25568

Nugget parameter set to avoid singular variance matrix.

Table D.7 Phase III Model Report Gaussian Process Model – Avg Response Time

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.010966	0.0969118	0.0937918	.	0.00076	4.5248e-5	0.0002161	0.0000365	0.0001232	0.000741	0.001198
2	0.0353113	0.1536039	0.1308526	0.00076	.	0.0003318	0.0078404	0.0028099	0.0058489	0.0001345	0.0050257
3	0.0049039	0.141841	0.1413115	4.5248e-5	0.0003318	.	6.2784e-5	2.0478e-5	3.0647e-5	1.7924e-5	2.0647e-5
4	0.0259506	0.1358656	0.1142491	0.0002161	0.0078404	6.2784e-5	.	0.0065117	0.0017774	0.0020743	0.0031338
5	0.0324943	0.1993426	0.1855804	0.0000365	0.0028099	2.0478e-5	0.0065117	.	0.0015788	0.0015148	0.00129
6	0.0240451	0.0859408	0.0743885	0.0001232	0.0058489	3.0647e-5	0.0017774	0.0015788	.	0.0021423	5.096e-5
7	0.0107589	0.1062049	0.0995186	0.000741	0.0001345	1.7924e-5	0.0020743	0.0015148	0.0021423	.	6.1385e-5
8	0.0360617	0.1254685	0.114688	0.001198	0.0050257	2.0647e-5	0.0031338	0.00129	5.096e-5	6.1385e-5	.

Mu	Sigma	Nugget
326.80689	5190.1337	1e-12

-2*LogLikelihood

253.58194

Table D.8 Phase III Model Report Gaussian Process Model – Avg % Time Occupied

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.0063427	0.0946855	0.0928848	.	0.0003792	1.0771e-5	0.0001164	1.7239e-5	0.00013	0.0006406	0.0005065
2	0.0270652	0.1633556	0.1417161	0.0003792	.	0.0002326	0.0071609	0.0012211	0.0058377	1.2617e-5	0.0067955
3	0.0034521	0.1344349	0.1340848	1.0771e-5	0.0002326	.	3.0033e-5	0.0000132	3.989e-6	2.5854e-5	0.0000337
4	0.0246764	0.1397561	0.1160969	0.0001164	0.0071609	3.0033e-5	.	0.0063922	0.0025735	0.0022525	0.0051337
5	0.0256975	0.1936698	0.1823729	1.7239e-5	0.0012211	0.0000132	0.0063922	.	0.0012565	0.0016442	0.0007525
6	0.0189866	0.0863736	0.073674	0.00013	0.0058377	3.989e-6	0.0025735	0.0012565	.	0.0028318	6.6072e-5
7	0.0102126	0.113661	0.1061884	0.0006406	1.2617e-5	2.5854e-5	0.0022525	0.0016442	0.0028318	.	6.5139e-5
8	0.0315163	0.1200368	0.1066837	0.0005065	0.0067955	0.0000337	0.0051337	0.0007525	6.6072e-5	6.5139e-5	.

Mu	Sigma	Nugget
1.2373163	0.0242258	1e-8

-2*LogLikelihood

-163.2284

Nugget parameter set to avoid singular variance matrix.

Table D.9 Phase III Model Report Gaussian Process Model – Avg % OCCT

Model Report

Column	Theta	Total Sensitivity	Main Effect	1 Interaction	2 Interaction	3 Interaction	4 Interaction	5 Interaction	6 Interaction	7 Interaction	8 Interaction
1	0.0560811	0.1344391	0.1338668	.	7.1837e-5	6.3153e-5	6.9666e-5	0.0001183	6.0742e-5	0.0001077	0.0000809
2	0.0586532	0.1253209	0.1248481	7.1837e-5	.	7.4466e-5	0.0000742	6.5438e-5	5.8812e-5	0.0000632	6.4842e-5
3	0.0529925	0.1163298	0.1159087	6.3153e-5	7.4466e-5	.	6.7123e-5	0.0000557	5.4269e-5	5.3684e-5	5.2721e-5
4	0.0573962	0.1364912	0.1359697	6.9666e-5	0.0000742	6.7123e-5	.	9.3667e-5	6.0982e-5	8.8671e-5	6.7152e-5
5	0.0515539	0.1194078	0.1188831	0.0001183	6.5438e-5	0.0000557	9.3667e-5	.	8.3926e-5	5.0071e-5	5.7562e-5
6	0.0498928	0.1243352	0.1238713	6.0742e-5	5.8812e-5	5.4269e-5	6.0982e-5	8.3926e-5	.	8.7944e-5	5.723e-5
7	0.0514649	0.1239181	0.1234104	0.0001077	0.0000632	5.3684e-5	8.8671e-5	5.0071e-5	8.7944e-5	.	5.6367e-5
8	0.0532844	0.1217083	0.1212715	0.0000809	6.4842e-5	5.2721e-5	6.7152e-5	5.7562e-5	5.723e-5	5.6367e-5	.

Mu	Sigma	Nugget
10.561996	4.0318663	0.1

-2*LogLikelihood

101.73543

Nugget parameter set to avoid singular variance matrix.