Kriging as an alternative to polynomial regression in response surface analysis

by

Dean Ronald De Cock

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-Majors: Statistics; Industrial Engineering

Program of Study Committee:
Max Morris, Co-major Professor
Sigurdur Olafsson, Co-major Professor
Philip Dixon
William Meeker
Stephen Vardeman

Iowa State University

Ames, Iowa

2003

UMI Number: 3118221

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3118221

Copyright 2004 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Dean Ronald De Cock

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Response surface methodology (RSM) is a statistical tool sometimes used in the modeling and optimization of processes (Meyers and Montgomery, 1995). RSM relies upon empirical approximations to the true process relationships between factors or independent variables, and response variables, with first and second-order polynomial models being the most popular approximating functions. RSM is usually used in a controlled experimental context with two popular plans for second-order RSM being the Central Composite (Box and Wilson, 1951) and the Box-Behnken (Box and Behnken, 1960) designs. Estimates of the parameters in the first and second-order models are usually found via ordinary least squares (OLS) methods. Process optimization can be executed for first-order models by using iterative methods such as "the path of steepest ascent" or for second-order models by taking derivatives of the fitted model, setting them equal to zero, and solving for the factor values associated with the maximum or minimum predicted response.

As an example, a manufacturer of computer screens is interested in modeling the durability of the glare coating on their products. The durability can be directly measured by rubbing a cheese cloth across the surface under a known load until the coating is removed. The process engineer responsible for the product knows that durability is directly related to two process inputs, chemical concentration of polymer in the coating and oven bake time. Increased knowledge of the relationship between these inputs and the durability will allow manufacturing the ability to adjust the process to achieve desired quality levels.

Traditional RSM can be used to approximate these relationships in the manufacturing process. To study the process, durability levels are measured at various process input settings that have been determined via structured data collection plans. These data are then used in conjunction with regression modeling techniques to approximate the surface and predict durability at specified operating conditions. Adequate modeling of the surface is important as

a poorly fitted model may result in recommended process settings that yield sub-standard products. These poor approximations to the process relationships can commonly result from overly simplistic assumptions of model form.

One common difficulty in RSM is the lack of fit (LOF) of a specified model. This LOF can for example, come from using a polynomial model which does not include enough terms to accurately represent the true relationship. An under-specified model results in biased estimates of the parameters, which lead to biased estimates of the factor values associated with the process maximum. One solution to this problem is the use of Box-Cox transformations (Box and Cox, 1964) of the data. While these transformations are commonly used, they can lead to unusual measurement units and problems of heteroscedasticity.

We shall consider kriging as an alternative to polynomial regression in RSM. Kriging is a spatial statistics method, first developed for applications in the mining industry (e.g. Huijbregts and Matheron, 1971) where such methods are commonly referred to as geostatistical methods. A basic assumption in the spatial applications of kriging is that the response values collected at points that are relatively closer together will tend to be more similar than the response values collected at points that are farther apart. Following this idea, kriging uses a weighted average of the responses at local data points to predict the response in a nearby area where no data have been collected. In the case of process optimization, "points that are close together" will be design points for which experimental conditions are defined by similar factor levels. In geostatistical applications, kriging data are often collected in a much less structured manner than is typical in RSM experiments.

Kriging is a flexible prediction method that accommodates a wide variety of functional relationships. The simplest form of the method, "ordinary kriging", is based on the assumption that the function of interest is a constant plus the realization of a spatially correlated random deviation from this value. In comparison, the corresponding simplest regression or RSM method is based on the assumption that the function of interest is a

constant, approximated only by an intercept term. In both kriging and regression, deviations between observed values and values of the function of interest are modeled as uncorrelated homogenous random errors. To predict responses at locations where data are not collected, kriging uses a weighted average of the data observed at other locations. These weights are based on the variogram function, which relates the covariance between responses to the distance between their corresponding design points. Given a known variogram, the kriging predictions are the best linear unbiased predictions (BLUP) of the response at any point, under the model. While prediction under the kriging method requires the continued use of the data, predictors within RSM are of different form, relying only on the fitted polynomial.

A generalization of ordinary kriging, universal kriging, substitutes a polynomial model for the constant mean used in ordinary kriging. This substitution allows for adaptation to a simple trend in the data (first-order linear model) or more complex relationships (second- or higher-order models). The assumed universal kriging model differs from the polynomial regression model in that a spatially correlated random model element is employed. The kriging predictions are based on both the fitted polynomial and interpolated random process, while regression predictions are based only on the fitted polynomial.

The use of kriging to approximate unknown or complex functions in applications other than geostatistics is not new. Kriging has been used successfully in computer experiments (e.g. Sacks et al., 1989) to approximate the output from complicated computer programs. Computer experiments differ from traditional experiments in that deterministic computer programs yield the same response value for a given set of inputs, every time the code is run. There is no random physical variability or measurement error in the output. Any discrepancy between the output of a computer model and the prediction of a statistical model is due to bias or lack of fit resulting from the mis-specification of the statistical model.

As noted above, the kriging method can be adapted to accommodate data that include random measurement error. In geostatistics, the specification of the variogram (weighting

factor) allows for a parameter called the "nugget". This parameter allows for a random difference between two samples taken from the same location. In mining applications such a difference would occur if one sample had a nugget of ore and the other did not. From a RSM standpoint, the nugget effect can be viewed as measurement error, the variance of which can be estimated from multiple response values taken at the same combination of factor levels.

Our application of kriging will differ significantly from its use in computer experiments, due to the assumption that measurements include random error that is not part of the response function of interest. It will deviate from traditional RSM in that, in additional to the random measurement errors present in the measured response taken at any fixed combination of factor values, the model also includes a realization of a spatially correlated random term as part of the response function of interest.

The goal of this study is to investigate whether kriging methods can lead to better predictions of near quadratic relationships than traditional regression methods. This analysis will begin by offering theoretical evidence that kriging methods have the potential to outperform standard RSM under certain conditions and will conclude by reviewing the performance of various methods for several simulated data sets.

CHAPTER 2 METHODS

In this chapter we will examine three prediction methods: Ordinary Least Squares (OLS), Ordinary Kriging (OK), and Universal Kriging (UK). Each of the methods will be used to approximate a surface over a region of interest defined as the design space, *D*. This region will be defined as a *d*-dimensional cube representing *d* independent variables that have each been scaled from −1 to 1. Any point within *D* represents a possible combination of factor values, and can be specified by in its coordinates:

$$x = [x_1 \quad \dots \quad x_d] \quad -1 \le x_i \le 1.$$

The n specific values of x for which responses are observed in the space are denoted by $s_1 \cdots s_n$.

In each of the methods, the predictor for the response associated with any given value of x, $\hat{y}(x)$, is a linear combination of the responses observed in the experiment. This linear combination will be defined by an n-element vector of weights, c_x', referred to as the prediction vector:

$$\hat{y}(x) = c_x' Y \qquad \text{for} \qquad Y = \begin{bmatrix} y(s_1) \\ \vdots \\ y(s_n) \end{bmatrix}.$$

The form of these vectors will be derived for each of the three methods in this chapter. Following the derivation, a numerical example will be analyzed to show the potential superiority of the kriging methods.

2.1 Ordinary Least Squares

In traditional RSM, it is typically assumed that the function to be modeled can be adequately approximated by a second-order polynomial model in the region of interest. The model usually contains two components, one representing the polynomial surface η and another representing random error from the measurement process:

$$y(x) = \eta + \varepsilon = \beta_o + \sum_{i=1}^{d} \beta_i x_i + \sum_{i=1}^{d} \beta_{ii} x_i^2 + \sum_{i<j=1}^{d} \beta_{ij} x_i x_j + \varepsilon \qquad (2.1)$$

where

$\varepsilon \sim N(0, \sigma_\varepsilon^2)$ normally distributed error.

The n design points will be represented in a model matrix constructed from the x_{ij} values, where $s_i' = [x_{i1} \quad \cdots \quad x_{id}]$, i.e.

i indicates the design point $\qquad 1 \le i \le n$

j indicates the independent variable $\qquad 1 \le j \le d$.

The model matrix X for the quadratic polynomial model will be composed of four parts representing the intercept, linear, quadratic, and interaction components of the model.

$$X = [\mathbf{i} \mid L \mid Q \mid M]$$

where

$$\mathbf{i} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad L = \begin{bmatrix} x_{11} & .. & x_{1d} \\ \vdots & \vdots & \vdots \\ x_{n1} & .. & x_{nd} \end{bmatrix} \quad Q = \begin{bmatrix} x_{11}^2 & .. & x_{1d}^2 \\ \vdots & \vdots & \vdots \\ x_{n1}^2 & .. & x_{nd}^2 \end{bmatrix} \quad \text{and}$$

$$M = \begin{bmatrix} x_{11}x_{12} & .. & x_{11}x_{1d} & x_{12}x_{13} & .. & x_{1(d-1)}x_{1d} \\ : & : & : & : & : & : \\ x_{n1}x_{n2} & .. & x_{n1}x_{nd} & x_{n2}x_{n3} & .. & x_{n(d-1)}x_{nd} \end{bmatrix}.$$

Using this matrix notation, equation (2.1) can be written for all experimental runs simultaneously as

$$Y = X\beta + \varepsilon$$

where

$$\beta = \begin{bmatrix} \beta_o \\ \beta_1 \\ : \\ \beta_{(n-1)n} \end{bmatrix}.$$

and ε now represents an n-vector of homogeneous, independent errors.

When X is of full column rank, the vector of least-squares estimators (b), for the model coefficients (β), is based on the measured responses Y, and the model matrix,

$$b = (X'X)^{-1}XY.$$

As OLS methods produce the best linear estimators under these assumptions, they are traditionally used to find estimators for the parameters β. Prediction of the response at any location is a function of the values of the independent variable(s) at the location x. Let

$$a' = \begin{bmatrix} 1 & l & q & m \end{bmatrix}$$

where

$$l = \begin{bmatrix} x_1 & .. & x_d \end{bmatrix} \quad q = \begin{bmatrix} x_1^2 & .. & x_d^2 \end{bmatrix} \quad m = \begin{bmatrix} x_1x_2 & .. & x_1x_d & x_2x_3 & .. & x_{(d-1)}x_d \end{bmatrix}$$

then the OLS estimate of the expected response at x is :

$$\hat{y}(x) = a'b = a'(X'X)^{-1}X'Y .$$

Hence for OLS the form of the prediction vector is

$$c_x' = a'(X'X)^{-1}X' . \tag{2.2}$$

2.1.1 Response Transformation

One common method to address lack of fit in quadratic models is to transform the response variable y into a new variable y_{new} through a power transformation (Box and Cox, 1964):

$$y_{new} = y^{\lambda} \tag{2.3}$$

where λ can take on any real value with $\lambda \to 0$ representing the log transformation. Under the assumption that elements of ε are normally distributed, maximum likelihood methods can be used to estimate λ. A scale-adjusted form of the power transformation (y^*) is used for this estimation as direct use of y_{new} tends to result in larger estimated λ's having larger errors regardless of the quality of fit:

$$y^* = \left(\prod_{i=1}^{n} y_i\right)^{1/n} \ln y \quad \text{if } \lambda = 0 \tag{2.4}$$

$$y^* = \frac{y^{\lambda} - 1}{\lambda\left(\prod_{i=1}^{n} y_i\right)^{(\lambda-1)/n}} \quad \text{if } \lambda \neq 0 .$$

The optimal value for λ is found by fitting the model, calculating the residual sum of squares for the transformed responses, and choosing the λ that minimizes this sum of squares. For this analysis it is assumed that the surfaces are approximately quadratic, so only relatively mild transformations $(0 \le \lambda \le 2)$ of the data were considered.

Estimators and predictors for the transformed variables are of the same form as those found using the OLS techniques outlined in 2.1, with the substitution of Y^* for Y. The use of the transformation technique in this analysis will be referred to as the Box-Cox (BC) method.

It should be noted that the Box-Cox methodology cannot be used without modification if negative values of the response are present in the data. To compensate for this deficiency, all data sets are consistently shifted into the positive range before the BC method is used. Following the determination of λ and the calculations for the predictors, the data are re-shifted back to the original range.

2.2 Kriging

Kriging is based on a model in which measurements are composed of the surface, modeled as a sum of two components, and an additional error term. The first component is a polynomial function, $g(x)$, which represents the "global" trend of the surface. The second component, $Z(x)$, is a realization of a stationary gaussian random process with mean zero and stationary non-negative covariance that represents the "local" deviations of the surface from the polynomial. These two components represent the true physical surface relating $E(y(x))$ to x. The final term is random error with zero mean and covariance, which corresponds exactly to the error assumed in the quadratic polynomial model. The general form of the kriging model is:

$$y(x) = g(x) + Z(x) + \varepsilon(x).$$

Note that without $Z(x)$, the model would be a polynomial regression model.

Two different kriging methods, ordinary kriging and universal kriging, will be investigated in this paper. The general equation defining the predictors is the same for both methods, with differences in the particular forms noted in Section 2.2.2. The ordinary kriging method uses the model assumption that the underlying polynomial, *g(x)*, is a constant and that the only variation in the true surface results from the realization of the gaussian process *Z*. The universal kriging method used for this analysis is based the model assumption of an underlying quadratic polynomial in addition to the gaussian process.

2.2.1 Gaussian Process *Z(x)*

The gaussian process $Z(x)$ is what differentiates the kriging model from the quadratic polynomial model. This process is characterized by a correlation function *R*, which specifies the correlation between the responses at any two points in the design space, $t = (t_1, t_2, .. t_d)$ and $u = (u_1, u_2, ... u_d)$, is a function of the scaled distance between the points. In this study we shall limit attention to the gaussian correlation function:

$$R(t,u) = \exp\left(-\sum_{k=1}^{d} \theta_k (t_k - u_k)^2\right)$$

for which realizations of $Z(x)$ are infinitely differentiable. Further, we shall consider only the isotropic form of this correlation, for which the scaling parameter is common to all factors:

$$R(t,u) = \exp\left(-\sum_{k=1}^{d} \theta (t_k - u_k)^2\right). \qquad (2.5)$$

We will define *Z* to be a stationary process with mean zero and variance σ_z^2, and so need only multiply *R* by σ_z^2 to define the covariance function:

$$Cov(Z(t), Z(u)) = \sigma_z^2 R(t,u)$$

2.2.2 Estimation and Prediction

The vector of polynomial coefficient estimators b used here are the best linear unbiased estimators (BLUE) under the assumed model. In addition to the vectors and matrices already defined, we will need the covariance matrix V of the responses Y, at the locations specified by the design. V is a function of the variance and covariance parameters of ε and Z. Using $R(S)$ to denote the n x n correlation matrix for Z at the design points, i.e. $\{R(S)\}_{ij} = R(s_i, s_j)$,

$$V = Cov(Y, Y') = \sigma_z^2 R(S) + \sigma_\varepsilon^2 I.$$

Additionally, the covariances between Z at any new x and the responses at the design points will be denoted by the vector v_x.

$$v_x' = [Cov(y(s_1), z(x)), \ldots, Cov(y(s_n), z(x))]$$

where

$$Cov(y(s_i), z(x)) = \sigma_z^2 R(s_i, x).$$

The BLUE for the vector of polynomial coefficients in the kriging model, β, is

$$b = (A' V^{-1} A)^{-1} A' V^{-1} Y$$

where A is a function of the design with

$A = \mathbf{1}$ for ordinary kriging

and

$A = X = [\mathbf{1} \quad L \quad Q \quad M]$ for universal kriging.

When the indicated matrix inverse exists, the BLUP of the response at x is

$$\hat{y}(x) = a'b + v_x' V^{-1}(Y - Ab)$$

where a is a function of the point of interest with

$$a' = 1 \qquad \text{for ordinary kriging}$$

and

$$a' = \begin{bmatrix} 1 & l & q & m \end{bmatrix} \qquad \text{for universal kriging.}$$

The fitted value is a result of the sum of the Generalized Least Squares (GLS) fit of the polynomial and a local adjustment related to the residuals between the fitted polynomial and the experimental data.

By substituting the respective values for a' and A for the two kriging methods we find that the form of the prediction vector for ordinary kriging is

$$c_x' = \left[\left(1_N{}' V^{-1} 1_N \right)^{-1} 1_N{}' + v_x' - v_x' V^{-1} 1_N \left(1_N{}' V^{-1} 1_N \right)^{-1} 1_N{}' \right] V^{-1} \tag{2.6}$$

and the form for universal kriging is

$$c_x' = \left[a' \left(X' V^{-1} X \right)^{-1} X' + v_x' - v_x' V^{-1} X \left(X' V^{-1} X \right)^{-1} X' \right] V^{-1}.$$

The above equations assume that the covariance matrix V is known, which requires that the experimenter know θ and the ratio of σ_z^2 and σ_ε^2. Estimation of these quantities will be discussed later, but a reparameterization of the covariance function will be introduced now to

facilitate later analysis. The covariance matrix V can be rewritten as a function of a different design correlation matrix $W(S)$ which characterizes both the Z and ε;

$$V = Cov(Y, Y') = \sigma_z^2 R(S) + \sigma_\varepsilon^2 I = \sigma^2 W(S)$$

where

$$W(t,u) = 1 \qquad t = u \qquad \forall t,u = (1,...,n),$$

$$W(t,u) = p\exp\left(-\sum_{k=1}^{d}\theta(t_k - u_k)^2\right) \quad t \neq u,$$

$$\sigma^2 = \sigma_z^2 + \sigma_\varepsilon^2,$$

$$p = \frac{\sigma_z^2}{\sigma^2} = \frac{\sigma_z^2}{\sigma_z^2 + \sigma_\varepsilon^2}.$$

The parameter p, is directly related to the geostatistical concept of a "nugget" effect, which allows for variation in assays taken at the same location. In our model, p represents the proportion of the total variation that can be attributed to the stationary gaussian process Z. Since the relationship between (σ, p) and $(\sigma_z, \sigma_\varepsilon)$ is one-to-one, we can easily transform estimates from either parameterization to the other.

2.3 Preliminary Method Comparisons

Integrated Mean Square Error (*IMSE*) is often used to compare the methods in RSM (Box and Draper, 1987, p. 428). *IMSE* is based on the Mean Square Error (*MSE*), which is one of the indices most commonly used in evaluating the fit of a model; minimization of *MSE* is the basis for many statistical estimators. The *MSE* is the expected squared difference between an

estimate or prediction and the value it is intended to approximate. In our context we are interested in

$$MSE[\hat{y}(x)] = E[\hat{y}(x) - \eta(x)]^2 . \tag{2.7}$$

The general form of the true surface addressed in this analysis will be a polynomial plus a realization of a gaussian process. The polynomial terms will be represented by the vector f for individual x locations and by the matrix F for the whole set of design points. Hence, for our purposes, f and F in the model are analogous to a and A in the kriging formula. The model can be written as

$$\eta(x) = f'\beta + Z(x)$$

where

$$f' = [1 \quad f_1(x) \quad f_2(x) \quad .. \quad f_k(x)]$$

is the a vector for the new point, and

$$F = \begin{bmatrix} f'(s_1) \\ : \\ f'(s_n) \end{bmatrix}$$

is the model matrix for the design points. The elements of f are functions of x and the relative importance of Z is determined by the value of σ_z^2.

Substituting the model matrix, F, and the prediction vector, c_x, into (2.7) and simplifying the equation results in the following expression for *MSE* as a function of x:

$$MSE[\hat{y}(x)] = E[\hat{y}(x) - \eta(x)]^2 = E[c_x'Y - \eta(x)]^2 = E[c_x'Y - Ec_x'Y + Ec_x'Y - \eta(x)]^2$$

$$= E(c_x'Y - Ec_x'Y)^2 + E(Ec_x'Y - \eta(x))^2 + 2E(c_x'Y - Ec_x'Y)(Ec_x'Y - \eta(x))$$

$$= Var(c_x'Y) + E(c_x'F\beta - \eta(x))^2 + 2E(c_x'Y - c_x'F\beta)(c_x'F\beta - \eta(x))$$

$$= [c_x'Vc_x] + [(c_x'F\beta - f'\beta)^2 + \sigma_z^2] + [-2c_x'v_x]$$

$$= c_x'Vc_x + (c_x'F\beta - f'\beta)^2 - 2c_x'v_x + \sigma_z^2 . \qquad (2.8)$$

This equation is valid for the ordinary least squares (OLS), universal kriging (UK), and ordinary kriging (OK) methods. As F, β, v_x and σ_z^2 are characteristics of the given surface and the design, only the prediction vector c_x is method specific. It should be noted that the equation is not valid for the least squares method incorporating the Box-Cox transformation (BC) as its fitted values are based on the transformed response y^* rather than the original y's.

Since the interest in the fit of the model encompasses the entire range of x, the *MSE* will be integrated over the entire experimental region, D. *IMSE* will stand for the integral of *MSE* over the entire region of interest. Using (2.8), this can be written directly as:

$$IMSE = \Omega \int_{x \in D} c_x'Vc_x + (c_x'F\beta - f'\beta)^2 - 2c_x'v + \sigma_z^2 dx \qquad (2.9)$$

where

$$\Omega^{-1} = \int_{x \in D} dx . \qquad D = (-1 \le x_i \le 1)\ \forall i = 1,...,d$$

As *IMSE* is based only on expectations, it is not a function of the observed data. In order to compute *IMSE* for method comparison, it is necessary to specify the surface of interest, the variances, and the experimental design.

2.3.1 Example 1

The *IMSE*'s for OLS, OK, and UK, will be compared for an example. The true surface will consist of a fourth-order polynomial in one variable without the addition of a stationary gaussian process (i.e. $\sigma_z^2 = 0$). Observed responses will be the true surface plus normally distributed measurement error generated as:

$$Y(x) = \eta(x) + \varepsilon = 5 + 2x - 2x^2 - 1.5x^3 + 0.4x^4 + \varepsilon$$

with

$$\beta = \begin{bmatrix} 5.0 \\ 2.0 \\ -2.0 \\ -1.5 \\ 0.4 \end{bmatrix} \quad V = \sigma_\varepsilon^2 I \quad \nu_x' = 0 \qquad \sigma_\varepsilon^2 = 0.1 \quad \sigma_z^2 = 0 .$$

Figure 1 illustrates the shape of the surface over the experimental region. The surface is approximately quadratic in shape, but does deviate slightly from a second-order surface.

The design *S* will consist of seven equally spaced points taken along the x-axis with no repeated measurements:

$$S \in [-1, \ -2/3, \ -1/3 \ \ 0, \ 1/3, \ 2/3, \ 1].$$

Because *Z* does not appear in the model, the *IMSE* can be written in the simplified form:

$$IMSE = \Omega \int_{x \in D} \sigma_\varepsilon^2 c_x' c_x + (c_x' F\beta - f'\beta)^2 dx$$

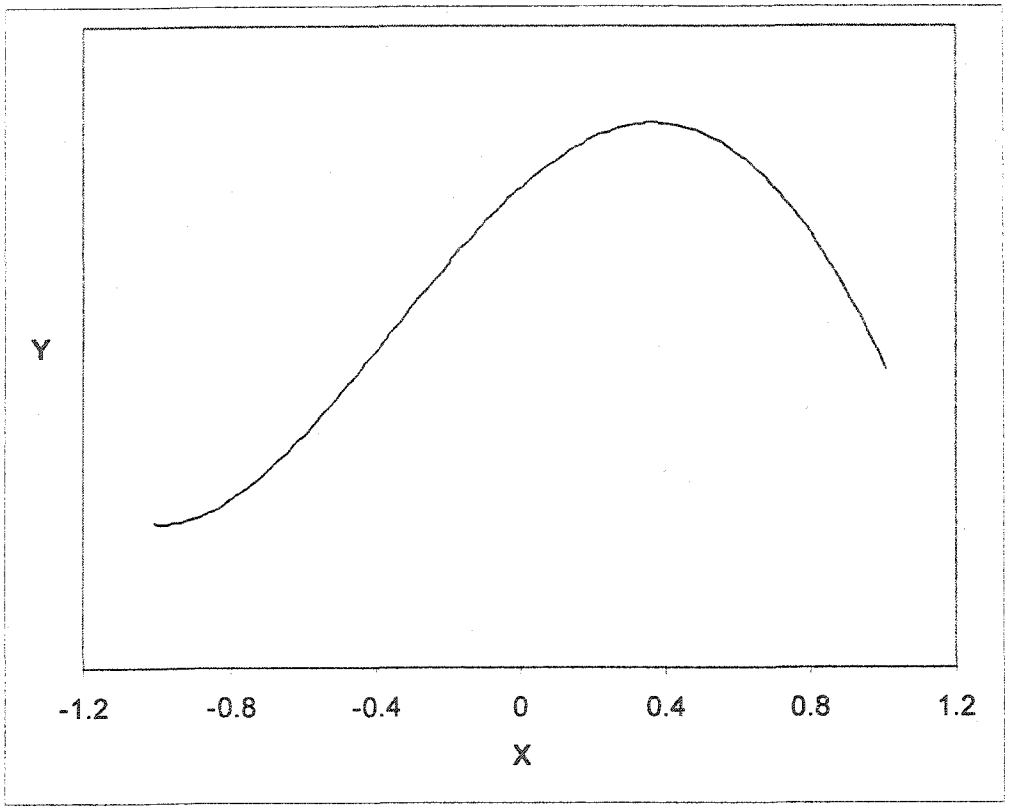

Figure 1. Fourth-order polynomial for Example 1

where

$$f' = \begin{bmatrix} 1 & x & x^2 & x^3 & x^4 \end{bmatrix}$$

$$F = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 \\ 1 & -2/3 & 4/9 & -8/27 & 16/81 \\ 1 & -1/3 & 1/9 & -1/27 & 1/81 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1/3 & 1/9 & 1/27 & 1/81 \\ 1 & 2/3 & 4/9 & 8/27 & 16/81 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

and each method specifies a particular form for the prediction vector c_x.

For the purposes of this analysis, the integral for each method is numerically approximated with the rectangular rule using 100 divisions over the range of x.

The calculation of *IMSE* for the OLS method is straightforward (Khuri,1996) as estimators for the parameters are easily expressed for any given situation. Calculation of *IMSE* for the kriging methods requires assumed values for θ and p, and the associated variance-covariance matrix. Two examples based on different assumed values of p will be used to demonstrate the differences in *IMSE* between the methods. For each example, θ will be varied throughout a range of values (0.1 to 10,000).

Example 1A ($p = 0.5$)

The *IMSE* of the OLS method is not a function of θ or p, so it remains constant at 0.160 throughout the range of θ and will serve as the benchmark for comparison (see Figure 2). For this evaluation of OK and UK, standard deviations of the gaussian process and measurement errors are assumed to be equal, resulting in $p = 0.5$. The *IMSE* of the OK method is much higher than that of the OLS method throughout the range of θ. The poor fit of OK is not unexpected as it is based on the assumption that the underlying surface is composed of random deviations from a constant. At $p = 0.5$, the magnitude of the random error appears to be large and some of the structured deviations from the constant surface tend to be interpreted as random error. This results in the kriging predictor that is too flat and does not follow the true surface.

The *IMSE* of the UK method is lower than that of OLS, indicating that UK gives a better fit to the quartic polynomial than the traditional OLS method based on the quadratic model.

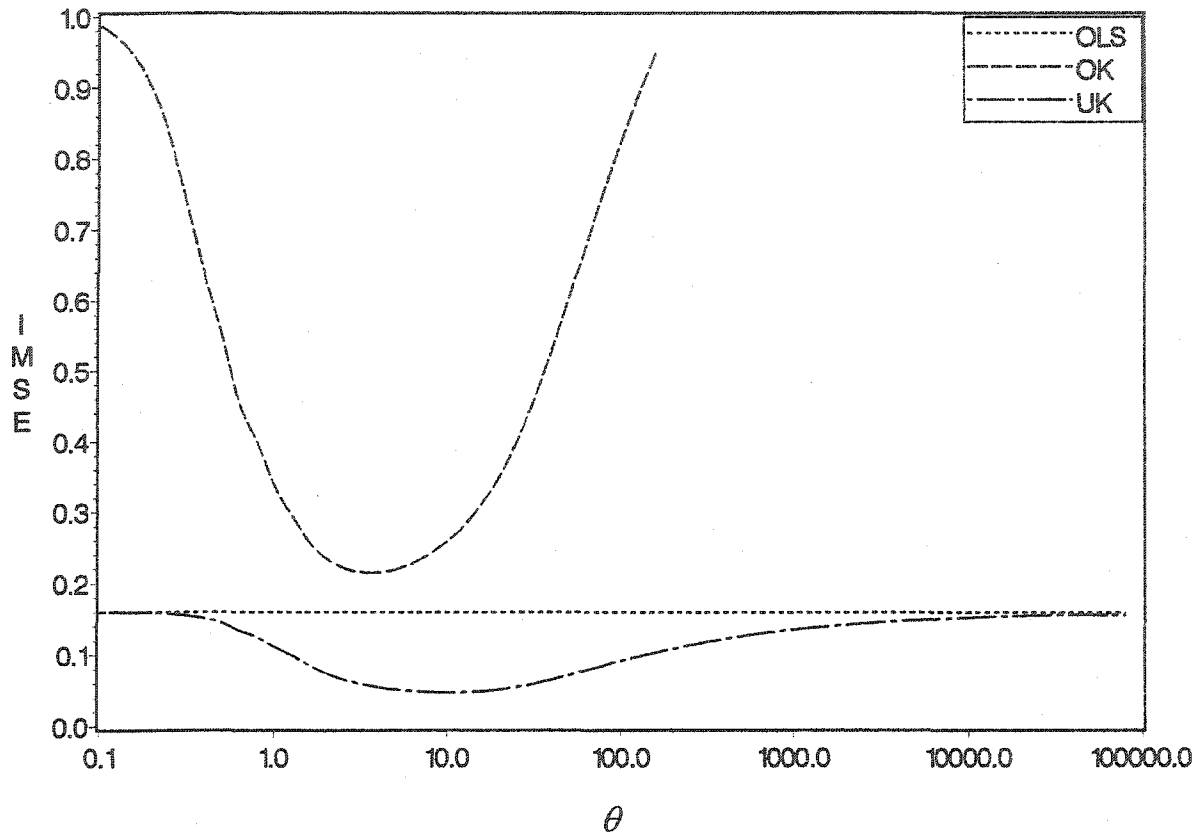

Figure 2. Example 1A, Integrated Mean Square Error for Ordinary Least Squares, Ordinary Kriging and Universal Kriging $(p = 0.5)$

Depending on the value of θ that is chosen; a substantial improvement in *IMSE* can be made. It is also interesting to note that the UK method approaches the OLS method at the extreme large and small values of θ. This is intuitively clear for large values of θ, as large θ values result in the covariance matrix approximating a multiple of the identity matrix, which reduces the UK method to a quadratic polynomial with no gaussian process (i.e. it is equivalent to the OLS method). A similar effect is seen as θ approaches 0, where all values of Z are so strongly correlated that the realization of the gaussian process becomes constant over the design region.

Example 1B $(p = 0.8)$

In the second situation, the kriging methods are evaluated under an assumption that the process variance σ_z^2 is four times the magnitude of the expected squared measurement errors, or equivalently, $p = 0.8$. Figure 3 displays *IMSE* for the three methods throughout the range of θ.

One interesting item to note is the strong improvement of the OK method relative to the results of Example 1A. We now see that for certain values of θ the OK method is actually superior in fit to the OLS method. The fit of the UK method has not changed significantly

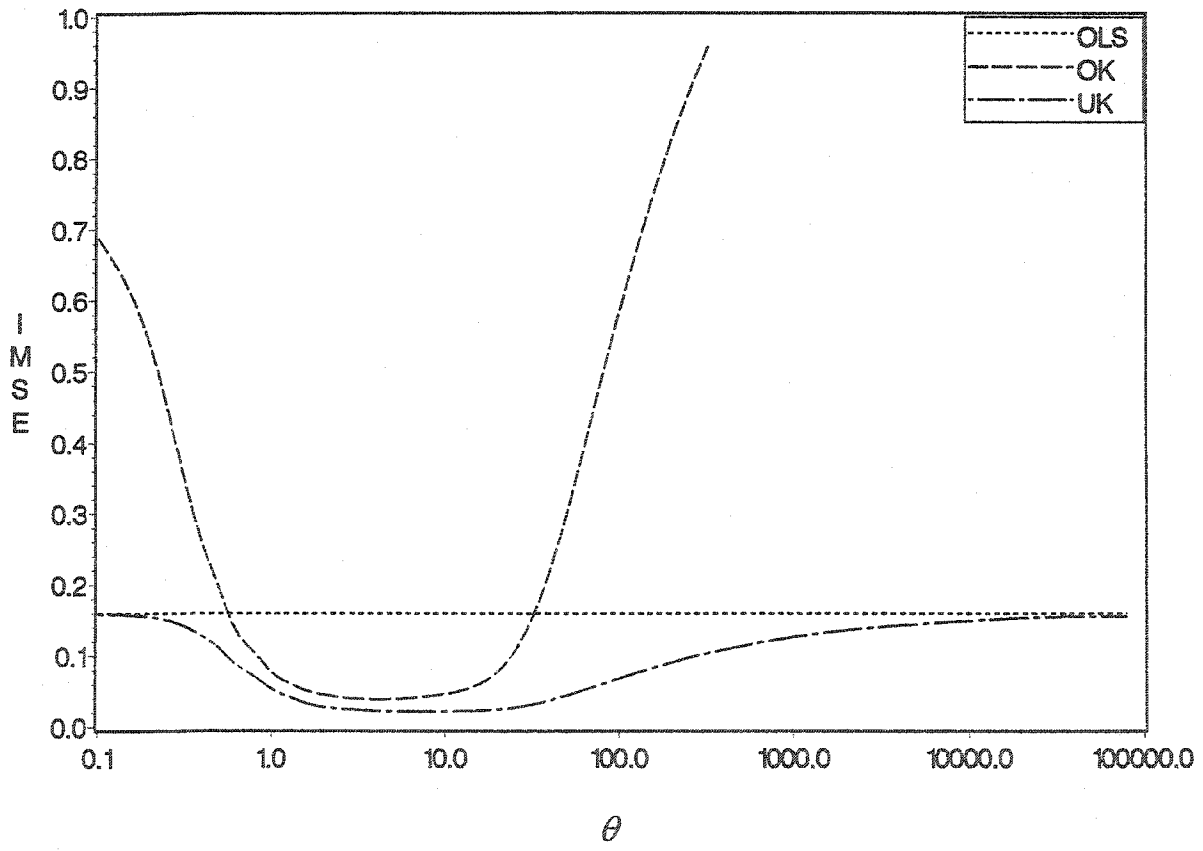

Figure 3. Example 1B, Integrated Mean Square Error for Ordinary Least Squares, Ordinary Kriging and Universal Kriging $(p = 0.8)$

from the previous example and is still superior to the OK method, but by a smaller margin for $\theta < 100$.

2.3.2 Simulations

To verify the potential of the kriging methods to outperform the OLS, a small set of simulations was run at the conditions specified for example 1. The responses generated at each of the seven x values can be found in the Table 1.

Table 1 – Response Y for 6 simulations under Example 1

	Trial					
X	1	2	3	4	5	6
-1	2.64	2.86	2.91	2.79	3.14	2.95
-0.67	3.19	2.94	2.92	3.11	3.16	3.19
-0.33	4.86	3.82	4.02	4.04	4.24	4.39
0	5.28	5.38	5.09	4.84	5.35	4.84
0.33	4.95	5.2	5.22	5.48	5.47	5.62
0.67	5.2	5.04	5.47	5.46	5.03	5.41
1	4.08	3.68	3.7	4.11	3.99	4.1

For each of these six data sets, the various methods were applied and the error of the fitted predictor with respect to the true surface was calculated. Before the kriging predictors can be fit and their errors calculated, values for the kriging covariance parameters must be chosen.

In the previous section, *IMSE* was used to show that the expected errors for the kriging methods are superior to that of OLS under certain conditions. The *IMSE* is indicative of the theoretical long run average error but it is not useful for evaluating the fit of the methods to actual surfaces, as data are not involved in its calculation. Comparisons of the methods using simulated data will require a new quantity of evaluation.

ISE

The *ISE*, integrated squared error, is the squared error of the fitted surface integrated over the area of interest, D. This represents how well the method fits the actual surface in a single application. The form is similar to the *IMSE* with the only change being the substitution of the actual squared error for the expected squared error.

$$ISE = \Omega \int_{x \in D} [\hat{y}(x) - \eta(x)]^2 \, dx \tag{2.10}$$

Method Comparisons, Example 1A continued $(p = 0.5)$

Using Figure 2, the values of the parameters for the first analysis were chosen as $\theta = 10$ and $p = 0.5$. These values were chosen as they indicate a situation where the fit of the UK method should be superior to the OLS fit and the OK fit should be inferior to OLS.

Table 2 – Integrated Squared Error assuming $\theta = 10$ and $p = 0.5$

	Trial					
Method	1	2	3	4	5	6
OLS	0.156	0.160	0.128	0.114	0.185	0.144
UK	0.156	0.048	0.015	0.032	0.058	0.050
OK	0.304	0.260	0.203	0.186	0.218	0.191

Table 2 indicates that the *ISE* achieved by the various methods is as expected. For the majority of the trials the UK method fits the data substantially better than OLS, while OK is worst in all cases. In trial 1, the OLS and UK methods have an equivalent fit, but this is not unusual, as the data contain random error and certain data configurations will favor different methods. The *IMSE* is based on an expected value and although the UK method should dominate OLS for the chosen situation in the long run, in certain simulations the OLS will provide better predictors than the UK method.

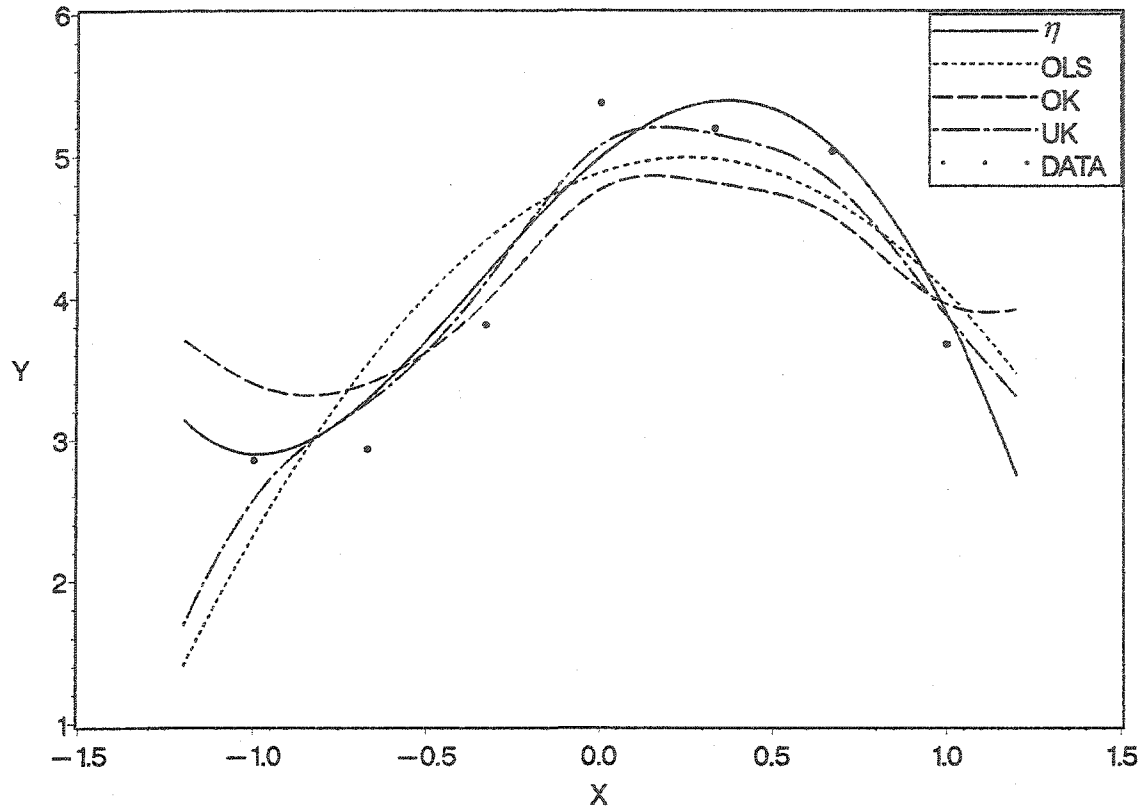

Figure 4. Comparison of Fit for OLS, UK, and OK for Example 1 Trial 2 Assuming Kriging Parameters of $\theta = 10$ and $p = 0.5$

The fit of the three methods for data from trial 2 is displayed in Figure 4. The graph illustrates the flexibility of the kriging methods to produce a predictor that deviates from an assumed quadratic polynomial.

Method Comparisons, Example 1B continued $(p = 0.8)$

Using Figure 3, the values of the parameters for the second analysis were chosen as $\theta = 10$ and $p = 0.8$. These values were chosen, as they indicate a situation where the fit of both the UK and OK methods should be superior to the OLS fit.

Table 3 – Integrated Squared Error assuming $\theta = 10$ and $p = 0.8$

	Trial					
Method	1	2	3	4	5	6
OLS	0.156	0.160	0.128	0.114	0.185	0.144
UK	0.186	0.073	0.034	0.044	0.044	0.047
OK	0.187	0.099	0.047	0.040	0.043	0.037

Table 3 indicates that for the majority of the trials the UK and OK methods fit the data significantly better than the OLS. In trial 1, the OLS method produces a better fit than either kriging method, due to the data configuration for this trial, but in all other trials the kriging methods dominate the OLS fit. A comparison of Tables 2 and 3 indicate a significant improvement in the fit of the OK method for $p = 0.8$ relative to $p = 0.5$ while little change is evident in the UK method. This trend would be expected based on the *IMSE*'s displayed in Figures 2 and 3.

The fit of the three methods for trial 2 when $\theta = 10$ and $p = 0.8$ is displayed in Figure 5. A comparison of Figures 4 and 5 reveals the effect of varying the choice of p in the kriging models for this particular data set. As p is increased, the kriging methods interpret most of the variation in the data as reflecting variation in the true surface, rather than as measurement error, which results in the kriging predictors following the data more closely. This trend is quite evident in the comparison between the two figures. Both of the kriging predictors are substantially closer to the data in Figure 5, where the value of p is closer to one.

It is important to note that these examples are based on a single quartic surface with an assumption of prior knowledge of p and θ for both kriging methods. It is quite evident in this particular example that the fit of the OK method is heavily influenced by the choice of p and θ. In an actual data analysis, these parameters would usually be estimated by the experimenter.

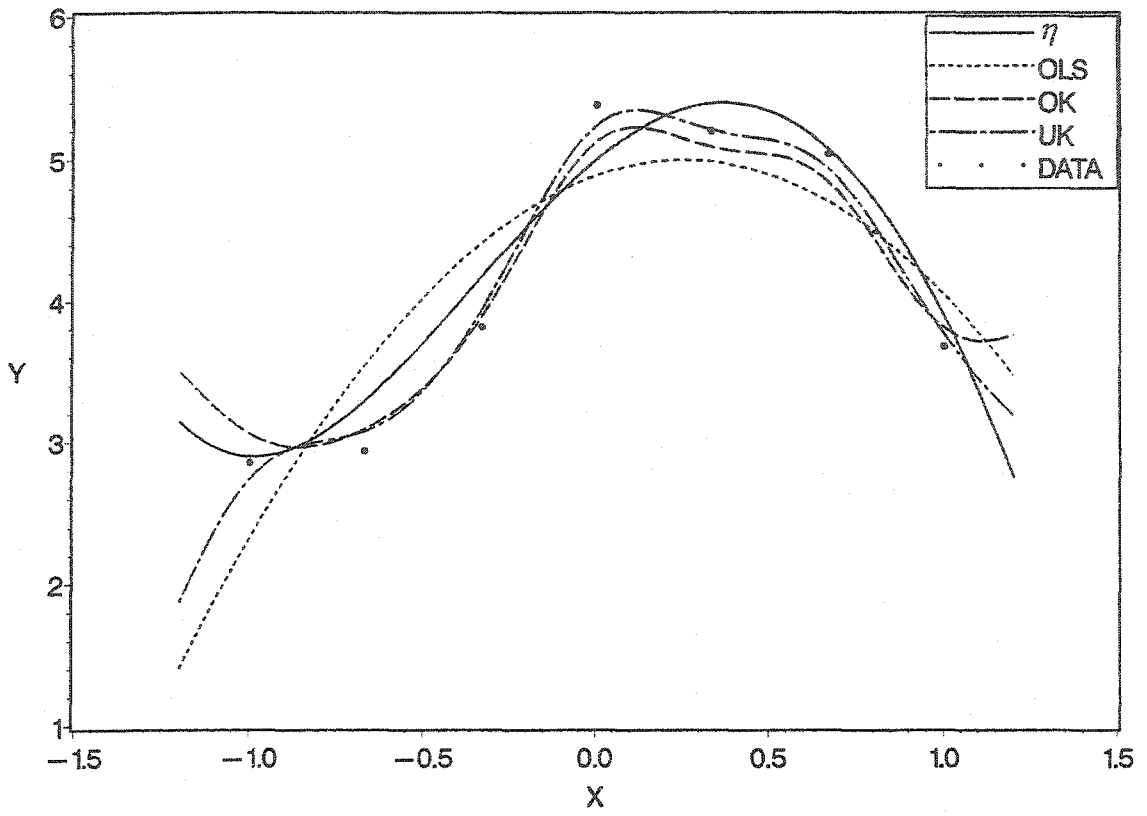

Figure 5. Comparison of Fit for OLS, UK, and OK for Example 1 Trial 2 Assuming Kriging Parameters of $\theta = 10$ and $p = 0.8$

Additionally, the UK method dominates the OK method with respect to *IMSE* for all combinations of p and θ examined here, but there is no guarantee that similar estimates of these parameters would be obtained under the OK and UK methods for any given data set. For example, estimates of p and θ might be 0.8 and 10, respectively, under the OK method, but 0.8 and 1000, respectively, under the UK method. Conditional on these values, *IMSE* for OK is actually less than that of UK under the conditions of Example 1. Hence it is important to understand the statistical properties of the estimates of p and θ and of $\hat{y}(x)$ when p and θ are estimated.

CHAPTER 3 NUMERICAL STUDIES I

It was noted that each of the predictors of $\eta(x)$ used in the previous examples is BLUP when the assumptions of the selected method are correct. While the OLS predictors are BLUP as they would commonly be used for data analysis, the kriging predictors are BLUP only in the case where the values of θ and p are known. In almost all real experiments, however, these parameters would be unknown and would be estimated by the experimenter. This chapter presents a comparison of methods under the more realistic situation in which estimates of θ and p, rather than known or assumed values, are used. We focus on maximum likelihood estimates of these parameters based on the same set of data used to fit the predictor.

3.1 Obtaining Estimates of p and θ

The *IMSE* values shown in Chapter 2 indicate the potential for the kriging methods to outperform OLS as used in RSM when the true surface differs from the form of the assumed model. Unfortunately, *IMSE* as defined in equation (2.9) cannot be used to prove the superiority of the kriging methods for practical purposes as it assumes prior knowledge of p and θ, which in practice must be estimated. In this analysis we will examine the performance of OK and UK (relative to OLS) when maximum likelihood estimates (MLE) of θ and p are used in place of assumed values. As Z and ε are assumed to be normal for each kriging method, we can write the likelihood function and log-likelihood function as

$$L = (2\pi)^{-n/2}|V|^{-1/2}\exp\left(\frac{-1}{2}(Y - X\beta)'V^{-1}(Y - X\beta)\right)$$

and

$$\ell = \frac{-1}{2}\left[n\ln 2\pi + \ln|V| + (Y - X\beta)' V^{-1}(Y - X\beta) \right]$$

respectively. Re-expressing the covariance matrix as

$$V = \sigma^2 W$$

yields:

$$\ell = \frac{-1}{2}\left[n\ln 2\pi + n\ln(\sigma^2) + \ln|W| + (Y - X\beta)' \frac{W^{-1}}{\sigma^2}(Y - X\beta) \right].$$

Setting the derivative of the log-likelihood with respect to σ to zero,

$$\frac{\partial \ell}{\partial \sigma} = \frac{-1}{2}\left[n\frac{2\sigma}{\sigma^2} - \frac{2}{\sigma^3}(Y - X\beta)' W^{-1}(Y - X\beta) \right]$$

$$0 = \left[n - \frac{1}{\hat{\sigma}^2}(Y - X\beta)' W^{-1}(Y - X\beta) \right]$$

$$\hat{\sigma}^2 = \frac{(Y - Xb)' W^{-1}(Y - Xb)}{n}.$$

Likewise, differentiating with respect to β and setting the result to zero,

$$\frac{\partial \ell}{\partial \beta} = -\frac{1}{\sigma^2} X' W^{-1}(Y - X\beta)$$

$$0 = X' W^{-1}\left(Y - X\hat{\beta}\right)$$

$$\hat{\beta} = \left(X' W^{-1} X\right)^{-1} X' W^{-1} Y.$$

These expressions allow for the determination of the MLEs for β and σ^2, given a known correlation matrix W.

Substitution of the MLE's of β and σ^2 into the log-likelihood results in a function that is only dependent on W (i.e. θ and p),

$$\ell = \frac{-1}{2}\left[n\ln(2\pi) + n\ln\left(\frac{\left(Y - X\left(X'W^{-1}X\right)^{-1}X'W^{-1}Y\right)W^{-1}\left(Y - X\left(X'W^{-1}X\right)^{-1}X'W^{-1}Y\right)}{n}\right) + \ln|W| + n\right]$$

and maximizing this quantity is equivalent to maximizing

$$\ell^* = -\left(n\ln\left(\frac{\left(Y - X\left(X'W^{-1}X\right)^{-1}X'W^{-1}Y\right)W^{-1}\left(Y - X\left(X'W^{-1}X\right)X'W^{-1}Y\right)}{n}\right) + \ln|W|\right).$$

The correlation matrix W and estimated coefficient vector b of the kriging methods are both functions of p and θ and the MLE's of these parameters do not have closed form solutions. As an alternative, a numerical optimization program utilizing first derivatives can be used to maximize the likelihood with respect to p and θ. Taking the derivative of the likelihood with respect to p and θ and using ω_Θ to represent the derivative of the correlation matrix with respect to either of these variables

$$\frac{\partial W}{\partial \Theta} = \omega_\Theta \qquad \Theta = \theta \quad or \quad p,$$

yields:

$$\frac{\partial \ell}{\partial \Theta} = tr\left(W^{-1}\omega_\Theta\right) + \frac{n}{\left(Y - X\left(XW^{-1}X\right)^{-1}XW^{-1}Y\right)'W^{-1}\left(Y - X\left(XW^{-1}X\right)^{-1}XW^{-1}Y\right)}$$

$$\left[Y'\left(-W^{-1}\omega_\Theta W^{-1}X\left(XW^{-1}X\right)^{-1}XW^{-1} + \right.\right.$$

$$\left.\left.W^{-1}\left(X\left(XW^{-1}X\right)^{-1}XW^{-1}\omega_\Theta W^{-1}X\left(XW^{-1}X\right)^{-1}XW^{-1} - X\left(XW^{-1}X\right)^{-1}XW^{-1}\omega_\Theta W^{-1}\right)\right)Y\right]$$

where the individual components of ω_Θ are

$$\omega_{ij} = \begin{cases} 0 & i = j \\ -p\Delta_{ij}\exp\left(-\theta\Delta_{ij}\right) & i \neq j \end{cases} \qquad \omega_{ij} = \begin{cases} 0 & i = j \\ \exp\left(-\theta\Delta_{ij}\right) & i \neq j \end{cases}$$

$$\Theta = \theta \qquad\qquad \Theta = p$$

and

$$\Delta_{ij} = \left(s_i - s_j\right)'\left(s_i - s_j\right)$$

is the square of the distance between the points i and j.

The log-likelihood function was numerically maximized with a NAG routine utilizing these two first-order differential equations. The algorithm uses 10 initial starting locations, with $\theta = 1$ and p varied from .01 to .99 in 10 equal increments, to find the estimates $\hat{\theta}$ and $\hat{p}$ which maximize the log-likelihood function.

Having defined a method of determining estimates of θ and p, we can now numerically compare the various methods as they would be used in practice under different true surfaces.

3.2 Simulated Surfaces and Measurement Error

We shall limit this investigation to situations in which the surfaces being estimated are usually uni-modal and approximately quadratic in shape. The focus of this research is a comparison of the performance of kriging and regression methods when the surface is approximately, but not exactly, quadratic. For this reason, the surfaces used in this analysis were chosen to follow a basic quadratic form with slight variation. The variation from the quadratic is achieved through the addition of a realization of a gaussian random process, i.e. through a model form corresponding to the assumed model used in UK. Throughout the study each of the methods will be applied to data, Y, generated from a model consisting of the true surface, η, plus random error, ε,

$$Y = \eta + \varepsilon$$

in which η consists of a polynomial trend and/or a spatial stochastic process.

As RSM analyses often involve more than one factor, all of the methods will be tested in one, two, and three dimensions.

3.2.1 Surfaces

Surface 1 - Constant

The first surface is the simplest with the observed response being the sum of a constant, β_o, and measurement error. This is the situation in which the response is not affected by the factors and the only source of differences between measurements is measurement error. It is of limited interest in practical data analysis, but is included in this study as it is the simplest model form generally considered in regression analysis:

$$\eta = \beta_o .$$

Surface 2 – Quadratic Polynomials

The second surface is a second-order polynomial, a form commonly used in RSM:

$$\eta = P_2(x),$$

where

$$P_2(x) = \beta_o + \sum_{i=1}^{d} \beta_i x_i + \sum_{i=1}^{d} \beta_{ii} x_i^2 + \sum_{i<j=1}^{d} \beta_{ij} x_i x_j .$$

Throughout the simulations P_2, the polynomial function, will be varied to allow for three different shapes of polynomial surfaces.

Shape 1

The first shape is a hyperplane, hence

$$\beta_{ii} = \beta_{ij} = 0 \quad \forall i, j = 1,2,...,d .$$

The quality of fit for OLS and the kriging methods is invariant to the value of β_o, while the choice of β_o will affect BC estimates (see Section 2.1.1). For consistency, all surfaces were constructed with a minimum value of 0 and linear coefficients set to a common value such that the range of the true surface throughout the design space is one, that is,

$$\max_{x \in D} \eta(x) - \min_{x \in D} \eta(x) = 1 .$$

The specific coefficient values used in generating data are listed in Table 4.

Table 4 – Coefficients for shape 1

Number of Factors	β_o	β_i	β_{ii}	β_{ij}
1	0.5	0.5	0	0
2	0.5	0.25	0	0
3	0.5	0.167	0	0

Shape 2

The second shape is a quadratic ridge along the $x_1 = x_2 = ... = x_d$ line. In this situation η is maximized at any point on this line and elsewhere is determined by the squared distance to this line. For such polynomials

$$\beta_i = 0 \ \ \forall i = 1,2,...,d .$$

A "ridge" is not meaningful in the one-dimensional case, so $d = 1$ is not considered. For $d =$ 2 the surface is a ridge along the $x_1 = x_2$ line on the plane. For $d = 3$, the surfaces of equal response are tubes along the $x_1 = x_2 = x_3$ line. As in shape 1, the specific coefficient values used in generating data were chosen such that the range of the surface, $\max \eta(x) - \min \eta(x)$, was set to one. These coefficients are listed in Table 5.

Table 5 – Coefficients for shape 2

Number of Factors	β_o	β_i	β_{ii}	β_{ij}
2	1.0	0	-0.25	0.5
3	1.0	0	-0.5	0.5

Shape 3

The final surface shape used is mound with maximum at $x = (0.5,0.5,...,0.5)$. This is the type of surface most commonly investigated with RSM. Contours of equal expected response values were set to be circular or spherical, so

$$\beta_{ij} = 0 \quad \forall i, j = 1,2,...,d .$$

In the one-dimensional case $\eta(x)$ simplifies to a parabola and in the three-dimensional case it is a function whose value is completely determined by the distance from x to the point $(0.5,0.5,0.5)$. The coefficients were chosen such that the maximum and minimum of η in the experimental region, now located at $x = (0.5,0.5,...,0.5)$ and $x = (-1,-1,...,-1)$ respectively, were once again one and zero, respectively. The specific coefficient values used in generating data are listed in Table 6.

Table 6 – Coefficients for shape 3

Number of Factors	β_o	β_i	β_{ii}	β_{ij}
1	0.888	0.444	-0.444	0
2	0.888	0.222	-0.222	0
3	0.888	0.148148	-0.148148	0

Surface 3 – Gaussian Process

The third surface is a realization of a stationary gaussian process, *Z(x)*. This random gaussian process has mean zero, correlation function *R* and the process variance σ_z^2:

$$\eta = Z(x)$$

where

$$E(Z(x)) = 0 \quad \text{and } Cov(Z(u), Z(t)) = \sigma_z^2 R(u,t) \quad u,t \in S .$$

For these simulations, a gaussian correlation function was chosen because its realizations are smooth functions, which is a reasonable assumption about η in most RSM contexts. The

correlation between any two points $t' = (t_1 \quad t_2 \quad .. \quad t_d)$ and $u' = (u_1 \quad u_2 \quad .. \quad u_d)$ is defined by the distance between the points, and a parameter θ:

$$R(t,u) = \exp\left(-\sum_{k=1}^{d} \theta (t_k - u_k)^2\right).$$

The choice of θ will determine the rate at which the correlation decreases with increasing distance between t and u, with large values of θ approaching *iid* departures from the process mean.

Note that apart from the requirement that the $E(\eta(x)) = 0$, Surface 3 corresponds exactly to the assumed model underlying OK, hence we may expect this method to perform well in this section of the study.

Surface 4 – Polynomial and Gaussian Process

The final surface is the most important in the analysis. The combination of the polynomial and gaussian processes utilizing relatively small values for σ_z defines surfaces that are slight deviations from polynomial functions. The form of the last surface is

$$\eta = P_2(x) + Z(x)$$

where P_2 is as described for Surface 2 and $Z(x)$ is as defined for Surface 3. Note that Surface 4 corresponds exactly to the assumed model underlying UK, hence we may expect this method to perform well in this section of the study.

3.2.2 Simulations

Strength of correlation in Surfaces 3 and 4

The strength of the correlation for the gaussian process is determined by the correlation parameter θ. Small values of θ result in highly correlated gaussian processes with realizations that change gradually within the design space, while large values of θ result in smaller correlations between any two fixed points and allow the gaussian process to vary substantially over relatively small distances. As was stated earlier, the focus of this analysis is on response surfaces which are approximately quadratic in shape, so the strength of spatial correlation was limited to maintain this structure.

To standardize the correlation over the considered values of d, θ is chosen such that the correlation between values of Z at the center point and the most remote point in the design space is a specific value, regardless of the dimension of the problem. The choice of θ results in a specified value for this correlation:

$$R(1_d, 0_d) = \exp\left(-\sum_{k=1}^{d} \theta(1-0)^2\right) = \exp(-\theta d).$$

This analysis uses three levels of correlation to allow for varying levels of complexity in Z over the design space D. Informal analysis suggested that the generated realizations are approximately quadratic when $R(1_d, 0_d)$ is set to the values of 0.4, 0.3, and 0.2. Hence for each value of d, the three values of θ are determined as

$$\theta = \frac{-\ln(R(1_d, 0_d))}{d} \quad \text{where } R(1_d, 0_d) = 0.2, 0.3, 0.4$$

Standard Deviation of the Gaussian Process

The typical magnitude of Z, which represents the departure of the response surface from the polynomial, is controlled by the magnitude of σ_z. For the simulations, the median range of the gaussian process

$$M(\theta,\sigma_z,d) = \underset{z}{median}\left\{\max_{x\in D}(Z(x)) - \min_{x\in D}(Z(x))\right\}$$

was standardized to be a percentage of the range of the underlying polynomial surface. The levels $M(\theta,\sigma_z,d)$ were chosen to represent small, medium, and large contributions by the gaussian process to the final surface shape. Median ranges of 10%, 20% and 30%, were chosen to represent the respective processes and although these ranges may appear large, the shape of gaussian process remains approximately quadratic.

Setting the values of σ_z to obtain the desired value of $M(\theta,\sigma_z,d)$ is not as simple as it would be for independent random deviations. The difficulty arises in the changing correlation structure for the gaussian processes. For a given experimental region D and σ_z, a gaussian process with $R(1_d,0_d)$ of 0.4 will typically have a smaller range of values than the same process with $R(1_d,0_d)$ of 0.2. To compensate for this effect, simulations were run to compare the maximum and minimum values of the gaussian processes over the three levels of correlations for each of the three dimensions.

The first step in the range simulations is to generate independent random errors for a standard normal distribution and assign them to a regular grid within D. For this study the regular grid consisted of 100 points for $d = 1$, 400 points for $d = 2$ and 8000 points for $d = 3$. These independent errors are then converted into the properly correlated gaussian process through the use of a Cholesky decomposition of the calculated correlation matrix for the grid locations (see next section).

The gaussian process was simulated 1000 times for each grid and $M(\theta,\sigma_z,d)$ was then used to calculate a scaling factor, equal to the inverse of $M(\theta,\sigma_z,d)$, for later simulations

with polynomial surfaces. A value of 1 for $M(\theta,\sigma_z,d)$ would indicate that the median value for the range of the gaussian process over D would approximately equal the magnitude of the standard deviation used to generate the uncorrelated random errors. In this case, to achieve the median range of the gaussian process to be 10% of the polynomial surface, the magnitude of the standard deviation for the generated random errors should be set at 10% of the range of the underlying polynomial. Similarly, a value of 3 would indicate that the standard deviation for the generation of the random error should be set at 3.3% of the range of the underlying polynomial to achieve a gaussian process that has a magnitude of approximately 10% of the polynomial. Table 7 below indicates that the median range of a gaussian process is approximately 1.5 to 3.6 depending on the dimension of the grid and correlation of the process.

Table 7 - Magnitude of the median ranges of the Gaussian Processes for 1000 simulations

$R(1_d,0_d)$	$d=1$	$d=2$	$d=3$
.2	1.74	2.75	3.59
.3	1.65	2.55	3.24
.4	1.46	2.26	2.94

Use of the above scaling factors on σ_z, result in a consistent contribution by the gaussian process to the shape of the surface for the various levels of correlation, regardless of the underlying polynomial or dimension of the simulation.

Simulating the Gaussian Process

As was mentioned previously, the gaussian process was constructed via the use of a Cholesky decomposition of the appropriate correlation matrix for each grid. To construct the surface, a vector G composed of the appropriate number of randomly generated standard normal values g_i was created.

$$G = \begin{bmatrix} g_1 \\ g_2 \\ \\ \\ g_k \end{bmatrix} \quad k = m^d \quad \text{where} \quad \begin{matrix} m = 100 & d = 1 \\ m = 20 & d = 2,3 \end{matrix}$$

A correlation matrix H is constructed for the grid using equation (2.5). For $d = 1$, H is a Toeplitz matrix with ones on the main diagonal and decreasing values for subsequent diagonals. For higher dimensional grids, the correlation matrix is a block Toeplitz matrix with ones on the main diagonal. As H is a symmetric positive definite matrix a Cholesky decomposition can be taken which yields a unique lower triangle matrix T:

$$H = TT'. \tag{3.1}$$

Multiplying the vector of independent random errors by the Cholesky matrix results in a vector of correlated errors, which are the values of a single realization of the gaussian surface on the grid,

$$Z = TG.$$

The large size of grids used in the simulations results in numerical difficulties in the construction of the gaussian process realizations. For $d = 3$, the 8000 point grid, 20 locations per dimension, results in H being an 8000 by 8000 matrix. Direct attempts at Cholesky decompositions of matrices of this size result in numeric instabilities making this process impossible. The special circumstance of the grid pattern and its resulting block Toeplitz

correlation matrix allows for an alternative method for the construction of the Cholesky matrix T.

As was stated earlier, for $d = 1$ H is a Toeplitz matrix now designated as h. For higher dimensions H is a block Toeplitz matrix with a special structure. Due to the regular grid pattern, H will be a kronecker product of the original Toeplitz matrix:

$$\begin{aligned} H &= h \qquad d = 1 \\ H &= h \otimes h \qquad d = 2 \\ H &= h \otimes h \otimes h \qquad d = 3. \end{aligned} \tag{3.2}$$

This special structure allows for an alternative method for data generation, which can be easily shown in the two-dimensional case. As h is a symmetric positive definite correlation matrix, it has a unique Cholesky decomposition

$$h = tt'$$

which can be substituted into (3.2)

$$H = h \otimes h = (tt') \otimes (tt') = (t \otimes t)(t' \otimes t') = (t \otimes t)(t \otimes t)'. \tag{3.3}$$

The kronecker product of two lower triangle matrices is also a lower triangle matrix. Combining equation (3.3) with (3.1) yields that the lower triangle matrix created by $t \otimes t$ must be the same unique lower triangle matrix T created with the Cholesky decomposition of H:

$$H = TT' = (t \otimes t)(t \otimes t)'.$$

This relationship allows for a significant improvement in construction of the gaussian process. Rather than attempting to decompose an 8000 by 8000 correlation matrix, the algorithm only requires the decomposition of a 20 by 20 matrix.

Standard Deviation of Random Error

In addition to the deviations from the underlying polynomial due to the correlated gaussian process, measurement error is also present in each measurement. All of the errors generated for the simulations will be obtained from a normal distribution.

$$Y = \eta + \varepsilon \quad \text{where } \varepsilon \sim iid\ N\left(0, \sigma_\varepsilon^2\right).$$

To standardize the size of error used in each simulation, σ_ε was defined to be a percentage of the range of the polynomial surface, P_2, within the experimental region

$$\sigma_\varepsilon \propto \left(\max_{x \in D} P_2(x) - \min_{x \in D} P_2(x) \right).$$

In this way, the choice of scale for the polynomial coefficients did not influence the method comparisons. To maintain a balance between the correlated gaussian process and the random error, three values of σ_ε were chosen such that the length of the central 95% probability interval for the random error would correspond to 10%, 20%, and 30% of the underlying polynomial range (i.e. σ_ε was set to be 2.5%, 5% and 7.5% of the polynomial range).

In the case of Surface 1, where P_2 is a constant, the magnitude of the error cannot be linked to the range of the polynomial and only one level of random error will be used. All simulations in this case will have the magnitude of the random error arbitrarily be set at $\sigma_\varepsilon = 0.10$.

3.3 Experimental Designs

Two-level factorial and fractional factorial experimental designs are widely used in industrial experimentation. These plans call for taking data at a combination of relatively high and low levels of the factors and are sometimes augmented with the addition of center points to allow for estimation of error variance and testing for LOF. While two-level factorial designs can detect LOF in some cases, they are inadequate for estimating all parameters in second-order regression models. For this reason, RSM practitioners often use other families of plans explicitly designed for the estimation of quadratic effects, with the Central Composite (Box and Wilson, 1959) and Box-Benhken (Box and Benhken, 1960) designs being the most popular for second-order modeling. This analysis will focus on a Central Composite Designs (CCD) and more specifically on a sub-class called Face-Centered Central Composite Designs (CCF).

A CCD is composed of two main components. The first component is a two level factorial or fractional factorial plan and the second component, added specifically to allow estimation of quadratic effects, consists of one or more measurements at the center point and a "star" formation. The star contains runs in which each factor individually is set at either a low or high level with all other factors held at central values. For this reason there are always twice as many star points as there are factors in the experiment. In a CCF the star points are taken at the same scaled level used for the factorial points, with +1 representing the maximum, -1 representing the minimum, and 0 representing the central value.

A CCF with $d = 3$ variables has the following design matrix S (shown on next page), whose points contained in the three-dimensional CCF are displayed in Figure 6.

$$S = \begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 \end{matrix} \\ \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

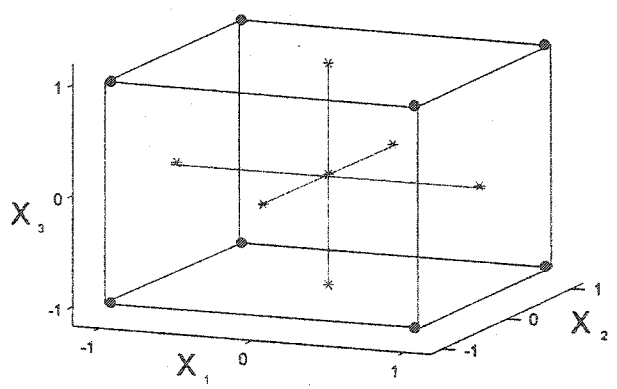

Figure 6 – CCF in three dimensions with $(\bullet)$ representing the factorial component and $(*)$ representing the star component of the design

The standard CCF will be the smallest experimental design studied in this investigation. Kriging methods, with their increased flexibility, require the estimation of additional covariance parameters, which may not be feasible under the standard design. Hence, we limit attention to designs for which the number of unique factor combinations is at least twice the number of coefficients included in a quadratic polynomial model. Additional design points will be appended to the standard CCF, when it does not meet this condition.

As the kriging methods are spatially oriented, designs that maximize the distance between design points are intuitively appealing. The standard CCF for three variables contains 15 design points, and so does not meet the requirement for the minimum number of design points, which is 18 points. To meet this requirement and maintain the symmetry of the design eight new locations are added to the design, for a total of 23 design locations. To maximize the distance between the points in each one-dimensional projection of the augmented design, the new locations are placed in factorial locations scaled to one-half of range the original factorial design. The additional design points S_{Δ} are listed below

$$S_{\Delta} = \begin{bmatrix} -0.5 & -0.5 & -0.5 \\ -0.5 & -0.5 & 0.5 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & -0.5 \\ 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix}.$$

This augmentation results in a cube being embedded within the original CCF. The augmented CCF in three factors is displayed in Figure 7.

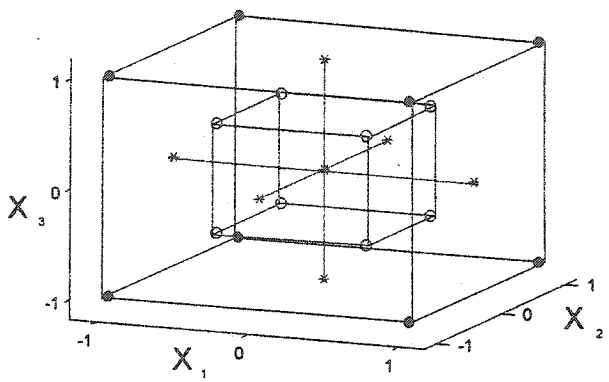

Figure 7 – Augmented CCF in three dimensions with $(\bullet)$ representing the factorial component, $(*)$ representing the star component, and $(\circ)$ representing the augmented component of the design

The design for two factors will follow the same general pattern as described for the three factor design, but requires fewer unique design points. The basic design is a CCF, which for $d = 2$ results in a regular 3-level grid pattern:

$$S = \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 0 & 0 \\ -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}.$$

As the standard CCF in two-dimensions does not meet the requirement for the minimum number of design points, this design is also augmented. Maintaining the symmetry of the design and maximizing the space between locations in one-dimensional projections of the design, four additional points are inserted at factorial points scaled to one-half the original factorial.

$$S_{\Delta} = \begin{bmatrix} -0.5 & -0.5 \\ -0.5 & 0.5 \\ 0.5 & -0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

The augmented CCF in two factors is displayed in Figure 8.

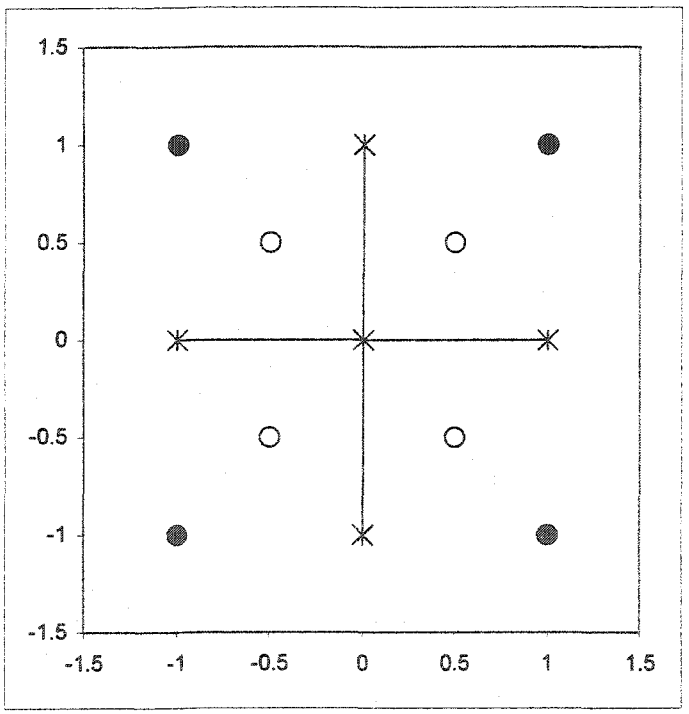

Figure 8 – Augmented CCF in two dimensions with $(\bullet)$ representing the factorial component, $(*)$ representing the star component, and $(\circ)$ representing the augmented component of the design

The pattern for a CCD degenerates in the one-dimensional case, because the star points replicate the factorial points. Hence, an alternative strategy will be used that maintains symmetry and equal spacing along the single axis. Design locations will be placed at equal intervals between the minimum and maximum values of the factor. For this analysis, the number of design locations will always be odd so that a center point is always included in the design. Doubling the number of coefficients in the quadratic polynomial requires six data points in the design so the number of equally spaced points in the design is seven.

$$S \in [-1, \ -2/3, \ -1/3 \ \ 0, \ \ 1/3, \ \ 2/3, \ \ 1].$$

CCD's include at least one measurement taken at the center of the design space, but it is standard practice to take replicate measurements at this location to facilitate estimation of the variance of measurement error in an experiment. For all of the simulations in the analysis, an additional five measurements (e.g. six in all) will be included at the center point .

3.4 Quantifying Error

In fitting a surface to our data, we are always concerned about the quality of the fitted surface as an approximation of the true surface. There are several indices that can be used to evaluate the fit of the method. Two such indices, *MSE* (2.8) and *IMSE* (2.9), were used in evaluations in Chapter 2. Both have shortcomings: *MSE* is defined only for a single specific value of x and both *MSE* and *IMSE*, as derived in Chapter 2, require specific values of θ and p. *ISE* was derived as an alternative to these methods (2.10) and it will be used exclusively throughout this analysis as it can be used to evaluate the fit of the method to simulated data. Additionally, two other indices of evaluation will be suggested.

3.4.1 Estimated *ISE* $\left(\hat{ISE}\right)$

ISE is a useful index for simulated studies, but cannot be used directly in data analysis because it is a function of the unknown true response function. The estimated Integrated Square Error $\left(\hat{ISE}\right)$ is the integral of the *MSE* as derived in (2.8), with maximum likelihood estimates substituted for parameters, under the assumption that the correct model has been selected. As all of the methods are unbiased under the correct model assumption, we find that the second term in the *MSE* as defined in (2.8) is equal to zero, that is:

$$\left(c'F\beta - f\beta\right) = 0$$

so, given the assumption of correct model form,

$$\hat{ISE} = \Omega \int_{x \in D} \left[c_x' \hat{V} c_x - 2c_x' \hat{v}_x + \hat{\sigma}_z^2\right] dx .$$

OLS

In this case the experimenter will be fitting the simple quadratic model. The method assumes independent errors and no spatial gaussian process, reducing the form of $\hat{ISE}$ significantly:

$$\hat{ISE}_{ols} = \Omega \int_{x \in D} \left[c_x' \hat{V} c_x - 2c_x' \hat{v} + \hat{\sigma}_z^2\right] dx = \Omega \int_{x \in D} \left[c_x' \hat{\sigma}_\varepsilon^2 I c_x - 2c_x' 0 + 0\right] dx$$

$$= \Omega \hat{\sigma}_\varepsilon^2 \int_{x \in D} \left[c_x' c_x\right] dx$$

BC

The form of $\hat{ISE}$ is exactly the same for BC as it was for OLS, but the estimate $\hat{\sigma}_\varepsilon^2$ has been calculated from the transformed responses y^λ (see 2.1.1). An adjustment must be made to estimate the *ISE* for the untransformed predictions using the transformed predictions $\hat{y}_{BC}$:

$$\hat{ISE}_{BC} = \Omega\hat{\sigma}_\varepsilon^2 \int_{x\in D} e^{2\hat{y}_{BC}}[c_x'c_x]\,dx \qquad \text{if } \lambda = 0$$

$$\hat{ISE}_{BC} = \Omega\hat{\sigma}_\varepsilon^2 \int_{x\in D} \left(\lambda^{-1}\hat{y}_{BC}^{\lambda^{-1}-1}\right)^2[c_x'c_x]\,dx \qquad \text{if } \lambda \neq 0.$$

Kriging

For OK and UK the experimenter does not know the true value parameter values of the gaussian process and estimates of the variance and covariance associated with Z are substituted for their respective parameters in defining $\hat{ISE}$:

$$\hat{ISE}_{krig} = \int_{x\in D}\left[c_x'\hat{V}c_x - 2c_x'\hat{v}_x + \hat{\sigma}_z^2\right]dx = \Omega\int_{x\in D}\left[c_x'\hat{V}c_x - 2c_x'\hat{V}\right]dx + \hat{\sigma}_z^2.$$

Ratio

Another quantity that will be examined in this analysis is the ratio of the of $\hat{ISE}$ and ISE.

$$\Gamma = \frac{\hat{ISE}}{ISE}$$

The ratio Γ represents how closely the quoted error index ($\hat{ISE}$) represents the true error index (ISE). If the $\hat{ISE}$ is a good estimate, we would expect a ratio very close to one most of the time. If the $\hat{ISE}$ is a conservative estimate of error, we expect a ratio typically greater than one and if it is too optimistic we expect a ratio less than one.

Having defined an appropriate measure for evaluating the fit of the various methods we can now proceed to simulating surfaces and data and determining which of the methods gives the best fit to each.

3.4.2 Calculation of Error

The integral required for calculating *ISE* and $\hat{ISE}$ will be numerically approximated for all methods via the rectangular rule. The number of elements used in the approximation depends on the number of dimensions in the simulation. In the one-dimensional case, the integral will be approximated with 100 rectangular areas of width 0.02 units. In the two-dimensional case a 20x20 grid, 400 volume elements, will be used to integrate the unit square resulting in 0.10 units between center points in each dimension. For the three-dimensional case a 20x20x20 grid, 8000 volume elements, will be used to integrate over the unit cube, which also results in 0.10 units between center points in each direction.

3.5 Method Comparisons

To compare the various methods, simulated data sets were generated for each surface. For each data set, all method were simultaneously applied to find predictions within the design area. As was mentioned before, the correlation parameters in the kriging methods do not have closed form estimators so these parameters are estimated numerically for use in calculating the *ISE*. The experimental designs described in Section 3.3 were used resulting in 7, 13, and 23 distinct design points with 5 additional center point replications respectively for the one-, two-, and three-dimensional cases. There were 1000 simulated data sets created for each case studied in the one- and two-dimensional cases and 500 simulated data sets for each three dimensional case. The average *ISE* indices and median ratio index Γ of these simulations are listed in Tables 8-15.

3.5.1 *ISE* and Γ – Polynomial Surfaces

All of the surfaces in this section consist of a polynomial function with no additional spatial processes.

Surface 1 - Constant

Data generated for surface 1 consists of a constant plus random error, with the random error representing measurement error.

$$\eta(x) = \beta_o$$
$$y = \beta_o + \varepsilon \qquad \varepsilon \sim N(0, \sigma_\varepsilon^2)$$

All four methods are based on correct models in this case but none of them are appropriate as OLS, BC, and UK include more polynomial terms than are necessary and both OK and UK include a structural stationary process that is not present. For the purposes of simulating Surface 1 the standard deviation σ_ε was arbitrarily set at 0.10. Changing this quantity would only affect the magnitudes of the *MSE, IMSE* and *ISE* by a common factor and so would not affect our comparisons.

Table 8 indicates that the OK method has a much lower *ISE* on average than the other three methods, which in turn, result in comparable values of *ISE*. The sharp contrast between OK and the fit of the other methods may come as no surprise as the latter three are based on assumptions of an approximate quadratic structure in the surface of interest and so can easily "overfit" the data due to the relatively small number of design points in *S*. The OK method is the least severely over-parameterized and apparently leads to the best method for this situation.

For all four methods, the median of the ratio Γ is greater than one indicating that all of the estimates of the quality of fit tend to be conservative. Each of the medians is approximately one except for OK, which tends to indicate a much worse fit than is actually achieved.

Table 8 – Average *ISE* [a] and median Γ [b] when surface is a constant $(d = 1,2,3)$

	ISE					Γ			
					Max[c]				
d	OLS	BC	OK	UK	Std Err	OLS	BC	OK	UK
1	.0647	.0648	.0391	.0658	.00100	1.09	1.39	1.53	1.16
2	.0942	.0941	.0489	.0952	.00122	1.03	1.16	1.56	1.08
3	.1382	.1380	.0617	.1393	.00218	1.01	1.06	1.62	1.09

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations for one and two dimensions and 500 for three dimensions.

[b] Median Γ is the median ratio of 1000 simulations for one and two dimensions and 500 for three dimensions.

[c] Maximum Standard Error is the largest standard error of the four average *ISE*'s.

[d] Dark cells represent the method with the minimum average *ISE* .

[e] Light cells represent methods within two standard errors of the minimum average *ISE*.

Surface 2- Quadratic Polynomials

All of the surfaces in this section consist of a quadratic polynomial function with independent random measurement error added to create the generated data.

$$\eta(x) = P_2(x)$$
$$y = P_2(x) + \varepsilon \quad \varepsilon \sim N(0, \sigma_\varepsilon^2).$$

The coefficients used to generate each surface vary depending on the desired shape and value of *d* and each simulation contains low, medium, or high levels of random error, as described in Section 3.2.

Shape 1 (hyperplane)

The true surface is a polynomial function with only the first-order coefficients being non-zero, representing a hyperplane within the design space. OLS, BC, and UK are over-

parameterized relative to the true model, but the OLS method may be anticipated as the most appropriate as it doesn't fit an additional gaussian process as the kriging methods do, or an unnecessary transformation parameter as with BC.

In the one-dimensional case, we see in Table 9 that OLS, BC, and UK methods dominate the OK with respect to *ISE*. The underlying polynomial structure of these three methods clearly gives a superior fit than that of OK. In sharp contrast, are the two and three- dimensional cases where OK is better for all levels of σ_ε with none of the other methods yielding results within two standard errors. One possible explanation for this discrepancy is the relative lack of design points in the higher dimensions combined with the random error, which results in OLS, BC and UK estimates of second-order polynomial coefficients that are substantially different from zero in some simulations.

Table 9 – Average *ISE* [a] and median Γ [b] when surface is a hyperplane $(d = 1,2,3)$

		ISE					Γ			
d	σ_ε	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
1	L	.0157	.0157	.0181	.0159	.000260	1.12	1.12	1.00	1.17
	M	.0329	.0328	.0382	.0332	.000530	1.06	1.07	0.95	1.10
	H	.0481	.0480	.0559	.0491	.000830	1.09	1.07	0.97	1.14
2	L	.0226	.0226	.0264	.0228	.000239	1.03	1.01	1.04	1.08
	M	.0453	.0452	.0535	.0458	.000471	1.01	1.00	1.02	1.08
	H	.0661	.0660	.0809	.0669	.000733	1.03	1.02	1.04	1.10
3	L	.0338	.0338	.0251	.0341	.000393	1.01	1.00	1.06	1.06
	M	.0668	.0667	.0496	.0671	.000745	1.01	1.00	1.07	1.12
	H	.0981	.0980	.0756	.0990	.001158	1.02	1.02	1.06	1.09

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations for one and two dimensions and 500 for three dimensions.

[b] Median Γ is the median ratio of 1000 simulations for one and two dimensions and 500 for three dimensions.

For all of the methods the median value of Γ is approximately one or slightly higher in almost all situations indicating that all of the methods tend to give the experimenter a well-calibrated estimate of the experimental error.

Shape 2 (ridge)

For shape 2, the true surface is a quadratic polynomial function representing a ridge or tube aligned along $x_1 = x_2 = ... = x_d$. While both kriging methods still incorrectly include a gaussian process, the OLS method is now the correct method for analyzing the data.

As would be expected, the OLS method dominates all the other methods with respect to *ISE*. UK performs almost as well as OLS in this situation, remaining within two standard errors for most situations. This might be anticipated as the UK model also contains the properly structured quadratic polynomial. Of interest though, is the relatively poor fit of OK in this case, as this method is optimal or near-optimal in many situations throughout this analysis. A likely explanation for this situation is the assumption of an isotropic correlation structure, indicating an equal correlation in all directions, for all kriging methods in this analysis. As the surface is a ridge, measurements taken along the ridge will be much more similar in value than measurements taken perpendicular to the ridge. This will create difficulties in estimating the correlation parameters in OK, as the assumed model on which it is based is not consistent with the structure of the ridge. Unlike Tables 8 and 9, BC yields a fit substantially different than OLS, being more than two standard errors worse for all levels of σ_ε. BC's additional parameter allows the method to overfit the data by warping the polynomial surface to fit the random error.

The OLS method, the correct choice, gives a median value of Γ very close to one. The Box-Cox transformation results in an optimistic estimate of error while the kriging methods tend to have Γ's close to one.

Table 10 – Average *ISE* [a] and median Γ [b] when surface is a ridge $(d = 2,3)$

		ISE					Γ			
d	σ_ε	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
2	L	.0215	.0237	.0298	.0217	.000280	1.03	0.88	0.96	1.08
	M	.0430	.0475	.0587	.0435	.000552	1.01	0.87	0.96	1.08
	H	.0628	.0676	.0855	.0636	.000812	1.03	0.92	0.97	1.10
3	L	.0322	.0350	.0404	.0323	.000456	1.02	0.90	1.09	1.06
	M	.0652	.0699	.0783	.0636	.000842	1.02	0.90	1.07	1.12
	H	.0930	.1005	.1128	.0939	.001271	1.02	0.92	1.06	1.10

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations for two dimensions and 500 for three dimensions.

[b] Median Γ is the median ratio of 1000 simulations for two dimensions and 500 for three dimensions.

Shape 3 (mound shaped)

Shape 3 is a quadratic surface with a peak at $x_i = 0.5\ \forall i = 1,2,...,d$ and the OLS method is the correct method for this scenario. The resulting fits are essentially the same as the ridge, Table 10, with OLS and UK offering fits substantially superior to both the BC and the OK.

Estimation of the prediction error is also very similar to results obtained with the ridge shaped surface, with OLS and the two kriging methods giving estimates of median Γ approximating one, while BC is optimistic.

Conclusions for Surfaces 1 and 2

If the second order coefficients in the true surface are zero (the situation where OLS is actually based on a correct but over-parameterized model) OK is most often the best method for estimating the surface. In situations where all second-order coefficients in the quadratic polynomial are non-zero, the average *ISE*'s are best for OLS, but almost as good for UK. The

Table 11 – Average *ISE* [a] and median Γ [b] when surface is a mound $(d = 1,2,3)$

		ISE					Γ			
d	σ_ε	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
1	L	.0156	.0190	.0224	.0159	.000260	1.12	0.85	0.97	1.17
	M	.0427	.0365	.0453	.0331	.000540	1.06	0.91	0.94	1.10
	H	.0479	.0512	.0658	.0488	.000823	1.09	0.94	0.94	1.14
2	L	.0222	.0246	.0282	.0224	.000269	1.03	0.88	0.98	1.08
	M	.0444	.0479	.0547	.0449	.000542	1.01	0.91	0.98	1.08
	H	.0648	.0683	.0802	.0657	.000811	1.03	0.94	0.98	1.10
3	L	.0331	.0354	.0380	.0334	.000437	1.01	0.91	1.03	1.06
	M	.0653	.0690	.0718	.0658	.000809	1.01	0.93	1.04	1.12
	H	.0962	.1002	.1040	.0971	.001289	1.02	0.96	1.05	1.09

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations for one and two dimensions and 500 for three dimensions.

[b] Median Γ is the median ratio of 1000 simulations for one and two dimensions and 500 for three dimensions.

fit of BC mimics OLS for the over-parameterized surfaces, but degrades substantially from the correctly parameterized OLS when the surfaces have non-zero quadratic terms.

For all situations, median Γ for OLS is approximately one indicating a well calibrated estimate for the true error. For most of the situations OK and UK give median Γ values that are slightly greater than one indicating conservative performance for $\hat{ISE}$. BC, on the other hand, gave somewhat optimistic estimates for the prediction error.

3.5.2 *ISE* and Γ - Polynomial Surfaces with Gaussian Process

All of the surfaces in this section consist of the sum of a polynomial function and a gaussian process, with the simulated data having the addition of random measurement error.

$$\eta(x) = P_2(x) + Z(x)$$
$$y = P_2(x) + Z(x) + \varepsilon \quad \varepsilon \sim N(0, \sigma^2).$$

The values for the polynomial coefficients, θ, and σ_z used to generate each surface, and the values of σ_ε used to generate measurement error were selected as described in Section 3.2. Each combination of true surface η and measurement error standard deviation σ_ε are now evaluated in nine comparisons corresponding to three levels of σ_z and three levels of θ.

Surface 3 - Constant with Gaussian Process

The surface is a constant with the addition of a gaussian process, corresponding to the assumed model upon which OK is based. It should be noted that the correlation structure was chosen such that the surface should typically not deviate substantially from quadratic polynomial (see Section 3.2), so the OLS method might also be expected to fit the surface adequately in many cases. The magnitude of the standard deviation σ_z was arbitrarily set at 0.10, 0.20, and 0.30 respectively for the low, medium and high levels of correlated gaussian process as no scaling to a polynomial is possible in this situation. A review of Table 12 indicates that, as expected, OK results in the smallest average *ISE* in almost all cases. The only situations in which OK is not the best method are those in which the magnitude of the gaussian process is relative small, which corresponds to random functions of relatively small magnitudes. Under this condition, for the one- and two-dimensional cases, OLS, BC, and UK are superior to OK. In situations where OK is the best method, UK is generally the second best method, although its fit is not within two standard errors of the best method.

The median value of Γ indicates that the kriging methods do a good job of estimating the quality of fit. The OLS and BC methods are substantially optimistic in this regard, with BC only approximating the correct error for small σ_z. This is of interest, as traditional RSM would use these two methods for approximating the surface.

Table 12 –

Average *ISE* [a] and median Γ [b] when surface is a constant plus gaussian process $(d = 1,2,3)$

			ISE					Γ			
d	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
1	L	L	.0900	.0893	.0936	.0906	.001186	0.85	1.16	0.87	0.93
	L	M	.0968	.0958	.0959	.0968	.001268	0.81	1.06	0.90	0.90
	L	H	.1068	.1064	.1013	.1045	.001415	0.78	1.02	0.87	0.94
	M	L	.1356	.1306	.0990	.1241	.001959	0.69	0.68	0.94	0.89
	M	M	.1514	.1473	.1009	.1357	.002238	0.63	0.63	0.94	0.87
	M	H	.1796	.1770	.1092	.1490	.002799	0.60	0.59	0.95	0.95
	H	L	.1870	.1770	.1033	.1468	.002997	0.60	0.61	0.94	0.95
	H	M	.2128	.2053	.1070	.1604	.003429	0.57	0.57	0.94	1.00
	H	H	.2561	.2515	.1175	.1773	.004315	0.54	0.54	0.93	1.06
2	L	L	.1151	.1154	.1238	.1161	.001237	0.91	1.02	0.92	0.98
	L	M	.1240	.1242	.1322	.1250	.001244	0.88	0.97	0.90	0.96
	L	H	.1363	.1369	.1395	.1367	.001269	0.82	0.89	0.90	0.91
	M	L	.1639	.1620	.1393	.1626	.001499	0.77	0.76	0.95	0.92
	M	M	.1860	.1855	.1496	.1822	.001724	0.72	0.77	0.97	0.92
	M	H	.2170	.2167	.1643	.2110	.001948	0.67	0.73	0.98	0.96
	H	L	.2229	.2184	.1517	.2132	.002159	0.69	0.68	0.95	1.04
	H	M	.2548	.2523	.1609	.2386	.002496	0.64	0.63	0.99	1.13
	H	H	.3057	.3053	.1812	.2854	.002946	0.61	0.59	1.02	1.21
3	L	L	.1630	.1640	.1590	.1640	.001850	0.93	0.99	0.97	1.01
	L	M	.1710	.1719	.1698	.1715	.002030	0.91	0.99	0.95	1.03
	L	H	.1875	.1884	.1753	.1878	.001754	0.86	0.90	0.96	0.96
	M	L	.2185	.2175	.1890	.2130	.002024	0.80	0.92	0.97	0.98
	M	M	.2389	.2387	.2043	.2324	.002155	0.78	0.83	0.95	0.91
	M	H	.2739	.2738	.2249	.2618	.002342	0.72	0.78	0.95	0.92
	H	L	.2853	.2830	.2160	.2582	.002618	0.74	0.73	0.98	1.03
	H	M	.3207	.3204	.2351	.2819	.003002	0.71	0.70	0.98	1.03
	H	H	.3795	.3793	.2626	.3224	.003292	0.67	0.67	0.99	1.11

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations for one and two dimensions and 500 for three dimensions.

[b] Median Γ is the median ratio of 1000 simulations for one and two dimensions and 500 for three dimensions.

Surface 4 Quadratic Polynomials with Gaussian Process

Shape1 (hyperplane) with Gaussian Process

The surface is a hyperplane plus a gaussian process. Only UK is based on a correct model, although the polynomial contains unnecessary second order terms. For each value of d, three different levels of random error standard deviation, gaussian process standard deviation, and gaussian process correlation are now investigated. This results in 27 comparisons in each value of d or 81 total comparisons.

Table 13a-c reveals a definite pattern in the relative performance of the methods depending on the contribution of the correlated gaussian process. Surfaces with relatively small values of σ_z and θ, will be referred to as "weak gaussian processes" as the small magnitude and strong correlation result in surfaces which do not deviate substantially from a constant. In the one-dimensional case with a weak gaussian process, the results suggest the use of the BC method, although both the OLS and UK methods yield similar fits. This seems intuitive as the gaussian component adds little to the surface in these instances and the surface is approximately a linear polynomial. As the magnitude of the gaussian process grows, the OK method dominates the other methods, which all yield similar fits. As the random error increases and the contribution from the gaussian process is overwhelmed, we find that the BC method, along with OLS and UK, tends to provide a better fit.

For the two and three-dimensional cases, the OK method performs substantially better than the other methods in all situations examined, typically resulting in *ISE* values more than two standard errors below those associated with the other methods. The most important trend to note is the performance of the UK method, which typically gives a fit equivalent to or better than the OLS and BC methods throughout these situations. Although UK is over-parameterized, it is the only method based on a correct model assumption and as such might have been expected to yield the best fit.

Table 13a –

Average *ISE* [a] and median Γ^b when surface is a line with a gaussian process $(d = 1)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
1	L	L	L	.0222	.0218	.0220	.0223	.000262	0.85	0.86	0.88	0.94
	L	L	M	.0235	.0232	.0234	.0236	.000282	0.81	0.82	0.86	0.90
	L	L	H	.0270	.0267	.0268	.0267	.000330	0.76	0.75	0.86	0.86
	L	M	L	.0333	.0318	.0340	.0310	.000471	0.68	0.69	0.89	0.88
	L	M	M	.0373	.0358	.0246	.0342	.000528	0.66	0.67	0.88	0.92
	L	M	H	.0444	.0430	.0264	.0372	.000627	0.59	0.60	0.83	0.95
	L	H	L	.0456	.0429	.0245	.0368	.000702	0.61	0.63	0.92	0.96
	L	H	M	.0520	.0494	.0256	.0397	.000802	0.57	0.59	0.92	1.04
	L	H	H	.0634	.0610	.0272	.0435	.000956	0.54	0.55	0.88	1.08
	M	L	L	.0353	.0352	.0397	.0360	.000526	1.01	1.03	0.95	1.09
	M	L	M	.0363	.0361	.0399	.0369	.000510	1.01	1.01	0.95	1.07
	M	L	H	.0391	.0389	.0417	.0396	.000516	0.93	0.94	0.93	1.00
	M	M	L	.0441	.0435	.0430	.0441	.000512	0.84	0.84	0.89	0.91
	M	M	M	.0478	.0472	.0450	.0480	.000574	0.81	0.81	0.88	0.91
	M	M	H	.0533	.0526	.0476	.0531	.000649	0.74	0.75	0.84	0.87
	M	H	L	.0545	.0528	.0461	.0535	.000697	0.75	0.77	0.92	0.89
	M	H	M	.0610	.0595	.0484	.0585	.000793	0.69	0.70	0.86	0.85
	M	H	H	.0707	.0694	.0501	.0648	.000938	0.65	0.66	0.87	0.87
	H	L	L	.0509	.0508	.0584	.0519	.000785	1.05	1.05	0.95	1.10
	H	L	M	.0508	.0507	.0582	.0519	.000775	1.06	1.06	0.95	1.12
	H	L	H	.0541	.0540	.0605	.0552	.000765	1.02	1.01	0.93	1.07
	H	M	L	.0574	.0571	.0620	.0583	.000740	0.94	0.94	0.89	0.99
	H	M	M	.0607	.0603	.0637	.0613	.000794	0.92	0.91	0.91	0.99
	H	M	H	.0655	.0652	.0666	.0660	.000772	0.87	0.86	0.89	0.95
	H	H	L	.0657	.0649	.0661	.0665	.000816	0.89	0.88	0.91	0.94
	H	H	M	.0708	.0697	.0669	.0711	.000826	0.82	0.83	0.90	0.90
	H	H	H	.0793	.0784	.0699	.0784	.000972	0.77	0.77	0.89	0.88

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 13b –

Average *ISE* [a] and median Γ[b] when surface is a hyperplane with a gaussian process $(d = 2)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
2	L	L	L	.0273	.0272	.0268	.0275	.000237	0.91	0.90	0.88	1.00
	L	L	M	.0288	.0286	.0284	.0290	.000242	0.88	0.88	0.84	0.97
	L	L	H	.0320	.0319	.0315	.0322	.000261	0.83	0.82	0.79	0.93
	L	M	L	.0386	.0381	.0353	.0384	.000358	0.76	0.76	0.78	0.91
	L	M	M	.0429	.0423	.0385	.0423	.000380	0.72	0.72	0.73	0.93
	L	M	H	.0506	.0500	.0441	.0493	.000432	0.67	0.67	0.69	0.95
	L	H	L	.0519	.0510	.0402	.0503	.000499	0.69	0.68	0.78	1.05
	L	H	M	.0592	.0585	.0448	.0564	.000550	0.65	0.65	0.73	1.15
	L	H	H	.0710	.0703	.0500	.0674	.000622	0.62	0.62	0.72	1.21
	M	L	L	.0477	.0477	.0460	.0483	.000472	1.00	0.99	0.96	1.08
	M	L	M	.0480	.0480	.0461	.0484	.000452	0.98	0.98	0.96	1.05
	M	L	H	.0506	.0505	.0493	.0509	.000482	0.93	0.92	0.90	1.00
	M	M	L	.0543	.0540	.0639	.0547	.000497	0.91	0.90	0.87	1.00
	M	M	M	.0578	.0576	.0568	.0584	.000489	0.87	0.86	0.84	0.97
	M	M	H	.0635	.0632	.0624	.0638	.000525	0.84	0.82	0.80	0.94
	M	H	L	.0655	.0652	.0623	.0655	.000575	0.82	0.81	0.82	0.93
	M	H	M	.0703	.0698	.0667	.0705	.000605	0.79	0.79	0.79	0.91
	M	H	H	.0819	.0814	.0750	.0813	.000680	0.74	0.73	0.74	0.90
	H	L	L	.0690	.0689	.0654	.0697	.000707	1.01	1.01	0.99	1.06
	H	L	M	.0690	.0688	.0652	.0698	.000726	1.00	0.99	1.00	1.07
	H	L	H	.0717	.0717	.0681	.0727	.000719	0.97	0.96	0.95	1.03
	H	M	L	.0743	.0741	.0736	.0751	.000713	0.95	0.94	0.92	1.02
	H	M	M	.0763	.0762	.0752	.0771	.000744	0.94	0.94	0.92	1.01
	H	M	H	.0814	.0811	.0797	.0819	.000693	0.90	0.90	0.87	0.99
	H	H	L	.0819	.0817	.0813	.0826	.000755	0.89	0.89	0.87	0.98
	H	H	M	.0865	.0863	.0856	.0872	.000759	0.88	0.86	0.85	0.96
	H	H	H	.0952	.0949	.0929	.0955	.000784	0.83	0.81	0.80	0.90

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 13c –

Average *ISE* [a] and median Γ [b] when surface is a hyperplane plus gaussian process $(d = 3)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
3	L	L	L	.0386	.0385	.0344	.0388	.000373	0.92	0.92	0.84	1.00
	L	L	M	.0399	.0398	.0360	.0401	.000385	0.92	0.91	0.83	1.00
	L	L	H	.0442	.0440	.0408	.0442	.000411	0.87	0.86	0.76	0.98
	L	M	L	.0517	.0514	.0482	.0503	.000496	0.82	0.80	0.75	0.95
	L	M	M	.0567	.0564	.0530	.0547	.000536	0.77	0.77	0.70	0.94
	L	M	H	.0659	.0655	.0613	.0625	.000579	0.71	0.71	0.64	0.93
	L	H	L	.0681	.0675	.0585	.0612	.000706	0.73	0.73	0.73	0.99
	L	H	M	.0765	.0758	.0643	.0675	.000724	0.70	0.70	0.72	1.00
	L	H	H	.0905	.0899	.0769	.0766	.000828	0.67	0.66	0.66	1.10
	M	L	L	.0693	.0693	.0558	.0696	.000742	1.02	1.02	1.00	1.11
	M	L	M	.0700	.0699	.0573	.0708	.000735	0.98	0.97	0.96	1.08
	M	L	H	.0715	.0714	.0599	.0720	.000740	0.99	0.99	0.93	1.07
	M	M	L	.0768	.0767	.0647	.0774	.000724	0.94	0.94	0.86	1.02
	M	M	M	.0815	.0813	.0737	.0817	.000785	0.90	0.89	0.83	1.00
	M	M	H	.0878	.0876	.0799	.0878	.000798	0.86	0.85	0.79	0.96
	M	H	L	.0899	.0897	.0829	.0895	.000887	0.87	0.86	0.81	0.96
	M	H	M	.0961	.0958	.0894	.0957	.000892	0.83	0.83	0.76	0.93
	M	H	H	.1078	.1073	.1006	.1061	.000969	0.78	0.78	0.71	0.90
	H	L	L	.1015	.1013	.0787	.1024	.001135	1.00	0.99	1.02	1.06
	H	L	M	.1040	.1038	.0815	.1052	.001126	0.99	0.98	1.02	1.05
	H	L	H	.1023	.1021	.0814	.1030	.001142	1.00	1.00	0.99	1.08
	H	M	L	.1059	.1059	.0909	.1071	.001148	0.99	0.98	0.95	1.09
	H	M	M	.1101	.1100	.0950	.1106	.001169	0.97	0.97	0.89	1.04
	H	M	H	.1139	.1137	.0981	.1144	.001127	0.92	0.91	0.88	0.99
	H	H	L	.1172	.1171	.1063	.1177	.001164	0.95	0.92	0.87	1.00
	H	H	M	.1218	.1216	.1090	.1218	.001117	0.90	0.90	0.85	0.99
	H	H	H	.1327	.1325	.1218	.1329	.001250	0.86	0.85	0.80	0.97

[a] Average *ISE* is the average Integrated Squared Error of 500 simulations.

[b] Median Γ is the median ratio of 500 simulations.

In the one-dimensional case it is interesting to note that the median value of Γ for OK and UK is between 0.85 and 1.15, but that the values for OLS and BC are significantly less as the relative magnitude of the gaussian process grows. In the two and three-dimensional cases the median Γ are all approximately one for small σ_z, but these values are less for larger σ_z especially for OLS, BC, and OK, indicating a tendency for optimism as the gaussian process grows in importance. Only UK maintains a median Γ close to one for most situations.

Shape2 (ridge) with Gaussian Process

The surface is a ridge or tube plus the addition of a correlated gaussian process. The UK method is now the correct method for the analysis. Unlike the previous shape, there are only 54 different situations for comparison as $d = 1$ is omitted.

Table 14a reveals a definite, but unexpected, pattern in the performance of the methods. Even though UK is the correct method for all cases, we find no situations where it is numerically best. In the two-dimensional case the OK method is better in the presence of a stronger gaussian process, while the OLS method is the dominant method for situations in which the gaussian process is weak. Additionally, surfaces with large amounts of random error tend to favor the OLS method. Although the UK method (based on the correct model) does not yield the best fit for any of the situations, it does yield a comparable fit to OLS, being within two standard errors for any situation where OLS is best. As was mentioned earlier, the kriging assumption of isotropic correlation may be a poor assumption in the case of a ridge, resulting in inferior fits for OK in many situations.

Table 14a –

Average *ISE* [a] and median Γ [b] when surface is a ridge plus a gaussian process $(d = 2)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
2	L	L	L	.0258	.0276	.0309	.0261	.000252	0.91	0.81	0.95	0.99
	L	L	M	.0279	.0290	.0319	.0275	.000263	0.88	0.80	0.93	0.97
	L	L	H	.0304	.0316	.0336	.0305	.000275	0.83	0.74	0.89	0.93
	L	M	L	.0367	.0373	.0350	.0365	.00034	0.76	0.71	0.89	0.91
	L	M	M	.0407	.0411	.0368	.0402	.000361	0.72	0.67	0.87	0.93
	L	M	H	.0480	.0476	.0406	.0468	.000411	0.67	0.63	0.83	0.95
	L	H	L	.0493	.0489	.0383	.0478	.000474	0.69	0.65	0.86	1.05
	L	H	M	.0562	.0552	.0422	.0535	.000523	0.65	0.62	0.84	1.15
	L	H	H	.0675	.0654	.0464	.0640	.000591	0.62	0.59	0.83	1.22
	M	L	L	.0453	.0491	.0606	.0459	.000547	1.00	0.88	0.93	1.08
	M	L	M	.0456	.0493	.0590	.0460	.000540	0.98	0.88	0.97	1.05
	M	L	H	.0480	.0513	.0607	.0483	.000546	0.93	0.84	0.94	1.00
	M	M	L	.0516	.0545	.0609	.0520	.000525	0.91	0.82	0.94	1.00
	M	M	M	.0549	.0573	.0626	.0555	.000544	0.87	0.80	0.94	0.97
	M	M	H	.0603	.0622	.0662	.0606	.000547	0.84	0.75	0.91	0.94
	M	H	L	.0622	.0636	.0643	.0623	.000549	0.82	0.76	0.91	0.93
	M	H	M	.0661	.0679	.0678	.0670	.000575	0.79	0.75	0.90	0.91
	M	H	H	.0778	.0775	.0726	.0773	.000646	0.74	0.70	0.89	0.90
	H	L	L	.0668	.0702	.0873	.0662	.000816	1.01	0.90	0.97	1.06
	H	L	M	.0655	.0700	.0859	.0663	.000823	0.99	0.89	0.99	1.07
	H	L	H	.0681	.0722	.0881	.0691	.000835	0.97	0.87	0.95	1.03
	H	M	L	.0706	.0747	.0880	.0713	.000817	0.95	0.87	0.96	1.02
	H	M	M	.0725	.0759	.0884	.0732	.000833	0.94	0.86	0.99	1.01
	H	M	H	.0773	.0803	.0922	.0778	.000778	0.90	0.83	0.94	0.99
	H	H	L	.0778	.0806	.0905	.0785	.000814	0.89	0.83	0.95	0.98
	H	H	M	.0822	.0844	.0930	.0829	.000800	0.88	0.81	0.93	0.96
	H	H	H	.0904	.0914	.0970	.0908	.000798	0.83	0.78	0.92	0.90

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 14b –

Average *ISE* [a] and median Γ [b] when surface is a ridge plus a gaussian process $(d = 3)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
3	L	L	L	.0368	.0392	.0420	.0368	.000421	0.92	0.84	1.06	1.00
	L	L	M	.0378	.0404	.0424	.0380	.000410	0.92	0.84	1.07	1.00
	L	L	H	.0419	.0442	.0445	.0419	.000402	0.87	0.79	1.06	0.98
	L	M	L	.0490	.0506	.0463	.0477	.000482	0.82	0.77	1.04	0.95
	L	M	M	.0537	.0549	.0503	.0519	.000519	0.77	0.73	0.97	0.94
	L	M	H	.0625	.0631	.0581	.0593	.000560	0.71	0.69	0.94	0.93
	L	H	L	.0646	.0653	.0543	.0581	.000671	0.73	0.70	0.98	0.99
	L	H	M	.0726	.0727	.0599	.0640	.000686	0.70	0.68	0.92	1.00
	L	H	H	.0858	.0856	.0702	.0726	.000791	0.67	0.64	0.89	1.10
	M	L	L	.0667	.0719	.0781	.0660	.000800	1.02	0.89	1.08	1.11
	M	L	M	.0664	.0722	.0797	.0671	.000909	0.98	0.89	1.06	1.08
	M	L	H	.0678	.0731	.0794	.0683	.000852	0.99	0.87	1.07	1.07
	M	M	L	.0728	.0774	.0821	.0734	.000875	0.94	0.86	1.04	1.02
	M	M	M	.0772	.0815	.0846	.0775	.000831	0.90	0.83	1.03	1.00
	M	M	H	.0830	.0873	.0878	.0832	.000876	0.86	0.82	1.00	0.96
	M	H	L	.0853	.0895	.0867	.0848	.000855	0.87	0.80	1.04	0.96
	M	H	M	.0912	.0945	.0918	.0907	.000891	0.83	0.78	0.98	0.93
	M	H	H	.1022	.1049	.0989	.1006	.000973	0.78	0.74	0.95	0.90
	H	L	L	.0963	.1042	.1131	.0971	.001307	1.00	0.90	1.07	1.06
	H	L	M	.0986	.1071	.1174	.0998	.001232	0.99	0.87	1.04	1.05
	H	L	H	.0970	.1049	.1145	.0977	.001269	1.00	0.90	1.07	1.08
	H	M	L	.1004	.1080	.1188	.1015	.001335	0.99	0.88	1.05	1.09
	H	M	M	.1044	.1118	.1193	.1049	.001300	0.97	0.88	1.04	1.04
	H	M	H	.1080	.1142	.1206	.1085	.001193	0.92	0.85	1.03	0.99
	H	H	L	.1111	.1178	.1225	.1116	.001292	0.95	0.87	1.03	1.00
	H	H	M	.1155	.1211	.1235	.1134	.001223	0.90	0.84	1.02	0.99
	H	H	H	.1258	.1310	.1312	.1260	.001309	0.86	0.80	1.00	0.97

[a] Average *ISE* is the average Integrated Squared Error of 500 simulations.

[b] Median Γ is the median ratio of 500 simulations.

In this two-dimensional case, we see that the median values of Γ for BC and OK indicate the methods are slightly optimistic for weak gaussian processes and substantially optimistic for surfaces where the gaussian processes contribute significantly to the shape of the surface. OLS follows this same pattern but tends to have a good estimate of error for small σ_z. UK gives a better estimate for the modeling error than the other methods in most cases.

Table 14b exhibits a consistent pattern of method preference within each level of σ_ε. As was seen in the two-dimensional case, for situations where σ_z and θ tend to be small and the gaussian process comprises a minor part of the surface, the best method is OLS. As the relative contribution of the gaussian process grows in the formation of the surfaces, the OK method begins to dominate. The level of random error present in the situation affects the methods, with OK dominating in the areas of small random error. As was seen in Table 14a, the UK method yields fits that are within two standard errors of the OLS method, when OLS is optimal. Unlike the two-dimensional situations, the UK method is optimal for a several situations, specifically when the method yielding the smallest *ISE* transitions from OLS to OK. Additionally, UK exhibits a superior fit to OLS in situations where OK is optimal.

The OLS and BC methods are substantially optimistic in the estimation of their modeling error and tend to be more optimistic as the gaussian process becomes more dominant. The kriging methods both perform better in this regard, with the majority of their median values of Γ falling between 0.9 and 1.1.

In the two and three-dimensional cases, OLS and BC continue to produce optimistic estimates of modeling error, while the kriging methods give more accurate estimates of *ISE*.

Shape3 (mound) with Gaussian Process

The final surface is an approximately quadratic surface which is traditionally modeled using OLS and BC methods, but for which UK is actually based on the correct model. Tables 15a and 15b indicate that for $d = 1$ and 2, OLS dominates when σ_z is small and the gaussian

process is highly correlated, BC is better for somewhat larger values of σ_z and less highly correlated gaussian processes and OK is the best for the large values of σ_z and θ where the surface resembles a higher order polynomial. Small amounts of random error favor the OK method whereas large amounts of random error favor the OLS and BC methods. While UK is not optimal in any of the situations, it does offer a fit within two standard errors of OLS when it is optimal.

For small values of σ_z, OLS estimates of model error are well calibrated with the median value of Γ approximating one, but the estimates are optimistic when the magnitude of the gaussian process is larger. The BC method is more optimistic than the OLS, though their median Γ values are similar for large σ_z. The kriging methods tend to be more conservative for small σ_z but follow the same trend of underestimating error as σ_z increases.
The pattern of relative performance is somewhat different but qualitatively similar in the three-dimensional case found in Table 15c. OLS is the best method for small σ_z, except for situations of small random error where the relatively minor contribution of the gaussian process to the shape of the surface is not masked by the random error. For larger σ_z, OK performs better than all other methods. In general OK performs best when the random error is low, OLS is better for the larger values of σ_ε coupled with smaller values of σ_z. While the BC and UK methods are never optimal for any of the three-dimensional situations evaluated here they typically are within two standard errors of the OLS fit.

As was found in the two-dimensional cases, Table 15c indicates that OLS and BC tend to underestimate the modeling error as the magnitude of the gaussian process increases and they substantially underestimate the modeling error with low amounts of random error. The kriging methods are much closer in their estimates with the majority of the median values of Γ being between 0.9 and 1.1.

Table 15a –

Average *ISE* [a] and Median Γ [b] when surface is a mound plus a gaussian process $(d = 1)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
1	L	L	L	.0221	.0211	.0240	.0222	.000279	0.85	0.81	0.92	0.94
	L	L	M	.0234	.0216	.0241	.0235	.000290	0.81	0.81	0.91	0.90
	L	L	H	.0269	.0249	.0248	.0266	.000328	0.76	0.75	0.91	0.85
	L	M	L	.0331	.0272	.0250	.0309	.000469	0.68	0.73	0.94	0.88
	L	M	M	.0371	.0304	.0252	.0340	.000525	0.66	0.68	0.92	0.92
	L	M	H	.0442	.0367	.0267	.0371	.000624	0.59	0.61	0.88	0.95
	L	H	L	.0454	.0367	.0252	.0366	.000699	0.61	0.65	0.92	0.96
	L	H	M	.0518	.0427	.0258	.0396	.000799	0.57	0.60	0.94	1.04
	L	H	H	.0631	.0536	.0271	.0435	.000952	0.54	0.55	0.91	1.08
1	M	L	L	.0352	.0364	.0445	.0358	.000546	1.01	0.93	0.96	1.09
	M	L	M	.0361	.0368	.0447	.0368	.000533	1.01	0.93	0.97	1.07
	M	L	H	.0390	.0388	.0461	.0394	.000554	0.93	0.89	0.95	1.00
	M	M	L	.0439	.0415	.0456	.0439	.000526	0.84	0.85	0.92	0.91
	M	M	M	.0476	.0439	.0469	.0478	.000572	0.81	0.80	0.94	0.91
	M	M	H	.0530	.0483	.0490	.0528	.000646	0.74	0.77	0.89	0.87
	M	H	L	.0543	.0489	.0480	.0533	.000694	0.75	0.78	0.93	0.89
	M	H	M	.0607	.0545	.0492	.0582	.000790	0.69	0.71	0.89	0.86
	M	H	H	.0704	.0634	.0512	.0644	.000934	0.65	0.66	0.89	0.87
1	H	L	L	.0503	.0524	.0665	.0517	.000823	1.05	0.96	0.95	1.10
	H	L	M	.0506	.0523	.0669	.0516	.000820	1.06	0.98	0.94	1.12
	H	L	H	.0539	.0547	.0673	.0550	.000796	1.02	0.95	0.95	1.07
	H	M	L	.0572	.0566	.0676	.0580	.000765	0.94	0.89	0.94	0.99
	H	M	M	.0605	.0589	.0687	.0610	.000804	0.92	0.89	0.92	0.99
	H	M	H	.0652	.0630	.0700	.0657	.000821	0.87	0.85	0.93	0.95
	H	H	L	.0654	.0624	.0693	.0663	.000820	0.89	0.88	0.93	0.94
	H	H	M	.0705	.0662	.0701	.0707	.000827	0.82	0.83	0.94	0.90
	H	H	H	.0789	.0741	.0720	.0781	.000968	0.77	0.77	0.94	0.88

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 15b –

Average *ISE* [a] and median Γ [b] when surface is a mound plus a gaussian process $(d = 2)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
2	L	L	L	.0266	.0275	.0294	.0270	.000256	0.91	0.83	0.99	0.99
	L	L	M	.0282	.0288	.0309	.0284	.000272	0.88	0.82	0.94	0.97
	L	L	H	.0314	.0318	.0326	.0316	.000275	0.83	0.77	0.92	0.93
	L	M	L	.0379	.0363	.0389	.0377	.000351	0.76	0.74	0.89	0.91
	L	M	M	.0421	.0401	.0362	.0416	.000372	0.72	0.70	0.87	0.93
	L	M	H	.0496	.0472	.0402	.0484	.000424	0.67	0.65	0.83	0.95
	L	H	L	.0510	.0473	.0373	.0493	.000490	0.69	0.68	0.87	1.05
	L	H	M	.0581	.0542	.0414	.0553	.000540	0.65	0.65	0.83	1.15
	L	H	H	.0697	.0655	.0460	.0662	.000611	0.62	0.61	0.81	1.21
2	M	L	L	.0468	.0496	.0575	.0474	.000562	1.00	0.88	0.94	1.08
	M	L	M	.0471	.0496	.0559	.0475	.000532	0.98	0.90	1.00	1.05
	M	L	H	.0499	.0516	.0577	.0499	.000551	0.93	0.85	0.95	1.00
	M	M	L	.0533	.0541	.0587	.0537	.000525	0.91	0.84	0.94	1.00
	M	M	M	.0568	.0570	.0603	.0573	.000546	0.87	0.82	0.93	0.97
	M	M	H	.0623	.0619	.0642	.0626	.000542	0.84	0.80	0.90	0.94
	M	H	L	.0643	.0628	.0627	.0643	.000564	0.82	0.79	0.92	0.93
	M	H	M	.0690	.0675	.0659	.0693	.000594	0.79	0.77	0.89	0.91
	M	H	H	.0804	.0781	.0720	.0798	.000668	0.74	0.72	0.87	0.90
2	H	L	L	.0677	.0705	.0819	.0684	.000788	1.01	0.93	1.00	1.06
	H	L	M	.0677	.0702	.0808	.0685	.000812	1.00	0.93	0.98	1.07
	H	L	H	.0704	.0729	.0840	.0714	.000824	0.97	0.91	0.94	1.03
	H	M	L	.0729	.0743	.0838	.0737	.000802	0.95	0.91	0.96	1.02
	H	M	M	.0749	.0763	.0843	.0756	.000808	0.94	0.90	0.99	1.01
	H	M	H	.0798	.0813	.0882	.0804	.000775	0.90	0.86	0.95	0.99
	H	H	L	.0804	.0809	.0868	.0811	.000814	0.89	0.85	0.95	0.98
	H	H	M	.0849	.0844	.0896	.0856	.000804	0.88	0.84	0.94	0.96
	H	H	H	.0934	.0930	.0947	.0938	.000813	0.83	0.79	0.91	0.90

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 15c –

Average *ISE* [a] and median Γ [b] when surface is a mound plus a gaussian process $(d = 3)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
3	L	L	L	.0379	.0390	.0406	.0381	.000400	0.92	0.87	1.02	1.00
	L	L	M	.0391	.0403	.0407	.0393	.000404	0.92	0.86	1.02	1.00
	L	L	H	.0433	.0434	.0434	.0433	.000403	0.87	0.84	0.98	0.98
	L	M	L	.0507	.0498	.0458	.0493	.000487	0.82	0.78	0.97	0.95
	L	M	M	.0556	.0543	.0490	.0537	.000526	0.77	0.76	0.92	0.94
	L	M	H	.0646	.0628	.0565	.0613	.000568	0.71	0.70	0.86	0.93
	L	H	L	.0668	.0639	.0533	.0600	.000692	0.73	0.73	0.91	0.99
	L	H	M	.0750	.0718	.0596	.0662	.000710	0.70	0.69	0.88	1.00
	L	H	H	.0887	.0853	.0701	.0751	.000812	0.67	0.66	0.82	1.10
3	M	L	L	.0679	.0707	.0722	.0683	.000757	1.02	0.94	1.07	1.11
	M	L	M	.0687	.0715	.0744	.0694	.000833	0.98	0.92	1.01	1.08
	M	L	H	.0701	.0729	.0746	.0706	.000804	0.99	0.92	1.04	1.07
	M	M	L	.0753	.0770	.0776	.0759	.000801	0.94	0.89	1.00	1.02
	M	M	M	.0799	.0806	.0793	.0801	.000782	0.90	0.87	1.00	1.00
	M	M	H	.0861	.0861	.0845	.0861	.000814	0.86	0.83	0.97	0.96
	M	H	L	.0882	.0876	.0841	.0877	.000870	0.87	0.83	0.97	0.96
	M	H	M	.0943	.0935	.0889	.0938	.000875	0.83	0.80	0.93	0.93
	M	H	H	.1057	.1044	.0970	.1040	.000950	0.78	0.76	0.91	0.90
3	H	L	L	.0998	.1027	.1038	.1004	.001271	1.00	0.95	1.08	1.06
	H	L	M	.1020	.1058	.1068	.1032	.001194	0.99	0.93	1.04	1.05
	H	L	H	.1003	.1038	.1058	.1010	.001282	1.00	0.93	1.06	1.08
	H	M	L	.1039	.1073	.1093	.1050	.001268	0.99	0.91	1.04	1.09
	H	M	M	.1080	.1108	.1107	.1085	.001218	0.97	0.92	1.04	1.04
	H	M	H	.1117	.1141	.1125	.1122	.001174	0.92	0.88	1.00	0.99
	H	H	L	.1149	.1168	.1163	.1154	.001264	0.95	0.89	1.00	1.00
	H	H	M	.1194	.1193	.1166	.1194	.001202	0.90	0.87	0.99	0.99
	H	H	H	.1301	.1307	.1270	.1303	.001292	0.86	0.83	0.95	0.97

[a] Average *ISE* is the average Integrated Squared Error of 500 simulations.

[b] Median Γ is the median ratio of 500 simulations.

3.6 Results

3.6.1 OLS

The OLS method is commonly used in RSM as it is assumed that the surfaces of interest are nearly quadratic. Throughout these simulations, OLS performed relatively well when the surface was based on a quadratic polynomial with non-zero second order coefficients. In Tables 10 and 11, where the surfaces are quadratic polynomials representing a ridge and a mound, OLS is optimal for all situations. Additionally, the OLS method performed well when these two shapes were enhanced with a gaussian process (see Tables 14a-15c), as long as the gaussian process did not overwhelm the basic quadratic structure of the surface. In general OLS's performance improved relative to the kriging methods, as the amount of random error increased.

The OLS method has a strong tendency to overestimate the quality of its fit whenever the true surface deviates from a quadratic polynomial. This is a two-fold problem in that OLS methods tend to fit poorly and lead to underestimation of probable error when the true surface deviates from the quadratic polynomial.

3.6.2 Box-Cox

A common approach to correcting for slight deviations from quadratic polynomials is the application of the BC method to transform the data. Although only minor transformations were allowed in this analysis, it was assumed that BC would tend to fit better than OLS, as it allows for more flexibility in modeling the surfaces. For situations where OLS was the correct method, the BC method resulted in fits that were substantially worse than OLS. This is most likely due to the BC method overfitting the random error present in the simulations. For situations in which OLS has an incorrect model assumption, the BC method tended to yield a comparable fit to OLS, except for the case of a ridge with a gaussian process, where

BC performed poorly. There was only one situation in this study where the BC method appeared to be valuable in compensating for deviations from quadratic polynomials. In Tables 15a-c, a mound with gaussian process, a general pattern is displayed within each level of σ_ε. Typically the OLS method fits best in the presence of small values of σ_z, the BC method fits better for larger σ_z, and the OK method yields the best fit in the presence of large σ_z. This pattern might be anticipated, as the surface will deviate from a near quadratic polynomial to one dominated by the correlated gaussian process as σ_z increases. BC would be expected to work better than OLS in mildly non-quadratic surfaces. Larger measurement error favors the OLS and BC methods, while higher dimensions favor the use of OK.

As with OLS there is the tendency for BC to have optimistic estimates for modeling error. This is especially true in the presence of a correlated gaussian process and low measurement error, where the BC method incorrectly models the lack of fit from the gaussian process as measurement error.

3.6.3 Universal Kriging

Although OLS methods and BC transformations are the traditional choice for RSM, the initial focus of this analysis was the evaluation of universal kriging. UK with its combination of underlying polynomial and correlated gaussian process appeared to have an advantage over both OLS and BC methods, as it is based on an assumed model which is correct under all situations considered. In the comparisons examined here, UK was almost never the optimal method in any situation, however UK was not substantially worse than the other methods. As the focus of the analysis is to determine if universal kriging is a viable alternative to the traditional RSM, the next section allow for the comparison of UK to OLS and BC (with the exclusion of OK).

Universal kriging versus traditional methods (OLS & BC)

A quick review of the polynomial surfaces, Tables 9-11, indicates that although UK does not yield the minimum error, it is often within 2 standard errors of BC and OLS. This indicates there is not a great loss in accuracy in using the UK methods to model quadratic surfaces. The addition of a gaussian process, Tables 12-15c, illustrate the advantage of the more flexible universal kriging method with slightly non-quadratic surfaces. Typically, UK offers a comparable fit to OLS and BC and outperforms them well in situations where the random error is small and the data closely represent the true surface. Additionally, UK was substantially superior to BC in modeling the ridge. There is no obvious reason for this occurrence, but it should be kept in mind that the designs and surfaces used for each of the dimensions may tend to favor certain methods. In general, the UK method appears to offer reasonable fits similar to traditional methods for quadratic surfaces and superior fits for non-quadratic surfaces, especially in the presence of minimal measurement error.

Unlike OLS and BC, which consistently underestimate modeling error, UK frequently overestimates modeling error.

3.6.4 Ordinary kriging

The previous section noted that contrary to initial expectation, UK rarely yielded the best fit amongst the various methods, primarily because of the superiority of OK. Though OK offers the flexibility of a gaussian process, the lack of structure associated with the assumption of an underlying constant was believed to be detrimental to its functionality, at least in comparison to UK. For ridge and mound surfaces, OK is substantially worse than OLS, resulting in *ISE* values more than two standard errors greater, but when the higher order polynomial terms are zero, and the other methods are over-parameterized, OK performs relatively well. With the addition of a gaussian process, OK becomes a viable method even in ridge and mound situations. As with apparent in the UK analysis, large levels of σ_ε favor the

traditional methods of OLS and BC, while small amounts of error favor OK. In general, OK outperformed the other methods substantially when the underlying polynomial was a constant or linear in nature and performed well for near quadratic polynomials as long as the random error did not grow so large as to mask the deviations from the pure quadratic.

The estimates for modeling error for OK are similar to that of OLS and BC, in that they tend to be optimistic, but the OK estimates were generally closer to the truth than estimates from either methods.

3.6.5 Summary

In general the results of this study indicate that OLS methods are optimal for true quadratic surfaces. For non-quadratic surfaces, the kriging methods performed well, tending to fit better than both OLS and the traditional BC alternative. OK showed strong potential for being the better choice than the traditional methods, as it was the optimal method in many situations and its fit was substantially better than OLS or BC for surfaces substantially different from quadratic polynomials. UK, while not optimal in any situations, tended to offer a fit that was reasonable in most situations. Although UK did not fit as well as OK for surfaces substantially different from polynomials, it offers a good alternative by fitting near quadratic surfaces better than OK, while fitting substantially non-quadratic surfaces better than OLS or BC.

Error estimation for the various methods was fairly consistent across the various shapes. For ridge and mound polynomials, OLS and UK gave slightly conservative estimates with all values being at or slightly above one, while OK had median Γ's approximately equal to one and BC offered optimistic estimates for error. When the higher order polynomial terms were zero, all of the methods gave estimates greater than one. With the addition of the gaussian process, median Γ's for OLS and BC were consistently below one, which is not unexpected as the experimental error is calculated under the incorrect assumption of an unbiased model.

There is a regular pattern to the error estimates within levels of σ_ε with Γ's being closer to one for small σ_z and becoming smaller as σ_z gets larger. Additionally, as the level of σ_ε grows, the median Γ's are closer to one. This indicates that error estimation for OLS and BC is more optimistic as the relative importance of the gaussian process in the data grows. Error estimation for the kriging methods is more difficult to summarize in the presence of the gaussian process. In general, the median Γ values for OK tend be slightly below one and are more consistent across the changing levels of σ_z and σ_ε for a given shape than OLS or BC. UK tends to yield the most conservative estimate of fit, as its median Γ's are consistently higher than the other methods.

CHAPTER 4 FURTHER EVALUATION OF UNIVERSAL KRIGING

The most marked result of the method comparisons in Chapter 3 is the lack of superior fit by the UK method, even when it would seem to be the most appropriate method. Preliminary investigations into the kriging methods, Section 2.3, indicated a strong potential for the UK method to outperform the other methods, while direct simulations revealed very few situations in which the UK method is optimal. One possible explanation for this discrepancy is poor estimation of the correlation parameters in the simulation study. To investigate this issue further, a surface corresponding to the UK model was chosen for detailed investigation.

4.1 Maximum Likelihood Estimation

To study the performance of the UK method when the correlation parameters are estimated via maximum likelihood, 100 simulations were conducted using data generated by a mechanism corresponding to the UK model. For $d = 2$, a uni-modal quadratic polynomial surface was used with a relatively small value of σ_ε (95% probability interval equal to 10% of the polynomial range), a relatively large value of σ_z ($M(\theta, \sigma_z, d)$ equal to 30% of the polynomial range), and a large value of θ ($R(1_d, 0_d) = 0.2$). Of the methods examined, only UK is based on the correct model. Further, this is a surface in which the gaussian process is a dominant component. This corresponds to the ninth row in Table 15b, in which we see that OK, rather than UK, is substantially better than the other methods, all of which are approximately equal in performance. The results of this smaller simulation, Table 16, are consistent with those reported in Chapter 3.

Table 16 – *ISE* [a] and Γ [b] when surface is a mound plus a gaussian process $(d = 2)$

				ISE					Γ			
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	Max Std Err	OLS	BC	OK	UK
2	L	H	H	.0683	.0653	.0448	.0657	.001897	0.63	0.61	0.80	1.26

[a] *ISE* is the average Integrated Squared Error of $n = 100$ simulations.

[b] Γ is the median ratio of $n = 100$ simulations.

4.1.1 Universal Kriging

The fit of UK is almost identical to that of BC and slightly superior to OLS, which would be expected, as the OLS model is purely quadratic while both UK and BC are based on models that can deviate from a quadratic polynomial. For this sample size, the OLS estimated *ISE* values are within 2 standard errors of the other two methods. More interesting is that the fit of UK is substantially inferior to that of OK. This is counterintuitive as UK assumes the correct polynomial structure while OK does not. A review of the estimate pairs of θ and p for UK, shown in Figure 9, indicates a likely source of the relatively poor fit of the UK method.

A large proportion of the MLE's for θ are effectively infinity, corresponding to correlations between distinct points that are relatively close to zero. As the correlation between points drops off, the correlation matrix approaches the identity matrix, which renders the UK method practically equivalent to the OLS method. This occurs in 75 of the 100 samples generated.

Even in situations where UK results in the estimation of a non-zero correlation, the fit of UK is only a slight improvement over OLS. Restricting the analysis to the 25 cases for which θ is not approaching infinity yields an $ISE_{OLS} = 0.0710$ and an $ISE_{UK} = 0.0625$. Although the UK *ISE* is approximately two standard errors better than OLS, UK is still not competitive with the OK method.

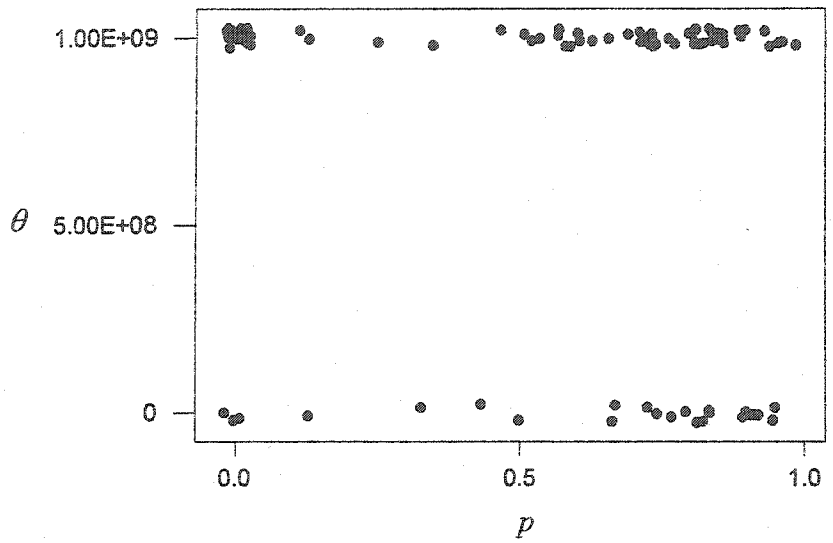

Figure 9 – MLE's of θ and p for UK when $d = 2$, $n = 100$, and $\eta = 0.888 + 0.222x_1 + 0.222x_2 - 0.222x_1x_2 + Z(x)$. (plot includes jitter).
Note $\theta = 10^9$ is an upper bound in the numerical procedure and represents $\hat{\theta}, \hat{p}$ pairings in which $\hat{\theta}$ is approaching infinity.

Trial and error selection of values of θ and p for several data sets indicated that some alternative values provide much smaller *ISE*'s than those found with the MLE's.

4.1.2 Ordinary Kriging

The fit of OK is substantially better than any of the other methods under consideration for this situation. A plot of the maximum likelihood estimate pairs of θ and p for OK, Figure 10, indicates that none of the estimates of θ are moving towards infinity. The values of p are large in this case, 0.99 to 0.999, but this would be expected as σ_ε is relatively small, σ_z is

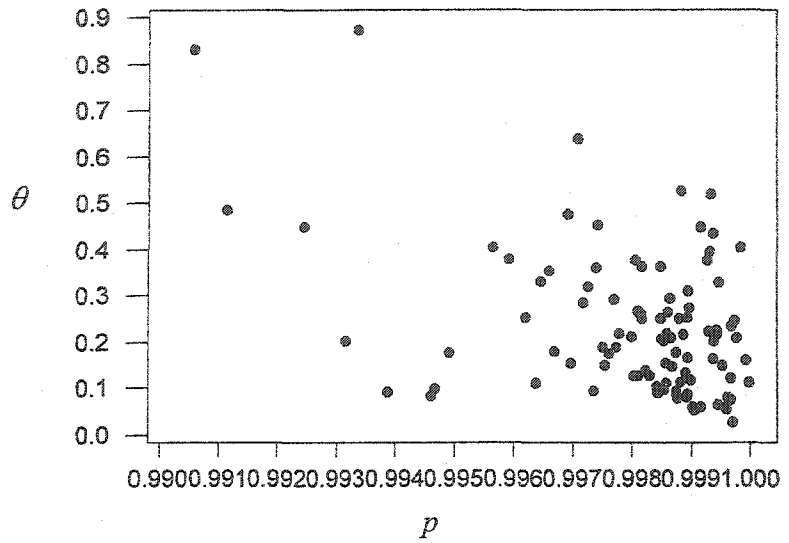

Figure 10 – MLE's of θ and p for OK when $d = 2$, $n = 100$, and $\eta = 0.888 + 0.222x_1 + 0.222x_2 - 0.222x_1x_2 + Z(x)$. (plot includes jitter)

relatively large, and the underlying polynomial is being interpreted as a component of the stationary gaussian process. Due to the relatively small values of $\hat{\theta}$, none of the correlation matrices resemble the identity matrix.

MLE techniques appear to be more stable for OK than UK in this situation. This is likely due to the large number of parameters that must be estimated for UK from a limited number of data points.

4.2 Alternative Estimators

The poor performance of UK based on MLE's of θ and p coupled with the improved fit achieved via trial and error substitutions for these parameters, suggests that other parameter estimators should be considered. A review of the preliminary example (see Figures 2 and 3),

indicates that the UK method dominated the OK method for given values of θ and p. Following this observation, it was hypothesized that substituting the OK MLE's of θ and p into the UK model might result in an improved UK fit. This new method will be called the hybrid (HY) method. Construction of the HY predictors and estimators will be the same as UK, except that the values of $\hat{\theta}$ and $\hat{p}$ used in the calculation of $\hat{V}$ and $\hat{\beta}$ will be the MLEs from OK.

Another method motivated by the UK model (combined polynomial structure and gaussian process) is the separation of the analysis of the two components (e.g. Chiles and Delfiner, 1999). The polynomial trend in the data can first be estimated by OLS. Assuming that this component of the surface is now "known", the OLS model is then used to create predictions for each of the data points, which are then subtracted from the data. The remaining residuals and their spatial relationships are then modeled using OK. This is analogous to de-trending a time series data and will be called residual kriging (RK). This method will result in biased estimates of the gaussian process parameters (Cressie, 1993). Predictions for this method will be made by combining the individual predictions from the separate components. Although this method will result in biased estimates, the interest of this study is not in the estimation of the individual components but in the quality of the predictions of the combined OLS and OK methods. To compare the fit of these new methods to the four previous methods, HY and RK were applied to same data sets evaluated in Tables 8-15.

As both of these two methods are kriging methods, estimation of experimental error will follow the equation developed in Section 3.4:

$$\hat{ISE}_{krig} = \int_{x \in D} \left[c_x' \hat{V} c_x - 2 c_x' \hat{v}_x + \hat{\sigma}_z^2 \right] dx = \Omega \int_{x \in D} \left[c_x' \hat{V} c_x - 2 c_x' \hat{V} \right] dx + \hat{\sigma}_z^2 .$$

For HY, the prediction vector is of the same form as UK

$$c_x' = \left[a'\left(X'\hat{V}^{-1}X\right)^{-1}X' + \hat{v}_x' - \hat{v}_x'\hat{V}^{-1}X\left(X'\hat{V}^{-1}X\right)^{-1}X' \right]\hat{V}^{-1},$$

but all estimates are formed using the OK MLE's for the correlation parameters. The prediction vector of RK is of the same form as OK

$$c_x' = \left[\left(1_N{}'\hat{V}^{-1}1_N\right)^{-1}1_N{}' + \hat{v}_x' - \hat{v}_x'\hat{V}^{-1}1_N\left(1_N{}'\hat{V}^{-1}1_N\right)^{-1}1_N{}' \right]\hat{V}^{-1},$$

but all estimates are based on using OK on the residual responses, after removing the OLS predictions.

4.2.1 Polynomials

We next compare these new methods in situations for which the true surface does not contain a gaussian process (see Tables 8-11). Tables 17-20 contain the *ISE* and median Γ values for HY and RK when applied to each of the respective surfaces. The HY performs poorly in all of these situations. (*ISE* values for OLS, BC, OK, and UK are from Tables 8-11 and are repeated here for convenience.)

In situations where the surface is a constant or multi-dimensional hyperplane and OK is superior, HY yields a fit similar to UK. The lack of improvement in the HY fit may be related to the fact that the HY method uses a quadratic surface as its model and is severely over-parameterized relative to the structure of the true model and susceptible to finding nonexistent patterns in the data. When the underlying surface is a quadratic polynomial and UK's fit is superior to OK, HY yields fits similar to OK. In general, HY rarely produces *ISE* values less than those associated with the UK method and there is no situation within Tables 17-20 where the HY method is optimal.

The error estimates for HY are approximately the same as those for UK for polynomials. The HY median Γ values in Table 17 are slightly higher than those found with UK (see Table 8). In the hyperplane cases, Table 18, the estimates of median Γ are all slightly lower than UK and typically close to one. For polynomial surfaces, Tables 19 and 20, the median estimate of modeling error is substantially optimistic for $d = 2$, though equivalent to BC, and approximately correct for $d = 3$. In general, substituting the OK correlation parameters into the UK model does not appear to adversely affect error estimation.

With respect to *ISE*, RK performed substantially better than HY for the quadratic surfaces. In general, the fit of RK is similar to UK in all situations, indicating that detrending the data and then using OK is approximately equivalent to UK. However, median values of Γ are extremely low indicating optimistic evaluations of modeling error. One possible explanation for this is the common assumption that the polynomial trend surface is known (though actually estimated for detrending purposes), which results in the error calculations for RK being based strictly on the OK component. Additionally, the estimation of the correlation parameters are negatively biased for both $\hat{\theta}$ and $\hat{p}$ resulting an underestimation of the strength of the gaussian component.

Table 17 – Average *ISE* [a] and median Γ [b] when surface is a constant $(d = 1,2,3)$

	ISE						Γ	
d	OLS	BC	OK	UK	HY	RK	HY	RK
1	.0647	.0648	.0391	.0658	.0661	.0651	1.20	0.61
2	.0942	.0941	.0489	.0952	.0969	.0944	1.14	0.45
3	.1382	.1380	.0613	.1393	.1433	.1385	1.11	0.39

[a] *ISE* is the average Integrated Squared Error of $n = 1000$ simulations for one and two dimensions and $n = 500$ for three dimensions.

[b] Γ is the median ratio of $n = 1000$ simulations for one and two dimensions and $n = 500$ for three dimensions.

Table 18 – Average *ISE* [a] and median Γ [b] when surface is a hyperplane $(d = 1,2,3)$

		ISE						Γ	
d	σ_ε	OLS	BC	OK	UK	HY	RK	HY	RK
1	L	.0157	.0157	.0181	.0159	.0176	.0158	1.03	0.62
	M	.0328	.0328	.0382	.0332	.0375	.0330	0.96	0.58
	H	.0481	.0480	.0559	.0491	.0549	.0486	0.97	0.61
2	L	.0226	.0226	.0204	.0228	.0229	.0226	1.02	0.82
	M	.0453	.0452	.0413	.0458	.0460	.0454	1.01	0.83
	H	.0661	.0660	.0609	.0669	.0677	.0663	1.02	0.83
3	L	.0338	.0338	.0251	.0341	.0338	.0339	1.02	0.33
	M	.0668	.0667	.0496	.0671	.0669	.0668	1.02	0.36
	H	.0981	.0980	.0756	.0990	.0990	.0985	1.02	0.35

Table 19 – Average *ISE* [a] and median Γ [b] when surface is a ridge $(d = 2,3)$

		ISE						Γ	
d	σ_ε	OLS	BC	OK	UK	HY	RK	HY	RK
2	L	.0215	.0237	.0298	.0217	.0277	.0215	0.91	0.82
	M	.0430	.0475	.0587	.0435	.0551	.0431	0.92	0.83
	H	.0628	.0676	.0855	.0636	.0804	.0630	0.94	0.83
3	L	.0320	.0350	.0404	.0323	.0385	.0322	1.04	0.33
	M	.0633	.0699	.0783	.0636	.0749	.0634	1.02	0.36
	H	.0930	.1005	.1128	.0939	.1094	.0934	1.02	0.35

Table 20 – Average *ISE* [a] and median Γ [b] when surface is a mound $(d = 1,2,3)$

		ISE						Γ	
d	σ_ε	OLS	BC	OK	UK	HY	RK	HY	RK
1	L	.0156	.0190	.0224	.0159	.0208	.0157	0.95	0.62
	M	.0327	.0365	.0453	.0331	.0424	.0328	0.92	0.58
	H	.0479	.0512	.0658	.0488	.0617	.0484	0.93	0.61
2	L	.0222	.0246	.0282	.0224	.0270	.0223	0.94	0.82
	M	.0444	.0479	.0547	.0449	.0527	.0446	0.96	0.83
	H	.0648	.0683	.0802	.0657	.0770	.0651	0.97	0.83
3	L	.0331	.0354	.0380	.0334	.0376	.0332	0.99	0.33
	M	.0653	.0690	.0718	.0658	.0727	.0655	1.01	0.36
	H	.0962	.1002	.1040	.0971	.1065	.0965	1.03	0.35

[a] *ISE* is the average Integrated Squared Error of $n = 1000$ simulations for one and two dimensions and $n = 500$ for three dimensions.

[b] Γ is the median ratio of $n = 1000$ simulations for one and two dimensions and $n = 500$ for three dimensions.

4.2.2 Gaussian Surfaces

The performance of the hybrid method improves substantially in the presence of a gaussian process. When the true surface is a constant plus a gaussian process and OK is the correct method (see Table 12), we can see from Table 21 that for $d = 1$ the HY method yields a fit similar but superior to OK for all situations. For $d = 2$ the HY method is superior to OK for small σ_ε, while giving fits similar to OK for higher levels of σ_ε. For $d = 3$ the OK method is superior throughout all situations. Although HY does not have the lowest *ISE* for all the situations, it does offer a substantial improvement fit over UK. In the higher dimensions, it is likely that both UK and HY are more susceptible to overfitting the random error than the OK method as these methods require more data points to accurately estimate the polynomial surface.

The error estimation for the HY method yields median Γ's which are approximately equal to that of OK and in many instances are closer to one than OK.

An unusual result for this surface is that HY yields a better fit than OK in many situations for $d = 1,2$ though the OK predictor with known covariance parameters is BLUP (see Section 2.2.2). The most likely reason for this apparent anomaly is the difference between fixed and estimated covariance parameters. The designs in this study used multiple measurements at center points only, which gives unusual influence to single data points at all other locations. The lack of an underlying polynomial may result in OK being more susceptible to over-fitting the random error.

As was seen with strict polynomial surfaces, RK yields fits similar to UK, with the exception for $d = 1$ where the fit of RK degrades relative to UK for large σ_z. Median Γ's for these surfaces exhibit a problem of substantial underestimation of error, which becomes more pronounced as d increases. The only situations where the median Γ approaches one is when σ_z is three times as large as σ_ε and θ is at its largest value, i.e. situations where the gaussian process is a major component in the shape of the surface.

Table 21 –

Average *ISE* [a] and median Γ [b] when surface is a constant plus gaussian process $(d = 1,2,3)$

			ISE						Γ	
d	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
1	L	L	.0900	.0893	.0936	.0906	.0876	.0897	0.99	0.50
	L	M	.0968	.0958	.0959	.0968	.0891	.0963	1.01	0.47
	L	H	.1068	.1064	.1013	.1045	.0949	.1044	0.96	0.50
	M	L	.1356	.1306	.0990	.1241	.0968	.1271	0.94	0.50
	M	M	.1514	.1473	.1009	.1357	.0996	.1391	0.93	0.50
	M	H	.1796	.1770	.1092	.1490	.1079	.1557	0.97	0.63
	H	L	.1870	.1770	.1033	.1468	.1023	.1574	0.94	0.65
	H	M	.2128	.2053	.1070	.1604	.1060	.1733	0.94	0.84
	H	H	.2561	.2515	.1175	.1773	.1163	.1931	0.91	0.90
2	L	L	.1151	.1154	.1238	.1161	.1171	.1151	1.01	0.40
	L	M	.1240	.1242	.1322	.1250	.1241	.1240	1.01	0.40
	L	H	.1363	.1369	.1395	.1367	.1328	.1356	0.99	0.37
	M	L	.1639	.1620	.1393	.1626	.1377	.1616	0.98	0.37
	M	M	.1860	.1855	.1496	.1822	.1476	.1811	0.97	0.38
	M	H	.2170	.2167	.1643	.2110	.1610	.2101	1.00	0.45
	H	L	.2229	.2184	.1517	.2132	.1532	.2127	0.95	0.56
	H	M	.2548	.2523	.1609	.2386	.1619	.2395	0.96	0.80
	H	H	.3057	.3053	.1812	.2854	.1808	.2857	1.02	0.91
3	L	L	.1630	.1640	.1590	.1640	.1680	.1634	1.03	0.33
	L	M	.1710	.1719	.1698	.1715	.1762	.1707	1.06	0.33
	L	H	.1875	.1884	.1753	.1878	.1853	.1872	1.04	0.30
	M	L	.2185	.2175	.1890	.2130	.1953	.2133	0.98	0.31
	M	M	.2389	.2387	.2043	.2324	.2103	.2325	0.97	0.30
	M	H	.2739	.2738	.2249	.2618	.2297	.264	0.96	0.33
	H	L	.2853	.2830	.2160	.2582	.2228	.2608	0.98	0.62
	H	M	.3207	.3204	.2331	.2819	.2376	.2857	1.00	0.71
	H	H	.3795	.3793	.2626	.3224	.2676	.3275	0.99	0.80

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations for one and two dimensions and 500 for three dimensions.

[b] Median Γ is the median ratio of 1000 simulations for one and two dimensions and 500 for three dimensions.

Shape 1

Tables 22a-c review the performance of the HY method when the surface is a hyperplane plus a gaussian process. In this case, the second-order coefficients in the assumed polynomial are zero and the original analysis, Tables 13a-c, indicated OK had lower *ISE* than the other three methods for most situations. The HY method yields a slight improvement on OK for $d = 1$, but HY and OK tend to offer poor fits relative to UK when the level of σ_z is less than the level of σ_ε. The fit of HY degrades relative to OK for larger *d*, but typically the HY method yields fits that are equivalent or better than UK for $d = 1,2,3$.

Error estimation for the HY method is similar to OK but more optimistic than UK, with the median Γ's all below one and consistently below the level of UK.

As was seen in Tables 17-21, RK yields fits similar to UK thus offering better fits than OLS when σ_z is large and σ_ε is small. Median Γ's for RK indicate substantial underestimation of error with values of median Γ's approaching one only when σ_ε is at its minimum and σ_z is at its maximum levels for this study.

Shape 2

Tables 23a-b review the performance of the HY method when the surface is a ridge plus a gaussian process. In this case, all of the first order coefficients in the quadratic polynomial are zero and the original analysis, Tables 14a-b, indicated OLS consistently had lower *ISE* than the other three methods, except when the level of σ_z was much larger than σ_ε. The HY model yields a lower *ISE* for all situations over OK. However HY does not result in a substantial improvements over OK when σ_z is less than σ_ε, the areas where OK fit poorly. This may indicate that the OK estimates of θ and p used in the HY model result in the HY predictors closely resembling the OK predictors.

The error estimation for HY is consistently more optimistic than UK or OK with median Γ's typically being between 0.8 to 1.0

Table 22a -

Average *ISE* [a] and median Γ [b] when surface is a line with a gaussian process $(d = 1)$

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
1	L	L	L	.0222	.0218	.0220	.0223	.0216	.0221	0.88	0.50
	L	L	M	.0235	.0232	.0224	.0236	.0223	.0234	0.85	0.48
	L	L	H	.0270	.0267	.0238	.0267	.0236	.0267	0.84	0.45
	L	M	L	.0333	.0318	.0240	.0310	.0239	.0316	0.88	0.49
	L	M	M	.0373	.0358	.0246	.0342	.0246	.0349	0.86	0.52
	L	M	H	.0444	.0430	.0264	.0372	.0264	.0390	0.82	0.64
	L	H	L	.0456	.0429	.0245	.0368	.0244	.0395	0.90	0.68
	L	H	M	.0520	.0494	.0256	.0397	.0255	.0427	0.88	0.81
	L	H	H	.0634	.0610	.0272	.0435	.0270	.0479	0.86	0.94
	M	L	L	.0353	.0352	.0397	.0360	.0388	.0356	0.96	0.58
	M	L	M	.0363	.0361	.0399	.0369	.0391	.0366	0.96	0.57
	M	L	H	.0391	.0389	.0417	.0396	.0411	.0392	0.91	0.54
	M	M	L	.0441	.0435	.0430	.0441	.0425	.0439	0.88	0.48
	M	M	M	.0478	.0472	.0450	.0480	.0448	.0476	0.86	0.48
	M	M	H	.0533	.0526	.0470	.0531	.0468	.0527	0.83	0.47
	M	H	L	.0545	.0528	.0461	.0535	.0459	.0531	0.89	0.48
	M	H	M	.0610	.0595	.0484	.0585	.0481	.0589	0.85	0.47
	M	H	H	.0707	.0694	.0501	.0648	.0498	.0658	0.84	0.47
	H	L	L	.0509	.0508	.0584	.0519	.0573	.0513	0.96	0.59
	H	L	M	.0508	.0507	.0582	.0519	.0569	.0514	0.97	0.60
	H	L	H	.0541	.0540	.0605	.0552	.0592	.0546	0.95	0.57
	H	M	L	.0574	.0571	.0620	.0583	.0609	.0576	0.90	0.52
	H	M	M	.0607	.0605	.0637	.0613	.0626	.0609	0.91	0.52
	H	M	H	.0655	.0652	.0666	.0660	.0655	.0653	0.88	0.50
	H	H	L	.0657	.0649	.0661	.0665	.0647	.0659	0.93	0.51
	H	H	M	.0708	.0697	.0669	.0711	.0660	.0706	0.90	0.49
	H	H	H	.0793	.0784	.0699	.0784	.0692	.0778	0.88	0.48

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 22b -

Average *ISE* [a] and median Γ [b] when surface is a hyperplane plus a gaussian process ($d = 2$)

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
2	L	L	L	.0273	.0272	.0268	.0275	.0271	.0273	0.92	0.41
	L	L	M	.0288	.0286	.0284	.0290	.0285	.0287	0.90	0.40
	L	L	H	.0320	.0319	.0313	.0322	.0314	.0319	0.84	0.39
	L	M	L	.0386	.0381	.0353	.0384	.0343	.0381	0.82	0.37
	L	M	M	.0429	.0423	.0385	.0423	.0374	.0420	0.78	0.38
	L	M	H	.0506	.0500	.0441	.0493	.0429	.0490	0.73	0.42
	L	H	L	.0519	.0510	.0402	.0503	.0391	.0500	0.82	0.52
	L	H	M	.0592	.0585	.0448	.0564	.0436	.0563	0.76	0.81
	L	H	H	.0710	.0703	.0500	.0674	.0488	.0672	0.73	0.92
	M	L	L	.0477	.0477	.0460	.0483	.0487	.0478	0.99	0.44
	M	L	M	.0480	.0480	.0460	.0484	.0485	.0480	0.99	0.43
	M	L	H	.0506	.0505	.0493	.0509	.0515	.0505	0.92	0.41
	M	M	L	.0543	.0540	.0539	.0547	.0540	.0543	0.92	0.41
	M	M	M	.0578	.0576	.0568	.0584	.0570	.0578	0.88	0.40
	M	M	H	.0635	.0632	.0624	.0638	.0621	.0633	0.85	0.38
	M	H	L	.0655	.0652	.0623	.0655	.0613	.0651	0.87	0.38
	M	H	M	.0703	.0698	.0667	.0705	.0656	.0699	0.84	0.37
	M	H	H	.0819	.0814	.0750	.0813	.0738	.0808	0.78	0.37
	H	L	L	.0690	.0689	.0654	.0697	.0709	.0691	1.01	0.44
	H	L	M	.0690	.0688	.0652	.0698	.0707	.0692	0.99	0.44
	H	L	H	.0717	.0717	.0681	.0727	.0732	.0720	0.97	0.42
	H	M	L	.0743	.0741	.0736	.0751	.0759	.0744	0.96	0.42
	H	M	M	.0763	.0762	.0752	.0771	.0778	.0765	0.95	0.41
	H	M	H	.0814	.0811	.0797	.0819	.0817	.0813	0.91	0.41
	H	H	L	.0819	.0817	.0813	.0826	.0817	.0819	0.91	0.40
	H	H	M	.0865	.0863	.0856	.0872	.0858	.0865	0.88	0.39
	H	H	H	.0952	.0949	.0929	.0955	.0927	.0948	0.86	0.37

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 22c -

Average *ISE* [a] and median Γ [b] when surface is a hyperplane plus a gaussian process ($d = 3$)

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
3	L	L	L	.0386	.0385	.0344	.0388	.0384	.0387	0.93	0.32
	L	L	M	.0399	.0398	.0360	.0401	.0397	.0399	0.93	0.32
	L	L	H	.0442	.0440	.0409	.0442	.0438	.0440	0.87	0.31
	L	M	L	.0517	.0514	.0482	.0503	.0485	.0504	0.85	0.30
	L	M	M	.0567	.0564	.0530	.0547	.0529	.0549	0.81	0.31
	L	M	H	.0659	.0655	.0615	.0625	.0610	.0628	0.75	0.37
	L	H	L	.0681	.0675	.0585	.0612	.0575	.0620	0.81	0.54
	L	H	M	.0765	.0758	.0645	.0675	.0634	.0681	0.77	0.63
	L	H	H	.0905	.0899	.0769	.0766	.0751	.0778	0.74	0.82
	M	L	L	.0693	.0693	.0558	.0696	.0693	.0694	1.02	0.35
	M	L	M	.0700	.0699	.0573	.0708	.0702	.0704	0.98	0.35
	M	L	H	.0715	.0714	.0599	.0720	.0716	.0716	0.99	0.34
	M	M	L	.0768	.0767	.0692	.0774	.0764	.0770	0.94	0.32
	M	M	M	.0815	.0813	.0735	.0817	.0806	.0814	0.91	0.32
	M	M	H	.0878	.0876	.0799	.0878	.0862	.0875	0.87	0.30
	M	H	L	.0899	.0897	.0829	.0895	.0860	.0893	0.90	0.30
	M	H	M	.0961	.0958	.0894	.0957	.0923	.0956	0.86	0.30
	M	H	H	.1078	.1073	.1006	.1061	.1022	.1060	0.81	0.29
	H	L	L	.1015	.1013	.0787	.1024	.1019	.1019	1.01	0.34
	H	L	M	.1040	.1038	.0818	.1052	.1049	.1045	0.99	0.34
	H	L	H	.1023	.1021	.0819	.1030	.1027	.1027	1.00	0.34
	H	M	L	.1059	.1059	.0909	.1071	.1069	.1065	0.98	0.36
	H	M	M	.1101	.1100	.0950	.1106	.1108	.1102	0.96	0.33
	H	M	H	.1139	.1137	.0981	.1144	.1136	.1140	0.92	0.32
	H	H	L	.1172	.1171	.1063	.1177	.1170	.1173	0.96	0.32
	H	H	M	.1218	.1216	.1090	.1218	.1199	.1212	0.93	0.31
	H	H	H	.1327	.1325	.1218	.1329	.1310	.1323	0.88	0.30

[a] Average *ISE* is the average Integrated Squared Error of 500 simulations.

[b] Median Γ is the median ratio of 500 simulations.

Table 23a -

Average *ISE* [a] and median Γ [b] when surface is a ridge plus a gaussian process ($d = 2$)

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
2	L	L	L	.0259	.0276	.0309	.0261	.0283	.0259	0.93	0.41
	L	L	M	.0273	.0290	.0319	.0275	.0292	.0273	0.92	0.40
	L	L	H	.0304	.0316	.0336	.0305	.0308	.0304	0.88	0.39
	L	M	L	.0367	.0373	.0350	.0365	.0343	.0362	0.87	0.37
	L	M	M	.0407	.0411	.0368	.0402	.0329	.0399	0.87	0.38
	L	M	H	.0480	.0476	.0406	.0468	.0360	.0466	0.84	0.42
	L	H	L	.0493	.0489	.0387	.0478	.0344	.0475	0.87	0.52
	L	H	M	.0562	.0552	.0422	.0535	.0376	.0535	0.83	0.81
	L	H	H	.0675	.0654	.0464	.0640	.0414	.0638	0.82	0.92
	M	L	L	.0453	.0491	.0606	.0459	.0564	.0454	0.90	0.44
	M	L	M	.0456	.0493	.0590	.0460	.0546	.0456	0.94	0.43
	M	L	H	.0480	.0513	.0607	.0483	.0566	.0480	0.91	0.41
	M	M	L	.0516	.0545	.0609	.0520	.0564	.0514	0.93	0.41
	M	M	M	.0549	.0573	.0626	.0555	.0579	.0549	0.90	0.40
	M	M	H	.0603	.0622	.0662	.0606	.0610	.0601	0.88	0.38
	M	H	L	.0622	.0636	.0643	.0623	.0592	.0618	0.91	0.38
	M	H	M	.0667	.0679	.0678	.0670	.0623	.0664	0.90	0.37
	M	H	H	.0778	.0775	.0726	.0773	.0664	.0768	0.87	0.37
	H	L	L	.0653	.0702	.0873	.0662	.0821	.0657	0.92	0.44
	H	L	M	.0653	.0700	.0859	.0663	.0800	.0657	0.96	0.44
	H	L	H	.0681	.0722	.0881	.0691	.0827	.0684	0.92	0.42
	H	M	L	.0706	.0747	.0880	.0713	.0822	.0707	0.93	0.42
	H	M	M	.0725	.0759	.0884	.0732	.0833	.0727	0.95	0.41
	H	M	H	.0773	.0803	.0922	.0778	.0856	.0773	0.91	0.41
	H	H	L	.0778	.0806	.0905	.0785	.0842	.0778	0.94	0.40
	H	H	M	.0822	.0844	.0930	.0829	.0863	.0822	0.93	0.39
	H	H	H	.0904	.0914	.0970	.0908	.0901	.0901	0.91	0.37

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 23b -

Average *ISE* [a] and median Γ [b] when surface is a ridge plus a gaussian process $(d = 3)$

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
3	L	L	L	.0366	.0392	.0420	.0368	.0396	.0367	0.99	0.32
	L	L	M	.0378	.0404	.0424	.0380	.0400	.0378	1.02	0.32
	L	L	H	.0419	.0442	.0445	.0419	.0422	.0417	0.99	0.32
	L	M	L	.0490	.0506	.0465	.0477	.0470	.0478	0.99	0.30
	L	M	M	.0537	.0549	.0503	.0519	.0470	.0520	0.95	0.31
	L	M	H	.0625	.0631	.0561	.0593	.0524	.0596	0.89	0.37
	L	H	L	.0646	.0653	.0542	.0581	.0501	.0588	0.93	0.54
	L	H	M	.0726	.0727	.0599	.0640	.0556	.0646	0.88	0.63
	L	H	H	.0858	.0856	.0702	.0726	.0642	.0738	0.86	0.82
	M	L	L	.0657	.0719	.0781	.0660	.0746	.0658	1.04	0.35
	M	L	M	.0664	.0722	.0797	.0671	.0765	.0668	1.00	0.35
	M	L	H	.0678	.0731	.0794	.0683	.0759	.0679	1.02	0.34
	M	M	L	.0728	.0774	.0821	.0734	.0783	.0730	0.99	0.32
	M	M	M	.0772	.0815	.0846	.0775	.0802	.0772	1.00	0.32
	M	M	H	.0832	.0873	.0878	.0832	.0832	.0829	0.96	0.30
	M	H	L	.0853	.0895	.0867	.0848	.0814	.0847	0.98	0.30
	M	H	M	.0912	.0945	.0918	.0907	.0863	.0906	0.96	0.30
	M	H	H	.1022	.1049	.0989	.1006	.0929	.1005	0.92	0.29
	H	L	L	.0963	.1042	.1131	.0971	.1097	.0966	1.03	0.34
	H	L	M	.0986	.1071	.1174	.0998	.1132	.0991	1.00	0.34
	H	L	H	.0970	.1049	.1145	.0977	.1104	.0974	1.03	0.34
	H	M	L	.1004	.1080	.1188	.1015	.1144	.1010	0.99	0.36
	H	M	M	.1044	.1118	.1193	.1049	.1150	.1045	1.02	0.33
	H	M	H	.1080	.1142	.1206	.1085	.1162	.1081	0.96	0.32
	H	H	L	.1111	.1178	.1225	.1116	.1183	.1113	0.98	0.32
	H	H	M	.1155	.1211	.1235	.1154	.1188	.1149	0.98	0.31
	H	H	H	.1258	.1310	.1312	.1260	.1255	.1255	0.96	0.30

[a] Average *ISE* is the average Integrated Squared Error of 500 simulations.

[b] Median Γ is the median ratio of 500 simulations.

In the situation of the ridge shaped surface, RK yields a fit similar to UK. As was evident in Tables 22a-c the fit is equivalent to OLS for most situations and superior to OLS when the gaussian process is a major component of the shape. Reasonable error estimation occurs only when σ_z is much larger than σ_ε.

Shape3

The performance of the HY method for the mound plus gaussian process, Tables 24a-c, is very similar to that of the ridge. All of the coefficients of the underlying polynomial are non-zero and the original analysis, Tables 15a-c, indicated a pattern in the method that yielded the lowest *ISE*, which was related to the relative importance of the gaussian component in the surface. The HY method yields an improvement on the OK, but does not offer a good fit when σ_ε is greater than σ_z. The RK method performs well for this shape, offering fits that are equivalent or superior to OLS in most situations and fits that are superior to OK and HY when σ_ε is larger than σ_z.

Error estimation for HY is optimistic and tends to be similar to OK with all values of median Γ being at or below the level of UK, while median Γ's for RK are substantially less than one in most situations.

4.3 Summary

In general, the fit of the HY method appears to be more similar to OK than UK. This would not seem unreasonable as although HY has the same underlying polynomial structure as UK, the OK estimates for θ and p will result in a much stronger correlation structure in the HY gaussian process. The HY method excels in situations similar to OK, each offering a better fit than UK whenever the gaussian process is a substantial component to the shape of the surface. Small levels of σ_ε tend to favor the HY method over UK. Error estimation for both HY and OK tend to be slightly optimistic. The results of this study indicate that

Table 24a -

Average *ISE* [a] and median Γ [b] when surface is a mound plus a gaussian process $(d = 1)$

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
1	L	L	L	.0221	.0211	.0240	.0222	.0225	.0220	0.91	0.50
	L	L	M	.0234	.0216	.0241	.0235	.0228	.0233	0.91	0.48
	L	L	H	.0269	.0240	.0248	.0266	.0253	.0266	0.89	0.45
	L	M	L	.0331	.0272	.0250	.0309	.0239	.0315	0.92	0.49
	L	M	M	.0371	.0304	.0252	.0340	.0242	.0347	0.90	0.52
	L	M	H	.0442	.0367	.0267	.0371	.0257	.0389	0.85	0.64
	L	H	L	.0454	.0367	.0252	.0366	.0243	.0393	0.90	0.68
	L	H	M	.0518	.0427	.0258	.0396	.0248	.0425	0.91	0.81
	L	H	H	.0631	.0536	.0271	.0435	.0263	.0477	0.88	0.94
	M	L	L	.0352	.0364	.0445	.0358	.0417	.0355	0.96	0.58
	M	L	M	.0361	.0368	.0447	.0368	.0418	.0364	0.96	0.57
	M	L	H	.0390	.0388	.0461	.0394	.0435	.0390	0.94	0.54
	M	M	L	.0439	.0413	.0456	.0439	.0433	.0437	0.91	0.48
	M	M	M	.0476	.0439	.0469	.0478	.0452	.0474	0.89	0.48
	M	M	H	.0530	.0483	.0490	.0528	.0473	.0525	0.85	0.47
	M	H	L	.0543	.0489	.0480	.0533	.0463	.0529	0.91	0.48
	M	H	M	.0607	.0545	.0492	.0582	.0476	.0586	0.85	0.47
	M	H	H	.0704	.0634	.0512	.0644	.0495	.0655	0.86	0.47
	H	L	L	.0507	.0524	.0665	.0517	.0627	.0511	0.95	0.59
	H	L	M	.0506	.0523	.0669	.0516	.0630	.0511	0.93	0.60
	H	L	H	.0539	.0547	.0673	.0550	.0632	.0543	0.95	0.57
	H	M	L	.0572	.0566	.0676	.0580	.0639	.0574	0.91	0.52
	H	M	M	.0605	.0589	.0687	.0610	.0651	.0606	0.92	0.52
	H	M	H	.0652	.0630	.0700	.0657	.0669	.0651	0.91	0.50
	H	H	L	.0654	.0624	.0693	.0663	.0663	.0656	0.93	0.51
	H	H	M	.0705	.0662	.0701	.0707	.0674	.0703	0.91	0.49
	H	H	H	.0789	.0741	.0720	.0781	.0693	.0775	0.89	0.48

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 24b -

Average *ISE* [a] and median Γ [b] when surface is a mound plus a gaussian process ($d = 2$)

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
2	L	L	L	.0268	.0275	.0294	.0270	.0280	.0269	0.96	0.41
	L	L	M	.0282	.0288	.0309	.0284	.0292	.0283	0.93	0.40
	L	L	H	.0314	.0318	.0326	.0316	.0311	.0313	0.88	0.39
	L	M	L	.0379	.0363	.0339	.0377	.0320	.0374	0.89	0.37
	L	M	M	.0421	.0401	.0362	.0416	.0339	.0412	0.85	0.38
	L	M	H	.0496	.0472	.0402	.0484	.0375	.0481	0.81	0.42
	L	H	L	.0510	.0473	.0375	.0493	.0350	.0491	0.86	0.52
	L	H	M	.0581	.0542	.0414	.0553	.0390	.0552	0.81	0.81
	L	H	H	.0697	.0655	.0460	.0662	.0432	.0659	0.81	0.92
	M	L	L	.0468	.0496	.0575	.0474	.0546	.0469	0.94	0.44
	M	L	M	.0471	.0496	.0559	.0475	.0529	.0471	0.98	0.43
	M	L	H	.0496	.0516	.0577	.0499	.0552	.0496	0.94	0.41
	M	M	L	.0533	.0541	.0587	.0537	.0555	.0535	0.94	0.41
	M	M	M	.0568	.0570	.0603	.0573	.0578	.0568	0.91	0.40
	M	M	H	.0623	.0619	.0642	.0626	.0613	.0621	0.89	0.38
	M	H	L	.0643	.0628	.0627	.0643	.0598	.0639	0.91	0.38
	M	H	M	.0690	.0675	.0659	.0693	.0625	.0686	0.89	0.37
	M	H	H	.0804	.0781	.0720	.0798	.0686	.0793	0.85	0.37
	H	L	L	.0677	.0705	.0819	.0684	.0789	.0679	0.97	0.44
	H	L	M	.0677	.0702	.0808	.0685	.0773	.0679	0.99	0.44
	H	L	H	.0704	.0729	.0840	.0714	.0806	.0706	0.94	0.42
	H	M	L	.0729	.0743	.0838	.0737	.0803	.0731	0.96	0.42
	H	M	M	.0749	.0763	.0843	.0756	.0813	.0751	0.96	0.41
	H	M	H	.0798	.0813	.0882	.0804	.0842	.0798	0.93	0.41
	H	H	L	.0804	.0809	.0868	.0811	.0831	.0804	0.94	0.40
	H	H	M	.0849	.0844	.0896	.0856	.0862	.0849	0.92	0.39
	H	H	H	.0934	.0930	.0947	.0938	.0908	.0930	0.88	0.37

[a] Average *ISE* is the average Integrated Squared Error of 1000 simulations.

[b] Median Γ is the median ratio of 1000 simulations.

Table 24c -

Average *ISE* [a] and median Γ [b] when surface is a mound plus a gaussian process ($d = 3$)

				ISE						Γ	
d	σ_ε	σ_z	θ	OLS	BC	OK	UK	HY	RK	HY	RK
3	L	L	L	.0379	.0390	.0406	.0381	.0393	.0379	0.97	0.32
	L	L	M	.0391	.0403	.0407	.0393	.0397	.0391	0.98	0.32
	L	L	H	.0433	.0434	.0434	.0433	.0434	.0432	0.95	0.31
	L	M	L	.0507	.0498	.0458	.0493	.0441	.0494	0.94	0.30
	L	M	M	.0556	.0543	.0499	.0537	.0478	.0538	0.90	0.31
	L	M	H	.0646	.0628	.0565	.0613	.0540	.0616	0.85	0.37
	L	H	L	.0668	.0639	.0538	.0600	.0514	.0608	0.90	0.54
	L	H	M	.0750	.0718	.0596	.0662	.0570	.0668	0.85	0.63
	L	H	H	.0887	.0853	.0701	.0751	.0665	.0763	0.82	0.82
	M	L	L	.0679	.0707	.0722	.0683	.0731	.0680	1.02	0.35
	M	L	M	.0687	.0715	.0744	.0694	.0749	.0690	0.97	0.35
	M	L	H	.0701	.0729	.0746	.0706	.0746	.0702	1.01	0.34
	M	M	L	.0755	.0770	.0776	.0759	.0771	.0755	0.99	0.32
	M	M	M	.0799	.0806	.0793	.0801	.0792	.0798	0.98	0.32
	M	M	H	.0861	.0861	.0845	.0861	.0837	.0858	0.93	0.30
	M	H	L	.0882	.0876	.0841	.0877	.0826	.0875	0.97	0.30
	M	H	M	.0943	.0935	.0889	.0938	.0875	.0937	0.93	0.30
	M	H	H	.1057	.1044	.0970	.1040	.0951	.1039	0.90	0.29
	H	L	L	.0996	.1027	.1038	.1004	.1076	.0999	1.02	0.34
	H	L	M	.1020	.1058	.1068	.1032	.1103	.1025	1.01	0.34
	H	L	H	.1003	.1038	.1058	.1010	.1083	.1007	1.03	0.34
	H	M	L	.1039	.1073	.1093	.1050	.1110	.1044	0.99	0.36
	H	M	M	.1080	.1108	.1107	.1085	.1131	.1080	1.00	0.33
	H	M	H	.1117	.1141	.1125	.1122	.1149	.1118	0.97	0.32
	H	H	L	.1149	.1168	.1163	.1154	.1180	.1151	0.98	0.32
	H	H	M	.1194	.1193	.1166	.1194	.1184	.1189	0.97	0.31
	H	H	H	.1301	.1307	.1270	.1303	.1271	.1297	0.94	0.30

[a] Average *ISE* is the average Integrated Squared Error of 500 simulations.

[b] Median Γ is the median ratio of 500 simulations.

although the HY method does offer improvements over OK and UK in certain situations, different estimators of θ and p may be necessary to achieve adequate modeling of near quadratic surfaces.

The RK results indicate that it offers a fit that is equivalent to UK for most situations. The RK method gives fits approximating OLS for strict quadratic polynomials and offers slight improvements for polynomials with gaussian processes. Although this does indicate RK is a viable alternative to UK for prediction purposes, the biases in the estimation of the correlation and error parameters make inference difficult with RK.

4.3.1 Future Work – Residual Kriging Error

Throughout the situations investigated in this study, the fit of RK was very similar to UK, validating the use of this simple ad hoc method over the more complicated UK method for fitting surfaces with drift. However, error estimation for RK was highly optimistic resulting in median Γ's that were typically less than 0.5, except when σ_z was substantial larger than σ_ε. This poor performance in error estimation is likely due to the assumption that the underlying polynomial is known and the resulting bias introduced into estimates of the covariance structure based on the residuals. Many researchers appear to circumvent the issue by using cross-validation to evaluate the fit of the models, rather than *IMSE*. An alternative error calculation will be suggested that follows the form of the prediction vector method used throughout this analysis. This method will calculate MSE and IMSE under the assumption that underlying polynomial is estimated, rather than assuming it is known.

As the predicted value for RK is a linear combination of the predicted value of OLS and predicted values of OK on the residuals of OLS (another linear operation on Y), the prediction for RK is also linear in Y. Combining (2.2) and (2.6) yields the prediction vector for RK:

$$c_x' = \left[a'(X'X)^{-1}X' + \left[\left(1_N{}' V^{-1} 1_N\right)^{-1} 1_N{}' + \nu_x' - \nu_x' V^{-1} 1_N \left(1_N{}' V^{-1} 1_N\right)^{-1} 1_N{}' \right] V^{-1} \right.$$
$$\left. - \left[\left(1_N{}' V^{-1} 1_N\right)^{-1} 1_N{}' + \nu_x' - \nu_x' V^{-1} 1_N \left(1_N{}' V^{-1} 1_N\right)^{-1} 1_N{}' \right] V^{-1} X (X'X)^{-1} X' \right].$$

Using (2.8), $\hat{ISE}$ can now be directly calculated using the prediction vector. It is important to note that the estimations for the correlation parameters will be biased when using this method, even if the experimenter has correctly assumed that the true surface is a polynomial with a gaussian process.

CHAPTER 5 CONCLUSIONS

The focus of this dissertation is the evaluation of kriging methods as an alternative to ordinary least squares or Box-Cox transformations in modeling near quadratic surfaces. In Chapter 2, *IMSE* is used to show that in the case of lack of fit caused by higher order polynomial terms, there is strong potential for the kriging methods to outperform OLS. In order to directly compare the methods, four distinct surfaces shapes were compared in three dimensions for varying levels of random error and gaussian process.

The results in Chapter 3 indicate that no method is preferred for all situations. In general, OLS fits best for pure quadratics and in situations where σ_ε is large. In situations where the gaussian process is a major component of the surface shape, OK offers a better fit, especially when σ_ε is small. There were almost no situations in this study where UK yielded the lowest *ISE*, but UK often was within two standard errors of the method that did yield the lowest *ISE*. The BC method did result in slight improvements over OLS in the presence of a gaussian process, but often yielded substantially worse results when the surface was strictly polynomial.

In Chapter 4, the performance of UK was investigated further, determining that the MLE's of the correlation parameters may be unstable for the designs and shapes investigated here. The alternative HY method was proposed which substitutes OK correlation estimates into the UK model. The fit of this method was superior to OK in most situations, but did not maintain the UK characteristic of offering a reasonable fit in a diverse range of situations. There is some indication that alternative methods for estimating the correlation parameters in UK could offer better fits, but that substitution of the OK parameter estimates does not greatly improve the performance of UK. Also discussed in Chapter 4 is the common practice of detrending spatial data. The RK method yielded fits that were similar to UK throughout the study, indicating the two methods were approximately equivalent. Error estimates for this

method were severely optimistic, most likely due to the biased correlation parameter estimates that result from the detrending. An alternative measure of error was suggested for future investigation.

Error estimation for the original four methods was relatively good for strict polynomials with median Γ's for all methods being approximately one. With the addition of a gaussian process to the surfaces, the error estimation for these methods was optimistic. This was especially true for OLS and BC, which for surfaces where the gaussian process was a major component commonly yielded median Γ's of 0.6 to 0.8.

The designs considered in this study are traditional RSM designs, or slightly larger plans of similar geometric structure. These may not be particularly effective for spatial modeling. Alternative data collection methods, including reallocation of center point measurements, could result in more situations where kriging is superior to OLS.

Evaluation of the methods in this study was primarily conducted by comparing the fits for various surfaces, *ISE*, and the estimated error of these fits, $\hat{ISE}$, through the ratio Γ. While these measures do give a good indication of how well the fitted surfaces conform to the true surfaces, they may not be the preferred method of evaluating the methods. As practitioners of RSM are typically interested in process maximization, measures relating the quality of fit at the maximum, such as distance from the estimated maximum to the true maximum, may be more informative.

In general, the results of this study indicate that OK offers a good alternative to OLS when the surface is not strictly quadratic, especially when measurement error is relative small compared to the range of the surface. UK while rarely offering the best fit, tends to offer a reasonable fit in a broad range of surfaces and the determination of alternative estimators for the correlation parameters may improve the model substantially.

BIBLIOGRAPHY

Atkinson, A. C. and Donev, A. N. (1996). Optimum Experimental Design. Clarendon Press, Oxford.

Box, G. E. P. and Benken, D. W. (1960) Some new three level designs for the study of quantitative variables. Technometrics, 2, 455-475.

Box, G. E. P. and Cox, D. R. (1964) An analysis of transformations. Journal of Royal Statistical Society, B26, 211-243.

Box, G .E. P. and Draper, N. R. (1987). Emperical Model-building and Response Surfaces. John Wiley & Sons, New York.

Box, G. E. P. and Wilson, K. B. (1951) On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, B13, 1-38.

Carr, J. R. (1995). Numerical Analysis for the Geological Sciences. Prentice Hall, New Jersey.

Chiles, J. and Delfiner, P. (1999). Geostatistics Modeling Spatial Uncertainty. John Wiley & Sons, New York.

Cressie, N. A. C. (1989). Geostatistics. The American Statistician, Vol. 43, No. 4, 197-202.

Cressie, N. A. C. (1993). Statistics for Spatial Data. John Wiley & Sons, New York.

Huijbregts, CH.J. and Matheron, G. (1971). Universal Kriging. The Canadian Institute of Mining and Metallurgy. Special Vol. 12, 159-169.

Journel, A.G. and Huijbregts, CH.J. (1978). Mining Geostatistics. Academic Press, London.

Khuri, A. and Cornell, J.A. (1996). Response Surfaces: Designs and Analyses. Marcel Dekker Inc., NewYork.

Koehler, J.R. and Owen, A.B. (1996). Computer Experiments. Handbook of Statistics, S. Ghosh and C.R. Rao, eds., Elsevier Science, New York, Vol. 13, 261-308.

Myers, R.H. and Montgomery D.C. (1995). Response Surface Methodology. John Wiley & Sons, New York.

O'Connell, M.A. and Wolfinger, R.D. (1997). Spatial Regression Models, Response Surfaces, and Process Optimization. Journal of Computational and Graphical Statistics, Vol. 6, No. 2, 224-241.

Ripley, B.D. (1981). Spatial Statistics. John Wiley & Sons, New York.

Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn,H.P. (1989). Design and Analysis of Computer Experiments. Statistical Science, Vol. 4, No. 4,409-435.

Simpson, T.W. (1998). Comparison of Response Surface and Kriging Models in the Multidisciplinary Design of an Aerospike Nozzle. ICASE Report No. 98-16.