Taylor Kriging Metamodeling For Simulation Interpolation,

Sensitivity Analysis And Optimization

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

_____
Heping Liu

Certificate of Approval:

_____   _____
Alice E. Smith   Saeed Maghsoodloo, Chair
Professor   Professor
Industrial and Systems Engineering   Industrial and Systems Engineering

_____   _____
Tin-Yau Tam   George T. Flowers
Professor   Dean
Mathematics and Statistics   Graduate School

Taylor Kriging Metamodeling For Simulation Interpolation,

Sensitivity Analysis And Optimization

Heping Liu

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

Auburn, Alabama
May 09, 2009

Taylor Kriging Metamodeling For Simulation Interpolation,

Sensitivity Analysis And Optimization

Heping Liu

_____

Signature of Author

_____

Date of Graduation

Heping Liu, son of Minglai Liu and Shuai Yuan who have three children together with two daughters, Minhua Liu and Minhong Liu, was born in Nanchang, Jiangxi Province, China. Heping graduated with a Bachelor of Science degree in Mathematics from Nanchang University in July, 1999. In July, 2003, he received a Master degree in Management Sciences and Engineering from Chinese Academy of Sciences. He entered the Graduate School at Auburn University in August, 2004 and earned a Master degree in Industrial and Systems Engineering in December, 2007. On October 8, 2006, Heping married Yanli Chen. Together they have one daughter born on June 14, 2008.

Dissertation Abstract

Taylor Kriging Metamodeling For Simulation Interpolation,

Sensitivity Analysis And Optimization

Heping Liu

Doctor of Philosophy, May 09, 2009
(M.S., Auburn University, 2007)
(M.S., Chinese Academy of Sciences, 2003)
(B.S., Nanchang University, 1999)

197 Typed Pages

Directed by Saeed Maghsoodloo

The dissertation explores Kriging metamodeling for the interpolation, sensitivity analysis and optimization of simulation models with costly computational or economic expenses. The key theoretical contribution is that a novel Kriging model based on Taylor expansion is developed and named Taylor Kriging (TK) where Taylor expansion is used to identify a drift function. Taylor expansion has excellent nonlinear function approximation capabilities, thus enhancing the interpolation potentials of Kriging. Another contribution is the use of sample standard deviation as the metric for influence distance of covariance, which makes simulations with data of differing magnitudes have a consistent measurement unit.

The partial differentiation equation of Kriging is developed and used together with analysis of variance to assist in sensitivity analysis on a simulation model. A physical simulation case based on cost estimation shows that Kriging sensitivity analysis is effective. While fitting metamodels, the dissertation compares the simulation interpolation accuracy

of Kriging with those of regression and artificial neural networks. A significant feature is that the comparison considers multicollinearity, heteroscedasticity and specification error.

A novel simulation optimization algorithm named SOAKEA is created. SOAKEA integrates Kriging with evolutionary algorithms to optimize simulation models with costly runtime expenses. The properties of SOAKEA are investigated, and some important empirical conclusions about parameter settings are obtained. Several typical multimodal benchmark functions are used to test and compare SOAKEA with other well-known metaheuristics. The results indicate that SOAKEA is a promising optimization tool for simulation models with expensive running costs.

The Kriging software is developed in order to satisfy application needs. The software is multi-functional and user-friendly. The software can be used not only for simulation interpolation but also for wider interpolation applications.

Style manual or journal used <u>Journal of Approximation Theory (together with the style</u> <u>known as "aums"). Bibliograpy follows van Leunen's *A Handbook for Scholars.*</u>

Computer software used <u>The document preparation package TeX (specifically LaTeX)</u> <u>together with the departmental style-file `auphd.sty`, Visual C++ 6.0, Visual Basic 6.0,</u> <u>Microsoft Excel, Minitab, and Matlab</u>

α_i A component of Vector $\boldsymbol{\alpha}$

β_i A component of Vector $\boldsymbol{\beta}$

$\boldsymbol{\alpha}$ The covariance coefficient vector of Dual Kriging

$\boldsymbol{\beta}$ The base-function coefficient vector of Dual Kriging

$\boldsymbol{\lambda}$ $[\lambda_1 \cdots \lambda_N]^T$

$\boldsymbol{\xi}$ The unknown point in Taylor expansion

$\chi^2(i)$ Chi-square distribution with the degree of freedom equal to i

η_i The regression coefficient of the ith term in the LLR model

$\gamma(r)$ The variogram with the distance r between two points

$\hat{\phi}$ The estimator of ϕ

λ_i The coefficient of $Z(\mathbf{X}_i)$ in the estimator of $Z(\mathbf{X})$

$\mu(\mathbf{X})$ A drift function (i.e., a process mean)

ω_{i+1} An assumed linear correlation coefficient between points Z_i and Z_{i+1}

ϕ The regression coefficient of Variable PC in the LLR model

$\pi_j(i)$ The ith integer of the jth permutation in the LHS algorithm

ρ_{ij} The correlation coefficient between variables $Z(\mathbf{X}_i)$ and $Z(\mathbf{X}_j)$

σ The standard deviation

σ^2 The variance of a normal distribution

σ_e The process standard deviation of interpolation errors

$\mathbf{C}(\mathbf{X})$ The partition matrix formed by the covariances of $\mathbf{X}$ and $\mathbf{X}_i$

$\mathbf{E}$ A partition matrix in the inverse of Kriging coefficient matrix

$\mathbf{F}$ The partition matrix formed by the base functions in Kriging coefficient matrix

$\mathbf{f}(\mathbf{X})$ $[f_l(\mathbf{X}) \cdots f_M(\mathbf{X})]^T$

$\mathbf{G}$ A partition matrix in the inverse of Kriging coefficient matrix

$\mathbf{H}$ A partition matrix in the inverse of Kriging coefficient matrix

$\mathbf{K}$ Kriging coefficient matrix

$\mathbf{P}_i$ The best position vector found by the ith particle in particle swarm optimization

$\mathbf{Q}$ A partition matrix in the inverse of Kriging coefficient matrix

$\mathbf{S}$ The partition matrix formed by covariances

$\mathbf{U}$ $[u_1 \cdots u_M]^T$

$\mathbf{V}_i$ The velocity vector of the ith particle in particle swarm optimization

$\mathbf{X}_0$ The position vector where Taylor expansion is conducted

$\mathbf{X}_i$ The ith position vector

$\mathbf{Z}$ $[Z(\mathbf{X}_1) \cdots Z(\mathbf{X}_N)]^T$

$\varepsilon(\mathbf{X})$ A random error term related to the point $\mathbf{X}$

ε_i A random error term

φ The mean of ε_i

φ_1 The cognition learning component in particle swarm optimization

φ_2 The social learning component in particle swarm optimization

a_l The coefficient of the lth base function

b The exponent in a Power-function variogram

b_i The regression coefficient of the ith term in a multivariable linear regression

$C(r)$ The covariance function of two points with the Euclidean distance equaling r

C_0 The constant coefficient in a covariance function

C_i The covariance between $Z(\mathbf{X}_i)$ and $Z(\mathbf{X})$

D The dimension number of a position vector $\mathbf{X}$

d The influence distance of covariance

D_N The number of points generated by LHS

$DBF(x_i)$ The derivative of Kriging base function with respect to x_i

E_0 The number of simulation inputs generated by computer experimental design in SOAKEA

F F statistic

$F(i, j)$ F distribution with the degrees of freedom equal to i and j

$F(x_1, \cdots, x_D)$ The multivariable linear regression model

$F_{0.05;ij}$ The inverse function of F distribution at the cumulative of 0.95 with numerator degrees of freedom equal to i and denominator degrees of freedom j

F_0 A specific value of F statistic

$f_l(\mathbf{X})$ The lth base function

G_0 A correction strategy in SOAKEA: a correction per G_0 updated generations

H_0 Null hypothesis

H_1 Alternative hypothesis

J A test in statistic

M The number of base functions

m The highest order of Taylor expansion

M_0 The size of an initial population of an evolutionary algorithm

m_0 The number of corrected individuals in SOAKEA

MC_i The model coefficient of the ith base function

N The number of observed points in a fitting set

P P-value

p The constant coefficient in Power-function and Logarithmic variograms

p_{ij} The jth component of $\mathbf{P}_i$

r The Euclidian distance between two points in a vector space

R^2 R-square

R_a^2 Adjusted R-square

r_i The Euclidian distance between $\mathbf{X}$ and $\mathbf{X}_i$

$rand()$ A random disturbance term

S Sample standard deviation

S_{ij} The covariance between $Z(\mathbf{X}_i)$ and $Z(\mathbf{X}_j)$

t t statistic

$t(i)$ t distribution with the degree of freedom equal to i

$U(0,1)$ A uniform distribution in the interval $[0,1]$

U_k^j A uniform variate generated from $U(0,1)$

u_l The Lagrange multiplier

v_{ij} The jth component of $\mathbf{V}_i$

x_i The ith component of the position vector $\mathbf{X}$

$x_i^{'}$ The normalized value of x_i

x_k^j The jth component of the kth position vector

y_a The actual value of dependent variable y

y_p The predicted value of dependent variable y

$Z(\mathbf{X})$ A stochastic function (or process)

Introduction

This chapter first introduces the research background and motivation for the dissertation, illustrates the objectives that the research needs to achieve, and then gives the research methods adopted. Finally, the chapter shows how the dissertation is organized.

1.1 Background and Motivation

Many scientific disciplines use simulation to describe and analyze real-world complicated systems in order to gain insights into such systems. Due to complexity of systems, analytical modeling methods generally become infeasible. Although a simulation model is only an abstraction of a real-world complicated system, it can consider numerous details and constraints in the system, which enables it to closely resemble the real system and thus become very useful.

However, simulation models are often computationally expensive. Under resource constrained environments, obtaining enough simulation outputs to satisfy the needs such as performing sensitivity analysis and finding optimal inputs becomes very difficult. Thus, it is desirable and necessary to use metamodeling methods (defined below) to establish a functional relationship for known simulation inputs and outputs and then use the functional relationship to estimate likely outputs for the inputs which have not yet been simulated. The estimated output values could be used for sensitivity analysis, optimization, and decision making.

According to Wikipedia encyclopedia [56], metamodeling is to analyze, construct and develop the frames, rules, constraints, models, and theories applicable that are useful for modeling a predefined class of problems. Metamodeling was initially developed from neuro-linguistics in the beginning of the twentieth century [79]. Kleijnen [73] discusses how to use metamodeling to perform sensitivity analysis, optimization, and validation and verification. From the perspective of simulation modeling, Kleijnen *et al.* [72] define a metamodel as an approximation of the true input and output function implicitly defined by a given simulation model. They think that a metamodel is much simpler than the underlying simulation model.

There are many different metamodeling methods such as design of experiments, regression, neural networks, and so on, to build a functional relationship of simulation inputs and outputs by bridging the gap between the known and unknown information. Essentially, these methods can be called predictive modeling. Considered from the viewpoint of statistical theory, predictive modeling is generally known as inference.

Kriging is a well-proven geostatistical metamodeling technique. Because of the precise interpolation feature, Kriging has been applied to some areas and achieved success. Kriging is named after Krige, a mining engineer in South Africa. In the 1950's, Krige first applied the technique to the estimation of gold deposits [17, 81]. Kriging estimates are Best Linear Unbiased Estimators (BLUE). By identifying and mathematically representing the underlying functional relationship in known observation points, Kriging can predict the function values of unobserved points.

In this dissertation, Kriging is proposed as a metamodeling method to build a functional relationship for simulation inputs and outputs to realize the precise interpolation and to assist the sensitivity analysis and optimization operations of simulation modeling. The

simulation models considered here are (assumed to be) computationally expensive, and they can be deterministic, stochastic or physical simulation models.

1.2   Research Objectives

The research objectives mainly include three aspects.

First, the dissertation explores Kriging metamodeling for simulation interpolation. The main objective is to use Kriging to interpolate simulation outputs to reduce computational expenses caused by running a simulation model. Kriging is improved by introducing Taylor expansion to act as a base function and sample Standard Deviation (SD) as the metric for influence distance of covariance. Simulation interpolation capabilities of Kriging are compared with those of regression and artificial neural networks. Multicollinearity, heteroscedasticity, and specification errors of metamodeling are emphasized.

The second objective is to use Kriging to assist in sensitivity analysis of a simulation model. The partial differentiation equation of Kriging is developed. A sensitivity analysis method based on the partial differentiation equation and Analysis Of Variance (ANOVA) is built for simulation modeling.

The third objective is involved in Kriging metamodeling for simulation optimization. A novel simulation optimization algorithm is created which can integrate Kriging with evolutionary algorithms to search for optimal inputs of a simulation model.

In addition, the multi-functional Kriging application software is developed in order to satisfy wide application needs of Kriging.

1.3 Research Methods

The theoretical research is an important feature. The dissertation theoretically improves Kriging models. A novel Kriging model based on Taylor expansion is developed to enhance the simulation interpolation capabilities of Kriging. In addition, sample standard deviation is used as the metric for influence distance of covariance, thus making different simulation problems have a consistent measurement.

Kriging models are explicit and differentiable, which makes it feasible to use their analytic features to assist the sensitivity analysis on simulation models. A methodology contribution to sensitivity analysis of simulation modeling is that a partial differentiation equation of Kriging is developed and used together with ANOVA to assist sensitivity analysis on simulation models.

Empirical research is another important aspect. The dissertation uses many empirical cases to investigate the interpolation capabilities of Kriging metamodeling in deterministic, stochastic and physical simulations. Kriging is empirically used to analyze the sensitivity of a simulation model to inputs. For the developed simulation optimization algorithm, several simulated benchmark functions are used to test its effectiveness.

Additionally, comparison and contrast is widely used. For a test case, different models or methods are used to implement simulation interpolation or optimization to analyze and compare their advantages and disadvantages.

1.4 Organization of Dissertation

The dissertation is divided into nine chapters, and the organization of contents is as follows.

In Chapter Two, the origins and applications of Kriging are reviewed. The Kriging origins reviewed include its early formulation and development process. The application review of Kriging focuses on simulation interpolation and optimization.

Chapter Three presents the Kriging methodology. The basic principles and structure components of Kriging are introduced. The introduced components include drift functions, covariance functions, neighbor structures, and variance of interpolation errors. The important properties of Kriging are further discussed.

Chapter Four introduces the theoretical contributions of the dissertation. A novel Kriging model based on Taylor expansion is developed and named Taylor Kriging. Its crucial properties are provided. Sample standard deviation is used to define influence distance of covariance in covariance functions of Kriging. A sensitivity analysis method using a partial differentiation equation of Kriging and ANOVA is established.

Chapter Five empirically investigates the interpolation effectiveness of Taylor Kriging for deterministic and stochastic simulation by using some simulated cases based on benchmark functions. The critical insights from the empirical investigation are given.

Chapter Six empirically analyzes how to use the sensitivity analysis method based on Kriging and ANOVA to perform sensitivity analysis on a simulation model. The chosen simulation model is a physical simulation related to cost estimation. The potentials of the sensitivity analysis method are compared with those of regression.

Chapter Seven establishes a novel hybrid simulation optimization algorithm by integrating Kriging and evolutionary algorithms. The properties and parameter settings of algorithm are analyzed in detail. The empirical applications of the algorithm to simulation optimization are investigated.

Chapter Eight introduces the developed Kriging application software. The software has a user-friendly graphic interface. The details about how to manipulate the multi-functional application software are demonstrated.

Chapter Nine concludes the dissertation and illustrates the future research directions.

CHAPTER 2

LITERATURE REVIEW

This chapter reviews the origins of Kriging and its applications in simulation interpolation and optimization. The review on the development process of Kriging will be presented in the third chapter.

2.1 Origins of Kriging

The roots of Kriging go back to the exploration research of Wold [158] and Kolmogorov [77, 78] and Wiener [157], although their research is directly related to regression instead of Kriging modeling. In 1951, Krige [81], a mining engineer in South Africa, first developed the approach therefore bearing his name, and used this approach to gold mining valuation problems.

The term "Kriging", named by Matheron [94], was initially coined as "krigeage" by Pierre Carlier [17]. Matheron [94] describes Kriging as a method which can predict the grade of a panel by computing the weighted average of available samples. He emphasizes that suitable weights should make variance smallest. Krige's own understanding about the term is that it is a multiple regression which is the best linear weighted moving average of the ore grade of an ore block of any size by assigning an optimum set of weights to all the available and relevant data inside and outside the ore block [82]. Ord [111] states that Kriging is a method of interpolation for random spatial processes in the Encyclopedia of Statistical Sciences. According to the opinion of Hemyari and Nofziger [52], Kriging is a form of weighted average, where the weights depend upon the location and structure

of covariance or semivariogram of observed points. The choice of weights must make the prediction error less than that of any other linear sum. A semivariogram is a function used to indicate spatial correlation in observations measured at sample locations [17].

In the early 1960's, Matheron [92, 93] developed the mathematical theory associated with Kriging models. Originally, Kriging was a linear predictor. In later developments in geostatistics, Kriging was extended to nonlinear spatial prediction. Matheron [98] also shows that Kriging and splines are formally equivalent. Dubrule [34] instead investigates the differences between splines and Kriging.

An early comprehensive introduction to the origins of Kriging is given by Cressie [18]. Cressie investigates the origins of Kriging from the perspective of different disciplines and states that Kriging is equal to spatial optimal linear prediction. An interesting historical table of researchers' contributions to Kriging is provided by Cressie. For Table 2.1, Cressie explains that an index of "1" for Component 1 means that all possible covariances are used to define weights, "$\frac{1}{2}$" means only some covariances are used, and "0" means some other weights are used; an index of "1" for Component 2 means that the weighted least-squares estimator of μ is used, "$\frac{1}{2}$" means that another estimator is used and "0" means μ is assumed known; an index of "1" for Component 3 means that prediction is performed in a spatial setting, "$\frac{1}{2}$" means a temporal setting is used and "0" means that neither a spatial nor a temporal setting is considered. Table 2.1 shows that Wold, Kolmogorov and Wiener did some early exploratory work, Krige established Kriging models, and Matheron [18] made comprehensive analysis on Kriging models.

Table 2.1: Historical Table of Researchers' Contributions to Kriging [18]

Research	1938	1941-1949	1951	1956	1952	1963
Wold	$(1,0,\frac{1}{2})$					
Kolmogorov		$(1,0,\frac{1}{2})$				
Wiener		$(1,0,\frac{1}{2})$				
Krige			$(\frac{1}{2},\frac{1}{2},0)$		$(1,\frac{1}{2},0)$	
Thompson				$(1,0,1)$		
Goldberger					$(1,1,0)$	
Matheron					$(1,1,1)$	
Gandin						$(1,1,1)$
Whittle						$(1,1,\frac{1}{2})$
Moritz						$(1,0,1)$
Henderson						$(1,1,0)$

2.2 Overview of Kriging Applications

The applications of Kriging cover some disciplines which range from the classical application fields of mining and geology to soil science, hydrology, meteorology, etc., and recently to engineering design, cost estimation, wireless sensing and networks, simulation interpolation, evolutionary optimization, etc..

The application of Kriging is mainly in geological settings [20, 21, 81, 65, 136]. Kriging is extensively used to produce contour maps [22, 23, 43, 109, 145]. More relevant references can be found in some review papers [33, 122]. Kriging has been used as a means to predict the values of soil attributes at unsampled locations [49, 105, 120, 135, 152, 156]. In the field of hydrology, Kriging has had wide applications [9, 13, 16, 26, 27, 112], and some related review papers are [51, 144, 150]. Additionally, Kriging is applied to meteorology [25].

The literature demonstrates that in recent years Kriging has been applied to engineering design [42, 89, 101, 132, 133], material sciences [35, 68, 91, 143], biomechanical engineering

[3, 28, 11], wireless wave propagation [86, 30, 31, 119, 147, 160, 87, 88], economic sensitivity analysis and cost estimation [151, 12], simulation interpolation [104, 6, 75], and optimization [63, 61, 60].

Next, the dissertation will in detail review the Kriging applications in simulation interpolation and optimization.

2.3 Kriging Applications in Simulation Interpolation

The application of Kriging to simulation interpolation first occurred in the area of deterministic simulation. The classic reference refers to Sacks *et al.* [124]. In this reference, Sacks *et al.* consider the computer experiments which are computationally expensive to run and whose outputs are deterministic. For such computer experiments, they use Kriging to fit an inexpensive but efficient predictor to reduce the computational cost.

In the reference [104], Mitchell and Morris investigate Kriging as an alternative to conventional response surface methodology for use in simulation experiments. Mitchell and Morris treat Kriging as a Bayesian method. They use Kriging to evaluate the importance of input parameters, consider how to use Kriging to optimize a dependent variable, and discuss the Kriging applications in inverse problems according to a simulation experiment on the model of groundwater flow. Although much of their focus is on deterministic simulations, they show how modifications can be made to handle the simulation interpolation with random outputs.

Trochu *et al.* [146] apply Dual Kriging as an interpolation method to simulating the macroscopic mechanic behavior of shape memory alloys. They use Dual Kriging to yield an

explicit equation of any partial cycle inside the main hysteretic domain, thus presenting a general material law for shape memory alloys.

The early application of Kriging to random simulations was proposed by Barton [6]. In his introduction to Kriging, he regards Kriging as a spatial correlation metamodel. He indicates that although the fitting capability of the spatial correlation method is exciting, it is based on a small set of examples, and the more extensive computational comparison of different methods would have to wait for more generally available computer codes. Additionally, Koehler and Owen [76] provide some extra introduction and discussion of the simulation interpolation based on Kriging.

Régnire and Sharov [118] discuss the application of Kriging to the interpolation of the spatial and temporal output data of the simulation model for male gypsy moth flight phenology. Universal Kriging (UK) and multivariate linear regression are compared in their simulation interpolation. According to experimental results, they think that these two methods are nearly equally precise in interpolating the output data of a simulation model; however, Kriging requires more computing time than does regression. But, they indicate that the success of regression may be due to the relatively simple physical processes simulated, and UK offers an alternative in cases where simple polynomial terms cannot mimic more complex response surfaces. As van Beers and Kleijnen [148] later noticed, the multivariate linear regression compared with UK by Régnire and Sharov is a rather complicated metamodel (involving terms of order six), which is perhaps one reason why its resulting prediction accuracy is similar to that of Kriging.

The recent representative work of Kriging applications to simulation interpolation is given by Kleijnen and van Beers [75, 74, 59, 148]. Their work mainly focuses on three areas:

1. Investigate the application of Kriging in random simulation when the variances of simulation outputs are not constant. They show that Ordinary Kriging (OK) is a robust interpolation method and seems not very sensitive to variance heterogeneity [75].

2. Propose a novel experimental design method based on Kriging for deterministic simulation interpolation. The method proposed [74] is sequential and application-driven or customized through cross-validation and jackknife. Kleijnen and van Beers [74] indicate that the method applies to other types of metamodels and stochastic simulation. Van Beers and Kleijnen [59] also establish a customized sequential design for interpolation in random simulation, especially discrete event simulation. The tool for this customization is bootstrapping which estimates the variances of Kriging predictions. The candidate input with the largest bootstrap variance is selected as the next actual input to be simulated. However, it is noticed that although what they discuss is discrete-event simulation, the interpolated variables are still continuous.

3. Discuss the Kriging interpolation for random simulation and in detail illustrate De-trended Kriging that they develop [148].

Additionally, van Beers and Kleijnen [149] review the Kriging application to simulation interpolation. They note that the Kriging interpolation is attractive in deterministic simulation, and the Kriging interpolation in discrete-event simulation has just started. Santner, *et al.* [125] provide some details of the application of Kriging to simulation interpolation. Meckesheimer *et al.* [62] develop the efficient methods to assess the validity of Kriging and other metamodels in simulation interpolation. They investigate computationally inexpensive assessment methods for metamodel validation based on leave-k-out cross validation,

and develop guidelines about how to select k. Based on the results from two sets of test problems, $k = 0.1N$ or the square root of N is recommended for the Kriging metamodel, where N is the number of sample points used to construct the metamodel. Hertog, *et al.* [53] discuss the Kriging variance. They prove that the classic Kriging variance formula widely used in geostatistics is wrong, and the formula underestimates the Kriging expected variance. And they develop a parametric bootstrapping method to estimate the Kriging variance.

The literature review shows that although some researchers applied Kriging to simulation interpolation, their work mainly focused on Simple Kriging (SK) and OK. The drift functions of SK and OK are a zero constant and an unknown nonzero constant, respectively. These drift functions are simple, and it is not possible for them to capture the non-constant mean drift of data in some simulation problems. Note that some references used UK, but the drift function of UK is a general polynomial, and its base function forms are not identified. There exist difficulties in choosing specific base functions for UK. How to give a non-constant drift function for Kriging with specified base function forms deserves exploration.

Similar to computer simulation with costly computational expense, physical simulation has the similar difficulties to obtain data with adequate sample size because of economic and time constraints. Exploring Kriging interpolation for physical simulation has been very few in the literature, and more work is needed in this area.

Although some references compared the interpolation capabilities of Kriging and regression, they did not consider multicollinearity, heteroscedasticity, and specification errors in regression. Methodologically, the comparison is incomplete, and the results can not be

trusted without any suspicion. It is needed to redo the comparison of Kriging and regression with these considerations.

Sensitivity analysis is another critical problem of simulation modeling. Kriging metamodels are explicit and differentiable, and it could be possible to use these analytic features of Kriging to assist sensitivity analysis on simulation models, which has not been seen in the literature. As a result, its exploration is justified.

2.4 Kriging Applications in Optimization

Recently Kriging has been applied to optimization areas. The applications in optimization can be divided into two types. One is that Kriging is used to assist evolutionary optimization by acting as temporal fitness functions; the other is called Sequential Kriging Optimization (SKO) where Kriging itself serves as an optimization tool to search for optimization solutions.

When evolutionary algorithms are used to solve optimization problems, explicit fitness functions may not exist. The application of Kriging in evolutionary algorithms is to build an approximate fitness function so that evolutionary algorithms can use this approximate function to guide further search. The literature indicates that the early research in this area was conducted by Ratle [115, 116]. Ratle proposed a hybrid algorithm by integrating Kriging and a real-coded Genetic Algorithm (GA). The new algorithm has the good approximation performance of Kriging with effective and robust evolutionary searching capability of GA. El-Beltagy *et al.* [36] investigate the same problem and suggest that the issue of balancing the concerns of optimization with those of design of experiments should be addressed. Song

et al. [137] couple a real-coded GA with Kriging for firtree structural optimization. Zhou *et al.* [110, 161] present a hierarchical surrogate-assisted evolutionary optimization framework.

Another form of Kriging applications to optimization is SKO, also called the Efficient Global Optimization (EGO) method. Its basic concept is to approximate the objective function of an optimization problem with a Kriging model, and then use the Kriging model to obtain the most promising point for sequential sampling. SKO focuses on solving expensive noisy black-box problems. Some related references are [61, 63, 127, 131]. If the system of interest is very complicated, it is often necessary to draw data with less cost from surrogate experimental systems used to mimic production systems, called "lower-fidelity systems", or from computer simulations used to approximate physical experiments. The system of interest is called the "highest fidelity" system. Leary *et al.* [85] use the low fidelity data as prior knowledge to be incorporated in the training of neural network and the generation of Kriging model. Huang *et al.* [70, 60] propose an extension of SKO that utilizes multiple fidelity data to reduce total evaluation cost.

The literature review shows that Kriging applications in optimization concentrate on using its interpolation capabilities to reduce the computational cost of optimizing a complex system. The main feature of the application is that Kriging temporarily replaces an objective function to assist optimization search. Note that the integration application of Kriging and evolutionary algorithms to optimization is just starting. The clear structure of the integration algorithm is not given and there are no extensive analysis and discussion on the algorithm itself. In the literature, the evolutionary algorithm used to be integrated with Kriging focuses on GA. However, evolutionary algorithms have been improved. The

improved algorithms with stronger optimization capabilities should be considered. In addition, Kriging may be improved, and its interpolation capabilities can be enhanced in the integration algorithm. The integration algorithm is seldom applied to simulation optimization which is a key problem of simulation modeling. Exploring its application to simulation optimization with costly computational expenses is meaningful.

KRIGING METHODOLOGY

From the perspective of methodology, this chapter briefly introduces the basic principles of Kriging, and gives some advanced discussion. Subsequently, some key components of Kriging are analyzed which are drift function, covariance function, neighbor structure, and variance of interpolation errors. Finally, the properties of Kriging are discussed.

3.1 Kriging Methodology

3.1.1 Basic Principles

Consider a stochastic function with the form below:

$$Z(\mathbf{X}) = \mu(\mathbf{X}) + \varepsilon(\mathbf{X}) \tag{3.1}$$

where $\mathbf{X}$ is a position vector, $\mu(\mathbf{X})$ is a mean term or a drift function showing the average behavior of $Z(\mathbf{X})$, and $\varepsilon(\mathbf{X})$ is a random error term with $E[\varepsilon(\mathbf{X})] = 0$. In geology, $Z(\mathbf{X})$ may represent the gold deposit in a location $\mathbf{X}$, $\mu(\mathbf{X})$ correspondingly the expected amount of gold deposit, and $\varepsilon(\mathbf{X})$ the associated random error from the mean. Suppose N observed values are $Z(\mathbf{X}_1), Z(\mathbf{X}_2), \cdots, Z(\mathbf{X}_N)$. Kriging uses a linear combination of the observed values to estimate the function value at an unobserved point $\mathbf{X}$.

$$\hat{Z}(\mathbf{X}) = \sum_{i=1}^{N} \lambda_i Z(\mathbf{X}_i) \tag{3.2}$$

where the coefficients λ_i are selected in such a manner that the above estimator satisfies the following constraints:

1. $\hat{Z}(\mathbf{X})$ must be the unbiased estimator of $Z(\mathbf{X})$;

2. The variance of estimation errors must be minimized, i.e., $\hat{Z}(\mathbf{X})$ must be the BLUE.

According to the unbiased requirement, the equation below can be obtained:

$$E[\hat{Z}(\mathbf{X})] = \sum_{i=1}^{N} \lambda_i E[Z(\mathbf{X}_i)] \tag{3.3}$$

which can be transferred to the following form:

$$\mu(\mathbf{X}) = \sum_{i=1}^{N} \lambda_i \mu(\mathbf{X}_i) \tag{3.4}$$

Without loss of generality, suppose the drift function $\mu(\mathbf{X})$ consists of M basis functions $f_l(\mathbf{X})$, $l = 1, 2, \cdots, M$, that is,

$$\mu(\mathbf{X}) = \sum_{l=1}^{M} a_l f_l(\mathbf{X}) \tag{3.5}$$

where a_l $(l = 1, 2, \cdots, M)$ are unknown constant coefficients that are later determined in Eq. (3.22) and $f_l(\mathbf{X})$ is named a base function in Kriging literature. Then the unbiased requirement leads to:

$$f_l(\mathbf{X}) = \sum_{i=1}^{N} \lambda_i f_l(\mathbf{X}_i) \qquad l = 1, \cdots, M \tag{3.6}$$

18

Thus, minimizing the variance of estimation error becomes a constrained optimization problem outlined below:

$$\text{Minimize: } Var[\hat{Z}(\mathbf{X}) - Z(\mathbf{X})] = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j Cov[Z(\mathbf{X}_i), Z(\mathbf{X}_j)]$$

$$-2\sum_{i=1}^{N}\lambda_i Cov[Z(\mathbf{X}_i), Z(\mathbf{X})] + Var[Z(\mathbf{X})]$$

$$\text{Subject to: } f_l(\mathbf{X}) = \sum_{i=1}^{N}\lambda_i f_l(\mathbf{X}_i) \qquad l = 1, 2, \cdots, M \qquad (3.7)$$

Kriging models transform the problem into the following unconstrained optimization by introducing Lagrange multipliers u_l, $l = 1, 2, \cdots, M$.

$$\text{Minimize: } Var[\hat{Z}(\mathbf{X}) - Z(\mathbf{X})] = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j Cov[Z(\mathbf{X}_i), Z(\mathbf{X}_j)]$$

$$-2\sum_{i=1}^{N}\lambda_i Cov[Z(\mathbf{X}_i), Z(\mathbf{X})] + Var[Z(\mathbf{X})]$$

$$+2\sum_{l=1}^{M}u_l[\sum_{i=1}^{N}\lambda_i f_l(\mathbf{X}_i) - f_l(\mathbf{X})] \qquad (3.8)$$

Partially differentiating the right side of Eq. (3.8) with respect to λ_i ($i = 1, \cdots, N$) and u_l ($l = 1, \cdots, M$) generates the well-known "Universal Kriging system" as follows:

$$\sum_{i=1}^{N}\lambda_i f_l(\mathbf{X}_i) = f_l(\mathbf{X}) \qquad l = 1, \cdots, M \qquad (3.9)$$

$$\sum_{l=1}^{M}u_l f_l(\mathbf{X}_i) + \sum_{j=1}^{N}\lambda_j Cov[Z(\mathbf{X}_i), Z(\mathbf{X}_j)] = Cov[Z(\mathbf{X}_i), Z(\mathbf{X})] \qquad i = 1, \cdots, N \qquad (3.10)$$

which can be expressed in matrix form as:

$$
\begin{bmatrix} \mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{S} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{f(X)} \\ \mathbf{C(X)} \end{bmatrix}
\tag{3.11}
$$

where

$$
\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & \cdots & \lambda_N \end{bmatrix}^T
\tag{3.12}
$$

$$
\mathbf{U} = \begin{bmatrix} u_1 & \cdots & u_M \end{bmatrix}^T
\tag{3.13}
$$

$$
\mathbf{C(X)} = \begin{bmatrix} C_1 & \cdots & C_N \end{bmatrix}^T
\tag{3.14}
$$

$$
\mathbf{f(X)} = \begin{bmatrix} f_1(\mathbf{X}) & \cdots & f_M(\mathbf{X}) \end{bmatrix}^T
\tag{3.15}
$$

$$
\mathbf{S} = \begin{bmatrix} S_{11} & \cdots & S_{1N} \\ \vdots & \ddots & \vdots \\ S_{N1} & \cdots & S_{NN} \end{bmatrix}
\tag{3.16}
$$

$$\mathbf{F} = \begin{bmatrix} f_1(\mathbf{X}_1) & \cdots & f_M(\mathbf{X}_1) \\ \vdots & \ddots & \vdots \\ f_1(\mathbf{X}_N) & \cdots & f_M(\mathbf{X}_N) \end{bmatrix} \tag{3.17}$$

In the above matrices, S_{ij} represents $Cov[Z(\mathbf{X}_i), Z(\mathbf{X}_j)]$ which is the covariance between the sample points $Z(\mathbf{X}_i)$ and $Z(\mathbf{X}_j)$, and C_i denotes $Cov[Z(\mathbf{X}_i), Z(\mathbf{X})]$. The coefficient matrix with dimensions $(N+M) \times (N+M)$ in the left side of Eq. (3.11) is termed Kriging matrix. If Kriging coefficient matrix is nonsingular, coefficients λ_i $(i = 1, \cdots, N)$ can be obtained and thus Eq. (3.2) can provide the estimate of $Z(\mathbf{X})$.

3.1.2 Advanced Analysis

According to equations (3.2) and (3.11), some researchers such as Hertog *et al.* [53] provide the following Kriging estimator whose proof is provided by Appendix A:

$$\hat{Z}(\mathbf{X}) = \hat{\boldsymbol{\lambda}}^T \mathbf{Z} = \hat{\mathbf{a}}^T \mathbf{f}(\mathbf{X}) + \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}(\mathbf{Z} - \mathbf{F}\hat{\mathbf{a}}) \tag{3.18}$$

where

$$\hat{\boldsymbol{\lambda}} = \mathbf{S}^{-1}[\mathbf{C}(\mathbf{X}) - \mathbf{F}\hat{\mathbf{U}}] \tag{3.19}$$

$$\hat{\mathbf{U}} = (\mathbf{F}^T \mathbf{S}^{-1} \mathbf{F})^{-1}[\mathbf{F}^T \mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) - \mathbf{f}(\mathbf{X})] \tag{3.20}$$

$$\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_M \end{bmatrix}^T \tag{3.21}$$

$$\hat{\mathbf{a}} = (\mathbf{F}^T \mathbf{S}^{-1} \mathbf{F})^{-1}\mathbf{F}^T \mathbf{S}^{-1}\mathbf{Z} \tag{3.22}$$

$$\mathbf{Z} = \left[\begin{array}{ccc} Z(\mathbf{X}_1) & \cdots & Z(\mathbf{X}_N) \end{array}\right]^T \tag{3.23}$$

An extensive analysis on the Kriging estimator leads to the following results:

1. Eq. (3.18) shows that the key parts of the Kriging estimator are the chosen observation points $\mathbf{X}_1, \cdots, \mathbf{X}_N$, the covariance matrix $\mathbf{S}$ whose core is actually a covariance function of Kriging, and the drift function $\mathbf{a}^T \mathbf{f}(\mathbf{X})$. The prerequisite of performing Kriging estimation is to have a set of observation points. This prerequisite means that it is necessary to design a good experiment to obtain pertinent observation points. Essentially, obtaining observation points is the selection problem of neighboring structures. Kriging adopts the method of Generalized Least Squares (GLS) to estimate the coefficients of the drift function. GLS is a robust method and also can consider heteroscedasticity. The drift function coefficients can be estimated accurately.

2. The Kriging estimator consists of two parts. One is the estimator of the drift function $\hat{\mathbf{a}}^T \mathbf{f}(\mathbf{X})$ based on the GLS estimator $(\mathbf{F}^T \mathbf{S}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{S}^{-1} \mathbf{Z}$; the other is the weighted correction part of prediction errors of the estimator of the drift function, where $\mathbf{Z} - \mathbf{F}\hat{\mathbf{a}}$ is a prediction error vector. If $\mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}$ is treated as weighted coefficients of the prediction error vector, a larger correlation between a predicted point and an observed point results in a larger weighted coefficient.

3. If a predicted point is far away from an observed point, the correlation between the predicted point and the observed point is very weak. That is to say, $\mathbf{C}^T(\mathbf{X})$ will become very small, which leads to $\mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}(\mathbf{Z} - \mathbf{F}\hat{\mathbf{a}})$ close to zero. Clearly, Kriging is actually using the regression of GLS, that is, $\hat{\mathbf{a}}^T \mathbf{f}(\mathbf{X})$, to estimate the value of $Z(\mathbf{X})$. In this case, Kriging can only estimate the mean drift behavior of the predicted point.

22

Note that $\hat{\mathbf{a}}$ is the GLS estimator of the vector $[a_1 \cdots a_M]^T$ in Eq. (3.5). Because of similarities of Kriging and regression, next the chapter methodologically compares Kriging with regression to better illustrate the features and properties of Kriging. Additionally, Artificial Neural Network (ANN) is a classic artificial intelligence technique to perform interpolation. Therefore, the comparison covers ANN to show the comparison comprehensiveness.

Regression uses observation values $\mathbf{X}_1, \cdots, \mathbf{X}_N$ and $Z(\mathbf{X}_1), \cdots, Z(\mathbf{X}_N)$ to establish a regression model. The direct objective is to find an explicit relationship between $\mathbf{X}_1, \cdots, \mathbf{X}_N$ and $Z(\mathbf{X})$. Based on regression assumptions, the methods of Ordinary Least Squares (OLS) or GLS are used to estimate the coefficients of polynomials in a regression equation. OLS or GLS can minimize the variances of coefficient estimators of polynomials and make the estimators unbiased. Similar to regression, ANN uses observation values $\mathbf{X}_1, \cdots, \mathbf{X}_N$ and $Z(\mathbf{X}_1), \cdots, Z(\mathbf{X}_N)$ to fit a network. The objective of ANN is to directly construct a relationship between $\mathbf{X}_1, \cdots, \mathbf{X}_N$ and $Z(\mathbf{X})$ by minimizing Mean Square Errors (MSE) of actual and predicted values. However, ANN can capture a nonlinear relationship between independent and dependent variables but in an implicit way. As for Kriging, its direct objective is to establish a functional relationship between $Z(\mathbf{X}_1), \cdots, Z(\mathbf{X}_N)$ and $Z(\mathbf{X})$ by minimizing the variance of prediction errors of $Z(\mathbf{X})$ and by making the estimator of $Z(\mathbf{X})$ unbiased, which is different from regression and ANN. However, similar to regression, Kriging assumes a functional form of $E[Z(\mathbf{X})]$, and uses the GLS to estimate the coefficients of polynomials in the assumed function. The final Kriging estimator of $Z(\mathbf{X})$ is actually a relationship between $\mathbf{X}_1, \cdots, \mathbf{X}_N$ and $Z(\mathbf{X})$.

23

3.2 Drift Functions

3.2.1 Literature Review

According to the differences of drift functions, Kriging models are generally divided into three types, which are Simple Kriging (SK), Ordinary Kriging (OK) and Universal Kriging (UK). The drift part of SK is the zero constant. SK is the most basic form of Kriging [109]. The Kriging model most often used is OK, which was developed by Matheron [93] in the early sixties. The drift function of OK is an unknown nonzero constant. The general form of Kriging models is UK, which was introduced by Matheron in 1969 [96]. UK is a non-stationary geostatistical method and its drift function is modeled as a general linear function of coordinates. When a drift function is defined externally through some auxiliary variables, the corresponding Kriging model is named Kriging with External Drift (KED)[14, 153]. Ahmed and de Marsily [2] suggest that the drift and residuals can be fitted separately and then summed. Odeh *et al.* [107, 108] call this form of Kriging Regression Kriging (RK), and Goovaets [50] names it "kriging after detrending".

Some other drift functions can refer to Disjunctive Kriging (DK)[97], Indicator Kriging (IK)[17], and Lognormal Kriging (LK)[82]. DK is for nonlinear problems and it is a nonlinear generalization of Kriging [97]. IK is used for the estimation of discrete variables and it uses indicator functions to estimate transition probabilities [64]. Multiple Indicator Kriging is another version of Indicator Kriging. Multiple Indicator Kriging works with a family of indicator functions instead of one indicator function. LK is assumed for highly skewed distributions of $\varepsilon(\mathbf{X})$ and it can only interpolate positive data [17].

When Kriging interpolation is based on blocks instead of points in space, the corresponding Kriging model is called Block Kriging (BK). BK provides a mean estimate for a

discrete area around an interpolation point. When Matheron [93] formulates Kriging, he presents it in the form of BK. The possible reasons why he chooses this form are the need of mathematical generality and his immediate concern about the estimation of average ore content of mining blocks. The only difference of BK from point Kriging is that the estimated point is replaced by a block. BK is usually more appropriate to be applied to the problems related to environment. In these two problems the block is defined as the rectangular area around a point.

Dual Kriging [145] is an alternative formulation to Kriging. The value estimated by Dual Kriging is a linear combination of covariance functions. Dual Kriging is a powerful mathematical interpolation method, and is particularly useful when estimations are needed for a global area with a large number of locations or when the availability of an analytical expression is advantageous. With an analytical expression, Dual Kriging is very useful for sensitivity analysis. By concentrating on local neighbors, Aunon *et al.* [5], present an alternative implementation of dual ordinary Kriging to provide an analytical expression within each subsection instead of the global area. And they suggest a procedure to remove the potential discontinuity across subsection boundaries.

Factorial Kriging Analysis (FKA) is one more form of Kriging, and it is developed by Matheron [99]. The literature review shows that it has wide applications [49, 138]. The main advantage of FKA is that it can be used to extract components from a variable which may be mapped separately for analysis. Galli and Sandjivy [44] theoretically compare FKA and spectral analysis and demonstrate their formal equivalence. However, they indicate that spectral analysis is not directly applicable to irregularly spaced data, which makes FKA advantageous.

3.2.2 Mathematical Development

In the stochastic process (3.1), if drift function $\mu(\mathbf{X})$ is the zero constant, there does not exist the constraint (3.9). The matrix form of Kriging system (3.11) correspondingly becomes:

$$
\begin{bmatrix}
S_{11} & \cdots & S_{1N} \\
\vdots & \ddots & \vdots \\
S_{N1} & \cdots & S_{NN}
\end{bmatrix}
\begin{bmatrix}
\lambda_1 \\
\vdots \\
\lambda_N
\end{bmatrix}
=
\begin{bmatrix}
C_1 \\
\vdots \\
C_N
\end{bmatrix}
\tag{3.24}
$$

The Kriging model is called SK. If drift function $\mu(\mathbf{X})$ is an unknown non-zero constant, the constraint (3.9) becomes a simple equation as follows:

$$
\lambda_1 + \cdots + \lambda_N = 1
\tag{3.25}
$$

The matrix form of Kriging system (3.11) is correspondingly formulated as:

$$
\begin{bmatrix}
0 & 1 & \cdots & 1 \\
1 & S_{11} & \cdots & S_{NN} \\
\vdots & \vdots & \ddots & \vdots \\
1 & S_{N1} & \cdots & S_{NN}
\end{bmatrix}
\begin{bmatrix}
u_1 \\
\lambda_1 \\
\vdots \\
\lambda_N
\end{bmatrix}
=
\begin{bmatrix}
1 \\
C_1 \\
\vdots \\
C_N
\end{bmatrix}
\tag{3.26}
$$

This Kriging model is named OK. Equation (3.11) gives the general form of Kriging in matrix form, and it is named UK.

Furthermore, assume the inverse of Kriging coefficient matrix in Eq. (3.11) exists. Then, the Kriging system in Eq. (3.11) can be solved and their solutions λ_i can be expressed

26

as:

$$
\begin{bmatrix} \hat{\mathbf{U}} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{G} \\ \mathbf{Q} & \mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{f}(\mathbf{X}) \\ \mathbf{C}(\mathbf{X}) \end{bmatrix} \tag{3.27}
$$

where $\mathbf{E}$, $\mathbf{G}$, $\mathbf{Q}$, and $\mathbf{H}$ are the partitioned matrices of the inverse of Kriging coefficient matrix. Then the estimation value of $Z(\mathbf{X})$ can be carried out by using the following formula:

$$
\hat{Z}(\mathbf{X}) = \begin{bmatrix} Z(\mathbf{X}_1) & \cdots & Z(\mathbf{X}_N) \end{bmatrix} \mathbf{Q}\mathbf{f}(\mathbf{X}) + \begin{bmatrix} Z(\mathbf{X}_1) & \cdots & Z(\mathbf{X}_N) \end{bmatrix} \mathbf{H}\mathbf{C}(\mathbf{X}) \tag{3.28}
$$

Noting that $Z(\mathbf{X}_1), \ldots, Z(\mathbf{X}_N)$, $\mathbf{Q}$ and $\mathbf{H}$ are all known, in order to simplify the above expression, the dissertation defines a new set of coefficients below:

$$
\hat{\boldsymbol{\alpha}} = \begin{bmatrix} \hat{\alpha}_1 & \cdots & \hat{\alpha}_N \end{bmatrix} = \begin{bmatrix} Z(\mathbf{X}_1) & \cdots & Z(\mathbf{X}_N) \end{bmatrix} \mathbf{Q} \tag{3.29}
$$

$$
\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 & \cdots & \hat{\beta}_N \end{bmatrix} = \begin{bmatrix} Z(\mathbf{X}_1) & \cdots & Z(\mathbf{X}_N) \end{bmatrix} \mathbf{H} \tag{3.30}
$$

Thus the estimator of $Z(\mathbf{X})$ is simplified as:

$$
\hat{Z}(\mathbf{X}) = \hat{\boldsymbol{\alpha}}\mathbf{f}(\mathbf{X}) + \hat{\boldsymbol{\beta}}\mathbf{C}(\mathbf{X}) \tag{3.31}
$$

This Kriging model is called Dual Kriging.

3.3 Covariance Functions

3.3.1 Literature Review

In Kriging models, selecting a proper covariance function is a crucial problem. The research on covariance functions mainly focuses on investigating the influence of misspecified covariance function on estimates. Diamond and Armstrong [29], Sukhatme [142], Warnes [155], Armstrong and Myers [4] investigate the influence of small perturbations of covariance function on a Kriging model. Yakowitz and Szidarovszky [159], Stein [140], and Stein and Handock [141] study the behavior of Kriging models caused by an incorrect covariance function as the number of observations in some fixed region increases. Stein [140] shows that when the number of observed points increases, the impact caused by an incorrect covariance function is asymptotically negligible if the adopted covariance function is "compatible" with the actual covariance function in the region of interest. Journal and Huijbregts [65] indicate that the spherical covariance function is inappropriate for most three-dimensional fields.

3.3.2 Typical Covariance Functions

If the covariance function of a stochastic process only depends on the distance between two points and is not influenced by their particular fixed locations, the process is said to be second-order stationary. When a stochastic process has the same covariance structure in all directions, it is said to be homogeneous and isotropic. Homogeneity means that any two points have the same finite covariance if the distance between them is equal. Isotropy means that the covariance has the property of being independent of direction. For simplicity, Kriging models generally assume that covariance functions are homogeneous and isotropic. Of course, the assumptions are not the prerequisites of Kriging models. Calculation of a

covariance function needs to have the knowledge of $\mu(\mathbf{X}_1)$ and $\mu(\mathbf{X}_2)$. In order to avoid this, geologists develop another measurement of joint variation that can be calculated without the means. This measurement is called variogram. Variogram is the variance of difference between $Z(\mathbf{X}_1)$ and $Z(\mathbf{X}_2)$. Mathematically, it is defined as [14]:

$$2\gamma(r) = Var(Z(\mathbf{X}_1) - Z(\mathbf{X}_2)) \tag{3.32}$$

where $\gamma(r)$ is named semi-variogram. The relationship between covariance and semi-variogram is

$$Cov(Z(\mathbf{X}_1), Z(\mathbf{X}_2)) = Var(Z(\mathbf{X}_1)) - \gamma(r) \tag{3.33}$$

Some typical covariance functions and variogram frequently used under the homogeneous and isotropic assumptions are listed below.

1. Pure nugget effect covariance:

$$C(r) = \begin{cases} C_0 & , \text{ if } r = 0 \\ 0 & , \text{ otherwise} \end{cases} \tag{3.34}$$

2. Triangle model or tent (linear) covariance:

$$C(r) = \begin{cases} C_0(1 - \frac{r}{d}) & , \text{ if } r \le d \\ 0 & , \text{ otherwise} \end{cases} \tag{3.35}$$

29

3. Cubic covariance (1):

$$C(r) = \begin{cases} C_0(1 - 3\frac{r^2}{d^2} + 2\frac{r^3}{d^3}) & \text{, if } r \leq d \\ \\ 0 & \text{, otherwise} \end{cases} \tag{3.36}$$

4. "Cubic" covariance (2):

$$C(r) = \begin{cases} C_0(1 - 7\frac{r^2}{d^2} + 8.75\frac{r^3}{d^3} - 3.5\frac{r^5}{d^5} + 0.75\frac{r^7}{d^7}) & \text{, if } r \leq d \\ \\ 0 & \text{, otherwise} \end{cases} \tag{3.37}$$

where Chiles *et al.* [14] name the above polynomial as the "cubic" covariance function.

5. Spherical covariance:

$$C(r) = \begin{cases} C_0(1 - \frac{3r}{2d} + \frac{r^3}{2d^3}) & \text{, if } r \leq d \\ \\ 0 & \text{, otherwise} \end{cases} \tag{3.38}$$

6. Exponential covariance:

$$C(r) = C_0 e^{(-\frac{r}{d})} \tag{3.39}$$

7. Gaussian covariance:

$$C(r) = C_0 e^{-(\frac{r}{d})^2} \tag{3.40}$$

8. Power-function variogram:

$$\gamma(r) = pr^b, \, b \in (0, 2) \tag{3.41}$$

9. Logarithmic or De Wijsian variogram [82]:

$$\gamma(r) = p \log{(r)} \tag{3.42}$$

where d is the influence distance of a covariance function.

Generally, the triangle covariance function is used in a one-dimensional space. Cubic and spherical covariance functions are used in the space one to three dimensions. Exponential and Gaussian covariance functions could be used in any space.

3.4 Neighbor Structures

The application of Kriging models involves how to select some proper data points to perform interpolation. Theoretically, the more points included, the better, because any smaller neighbors can be regarded as a constrained optimization with weights zero placed on the discarded points. However, the problem is that when more data points are used, the corresponding Kriging coefficient matrix will become dimensionally larger, which will result in the difficulties of calculating the inverse of the Kriging coefficient matrix. In general, Kriging models work well for up to 100 points. Davis and Grivet [24] further point out that for some special purposes Kriging models can accommodate up to 400 points. However, it is possible by using faster computers to accommodate more data points.

Two important guidelines about selecting neighbors are screening and relay effects. Mathematically, the screening effect is to concentrate non-zero Kriging weights on a subset of samples in the immediate vicinity of an estimated point, and block off the influence of all other data. The relay effect shows the necessity to include more data points. The relay effect indicates that data points can influence each other by virtue of relay even if

the distance between them is large. A simple example is that suppose there are a series of random points Z_i, Z_{i+1}, $\cdots$, and they follow a normal distribution $N(0, \sigma^2)$. If there is a function relationship $Z_{i+1} = \omega_i Z_i$ between Z_i and Z_{i+1}. Then the covariance between Z_i and Z_{i+1} can be calculated as follows:

$$Cov(Z_i, Z_{i+1}) = Cov(Z_i, \omega_i Z_i) = \omega_i \sigma^2 \tag{3.43}$$

However, according to the relay effect, it can be further noted that there is a covariance between Z_i and Z_{i+2} which is computed as follows:

$$Cov(Z_i, Z_{i+2}) = Cov(Z_i, \omega_{i+1} Z_{i+1}) = Cov(Z_i, \omega_{i+1} \omega_i Z_i) = \omega_{i+1} \omega_i \sigma^2 \tag{3.44}$$

It is important to balance the need of screening and relay effects to choose neighbors. The need is embodied by choosing a proper influence distance in a covariance function. When a larger influence distance is chosen, more points are included and the relay effect can be satisfied. However, including more points will increase the computational complexity of Kriging. The screening effect shows that it is feasible to limit the influence distance of covariance to decrease data points and to control the computational complexity of Kriging.

3.5 Variance and Accuracy

Kriging can provide the variance of interpolation errors $\hat{\sigma}_e^2$. The variance can be used to measure the accuracy of Kriging estimator. For the stochastic process (3.1), combining

equations (3.8, 3.9, 3.10) results in the variance of interpolation errors as follows:

$$\hat{\sigma}_e^2 = Var[Z(\mathbf{X})] - \sum_{i=1}^{N} \hat{\lambda}_i Cov[Z(\mathbf{X}_i), Z(\mathbf{X})] - \sum_{i=1}^{N} \hat{\lambda}_i \sum_{l=1}^{M} \hat{u}_l f_l(\mathbf{X}_i) \qquad (3.45)$$

The references [14, 53] provide similar versions of Eq. (3.45). In reference [14], the covariance in Eq. (3.45) is a specific Gaussian covariance function. For SK, the base function is a zero constant, and $\hat{\sigma}_e^2$ reduces to

$$\hat{\sigma}_e^2 = Var[Z(\mathbf{X})] - \sum_{i=1}^{N} \hat{\lambda}_i Cov[Z(\mathbf{X}_i), Z(\mathbf{X})] \qquad (3.46)$$

For OK, the base function is an unknown non-zero constant, and $\hat{\sigma}_e^2$ becomes

$$\hat{\sigma}_e^2 = Var[Z(\mathbf{X})] - \sum_{i=1}^{N} \hat{\lambda}_i Cov[Z(\mathbf{X}_i), Z(\mathbf{X})] - \hat{u}_1 \qquad (3.47)$$

where $\hat{u}_1$ is the only Lagrange multiplier estimator in the matrix system (3.26) of OK. According to (3.8, 3.9, 3.10), $\hat{\sigma}_e^2$ is greater than zero.

3.6   Analysis on Properties

3.6.1   Algorithm Complexity

The algorithm complexity of Kriging is dominated by calculating the inverse of Kriging coefficient matrix. Therefore, the analysis on its algorithm complexity will focus on computing this inverse matrix. The typical algorithm in linear algebra to compute the inverse of a matrix is Gaussian elimination or Gauss-Jordan elimination, named after mathematician Karl Friedrich Gauss and Wilhelm Jordan.

The computational complexity of Gaussian elimination is $O(N^3)$, which means that when the matrix size is $N \times N$, the computation required in Gaussian elimination is proportional to N^3. However, this algorithm is numerically unstable, at least on pathological examples. For example, the floating-point errors committed throughout the computation are accumulated, which makes the actual results far from correct solutions. For very large matrix systems, Gauss-Jordan elimination is not a very good method because of its numerical instability and prohibitive computational cost, and iterative methods are more preferred. For some special matrix systems, there exist some better computational methods. Some classic references about how to compute an inverse matrix can be found in [66, 67].

When N increases, it will become more difficult for Gaussian elimination to compute an inverse matrix. However, it is unnecessary for Kriging to use a large number of observed points to estimate the values of unknown points. According to covariance functions, the covariances between an estimated point and the observed points far away from the estimated point are close to zero, and thus these observed points can be removed. In addition, for a cluster of observed points close to each other, the similarity among them is strong. Thus, only several representative points are needed, because according to screening effect it is obviously redundant to use all of the observed points in this cluster. After these considerations, the known points used to estimate the unknown points become very limited even if the actual number of points is large. Thus the computational complexity of Kriging can be well controlled. Additionally, a finite covariance function often causes Kriging coefficient matrix to become sparse, which makes possible the use of sparse matrix solvers. A sparse matrix is a matrix populated primarily with zeros.

3.6.2 Parameter Settings

One of the main issues of applying Kriging is to choose its parameters. A significant advantage of Kriging is that its parameter settings are limited. For SK and OK, only two parameters which are covariance functions and influence distances of covariance need be set. In some practical problems, the selection of covariance functions and influence distances of covariance is very obvious or defined by the problems. In this case, no parameters need be set. Although parameter settings in Kriging are few, they are important and determine the accuracy and efficiency of Kriging interpolation.

3.7 Conclusions

This chapter introduced the basic principles of Kriging, and gave an advanced discussion. The discussion analyzed some important structure properties of Kriging estimator, and pointed out the similarities and differences of Kriging, regression, and ANN. Subsequently, the chapter investigated some key structure components of Kriging. Finally, some important properties of Kriging were discussed.

Chapter 4

Theoretical Developments Of Kriging For Simulation Interpolation

4.1   Introduction

The main goal of the chapter is to discuss three important theoretical developments of
Kriging which include a novel Kriging model based on Taylor expansion, the measurement
of influence distance of covariance, and the sensitivity analysis method based on Kriging
and Analysis of Variance (ANOVA). These developments can benefit Kriging interpolation
and sensitivity analysis for simulation models.

For Simple Kriging (SK) and Ordinary Kriging (OK), the drift functions of models are
zero and an unknown nonzero constant, respectively, which limit the simulation interpola-
tion capabilities of these two Kriging models. Mathematically, Taylor expansion has strong
nonlinear function approximation abilities. If Taylor expansion can be integrated with Krig-
ing to describe the mean drift behavior of data, the simulation interpolation abilities of the
integrated Kriging may be stronger. With this idea in mind, this chapter develops a novel
Kriging model based on Taylor expansion.

The influence distance in a covariance function impacts the simulation interpolation
performance of Kriging. Thus, it is essential to choose a proper value for it. However, this is
difficult because different simulation problems have different data magnitudes. This chapter
suggests sample standard deviation to be the metric for influence distance of covariance.
The objective is to simplify the setting of this parameter.

Sensitivity analysis can offer insights into the variation in simulation outputs caused
by inputs. The outcome of sensitive analysis is to quantify the influence of input factors

on outputs and to find critical input factors, which can help decision makers improve a simulation model. Kriging models are explicit and differentiable. Thus, it is feasible to apply partial differentiation to perform sensitivity analysis on independent variables of Kriging. When Kriging is used to construct a metamodel for inputs and outputs of a complicated simulation model, the sensitivity analysis on Kriging can assist model developers to analyze the output sensitivity to inputs. The partial differentiation equation of Kriging has not been seen in references. This chapter devises its mathematical form. Together with ANOVA, it is used to build a sensitivity analysis method for simulation models.

The organization of the chapter is as follows. First, a novel Kriging model based on Taylor expansion is developed, and its key properties are discussed. Second, the use of sample standard deviation to define influence distance of covariance is described. Then, the sensitivity analysis based on Kriging and ANOVA is presented. Finally, some conclusions are given.

4.2 A Novel Kriging Model Based on Taylor Expansion

4.2.1 Model Development

Suppose $\mu(\mathbf{X})$ has continuous derivatives up to the $(m+1)^{th}$ order at point $\mathbf{X}_0$. Then the Taylor expansion of $\mu(\mathbf{X})$ at the point $\mathbf{X}_0$ is:

$$\mu(\mathbf{X}) = \mu(\mathbf{X}_0) + \mu^{'}(\mathbf{X}_0)(\mathbf{X} - \mathbf{X}_0) + \frac{\mu^{(2)}(\mathbf{X}_0)}{2!}(\mathbf{X} - \mathbf{X}_0)^2 + \cdots$$
$$+ \frac{\mu^{(m)}(\mathbf{X}_0)}{m!}(\mathbf{X} - \mathbf{X}_0)^m + \frac{\mu^{(m+1)}(\boldsymbol{\xi})}{(m+1)!}(\mathbf{X} - \mathbf{X}_0)^{m+1} \qquad (4.1)$$

where $\boldsymbol{\xi}$ is a vector in the Taylor expansion such that each of its components lies within the corresponding components of $\mathbf{X}$ and $\mathbf{X}_0$. Obviously, the base functions in Eq. (4.1) are

$$f_l(\mathbf{X}) = (\mathbf{X} - \mathbf{X}_0)^l, \qquad l = 0, \cdots, m+1 \tag{4.2}$$

The Kriging matrix system therefore becomes:

$$\begin{bmatrix} \mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{S} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{pmatrix} 1 \\ \vdots \\ (\mathbf{X} - \mathbf{X}_0)^{m+1} \\ \mathbf{C}(\mathbf{X}) \end{pmatrix} \tag{4.3}$$

where

$$\mathbf{F} = \begin{bmatrix} 1 & \cdots & (\mathbf{X}_1 - \mathbf{X}_0)^m & (\mathbf{X}_1 - \mathbf{X}_0)^{m+1} \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & (\mathbf{X}_N - \mathbf{X}_0)^m & (\mathbf{X}_N - \mathbf{X}_0)^{m+1} \end{bmatrix} \tag{4.4}$$

Since this Kriging system is developed according to Taylor expansion, the Kriging model is termed Taylor Kriging (TK).

4.2.2 Model Properties

Next the properties of TK are investigated. Compared with SK and OK, TK has one more important parameter, m, to be determined. A Taylor expansion with high order (large m) has better approximation capabilities. However, the computation complexity of TK will quickly increase as m increases.

As discussed in Chapter 3, the computation complexity of Kriging is dominated by calculating the inverse of Kriging coefficient matrix $\mathbf{K} = \begin{bmatrix} \mathbf{0} & \mathbf{F}^T \\ \mathbf{F} & \mathbf{S} \end{bmatrix}$. If Gaussian elimination is used to invert $\mathbf{K}$, the computational complexity of Kriging is $O(N^3)$, where N denotes the matrix size. In TK, the size of $\mathbf{K}$ is determined by the number of known points used to estimate an unknown point plus the order of Taylor expansion and the dimension of independent variables. Suppose there are N known points, the order of Taylor expansion is m, and the dimension of independent variables is D; then, the computational complexity of TK is $O((N + f(m, D))^3)$. When the order in Taylor expansion is 1, the computational complexity of the corresponding TK is $O((N + (1 + D))^3)$. When the order in Taylor expansion is 2 or 3, the computational complexity of TK becomes $O((N + (1 + D + D(D + 1)/2))^3)$ or $O((N + (1 + D + D(D + 1)/2 + D^2 + D(D - 1)(D - 2)/(3 \times 2 \times 1))^3)$, respectively. Clearly, as the order of Taylor expansion m increases, the computation complexity of TK will increase rapidly. For example, when N is equal to 5, the order of Taylor expansion is 2, and the dimension of independent variables is 3, the size of $\mathbf{K}$ becomes 15 by 15, which is much larger than those of SK (5 by 5) and OK (6 by 6).

However, like regression, a too high order could cause the overfitting problem. Considering computational complexity and overfitting, generally the value of m is a positive integer less than 4. Another important parameter in the complexity formula is the dimension number D of variable $\mathbf{X}$. If D substantially increases, Taylor expansion will be complicated and the computational expenses of TK will increase rapidly.

4.3 The Measurement of Influence Distance of Covariance

The influence distance of covariance represents the parameter d in Eq. (3.34, 3.35, 3.36, 3.37, 3.38, 3.39, 3.40). Once a covariance function is chosen, the parameter d need be set. Note that a covariance function only has a parameter, that is d, which makes its selection crucial. Since different simulation problems have differing data magnitudes, the choosing process becomes very difficult. Here sample standard deviation denoted by S is adopted and used as the measurement unit of influence distance of covariance. S is calculated as follows:

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} [Z(\mathbf{X}_i) - \overline{Z}]^2} \tag{4.5}$$

where $\overline{Z} = \frac{1}{N} \sum_{i=1}^{N} Z(\mathbf{X}_i)$. The advantages of using sample standard deviation are as follows:

1. Sample standard deviation is easily understood and calculated;

2. Using sample standard deviation as the measurement unit of influence distance of covariance can make different problems have a consistent measurement.

In addition, sample standard deviation is a relative measurement unit instead of an absolute or fixed value. For example, if $S = 0.2$, two units of the sample standard deviation are equal to 0.4; however, if $S = 0.4$, two units of the sample standard deviation become 0.8.

4.4 The Sensitivity Analysis Method Based on Kriging and ANOVA

Sensitivity analysis on a simulation model implies how simulation inputs influence outputs. If a slight change in the value of an input variable has a large impact on its

corresponding output, the simulation model is said to be very sensitive to the input variable. This section uses the explicit and differentiable properties of Kriging to develop its partial differentiation equation which is used to perform sensitivity analysis on independent variables. Suppose that Kriging is utilized to fit a metamodel for inputs and outputs of a complex simulation model. The sensitivity analysis on dependent variables of the fitted Kriging metamodel can be used to analyze the output sensitivity of the simulation model to inputs.

The partial derivative of Eq. (3.31) with respect to variable x_j is calculated as follows:

$$\frac{\partial}{\partial x_j}\hat{Z}(\mathbf{X}) = \hat{\boldsymbol{\alpha}} \begin{bmatrix} \frac{\partial}{\partial x_j}f_1(\mathbf{X}) \\ \vdots \\ \frac{\partial}{\partial x_j}f_M(\mathbf{X}) \end{bmatrix} + \hat{\boldsymbol{\beta}} \begin{bmatrix} \frac{\partial}{\partial x_j}C_1 \\ \vdots \\ \frac{\partial}{\partial x_j}C_N \end{bmatrix} \tag{4.6}$$

If the covariance function below is adopted to compute the covariance between two points in a space,

$$C_i = C(r_i) \qquad i = 1, \cdots, N \tag{4.7}$$

where

$$r_i = \sqrt{\sum_{j=1}^{N}(x_j - x_{ij})^2} \tag{4.8}$$

the partial derivative of C_i with respect to x_j is given by:

$$\frac{\partial C_i}{\partial x_j} = \frac{\partial C(r_i)}{\partial r_i}\frac{\partial r_i}{\partial x_j} = \frac{\partial C(r_i)}{\partial r_i}\frac{(x_j - x_{ij})}{\sqrt{\sum_{j=1}^{N}(x_j - x_{ij})^2}} \tag{4.9}$$

41

Once the specific covariance and base functions are chosen, $\frac{\partial}{\partial x_j}\hat{Z}(\mathbf{X})$ will become explicit. For example, assume that the chosen covariance function is a cubic covariance (2) function. Then the partial derivative of the covariance function with respect to x_j becomes

$$\frac{\partial C_i}{\partial x_j} = \frac{(x_j - x_{ij})}{\sqrt{\sum_{j=1}^{N}(x_j - x_{ij})^2}} C_0 (1 - 14\frac{r_i}{d^2} + 26.25\frac{r_i^2}{d^3} - 17.5\frac{r_i^4}{d^5} + 5.25\frac{r_i^6}{d^7}) \qquad (4.10)$$

For SK and OK, the base functions are constants and thus their partial derivatives with respect to x_j become zero. For Taylor Kriging, although the base functions are not constants, their partial derivatives with respect to x_j can be calculated easily.

When a fitted Kriging metamodel is given, its partial derivative can be used to assist in sensitivity analysis of a simulation model. However, it is necessary to first determine which fitted Kriging metamodel should be chosen. Here analysis of variance (ANOVA) is used to select a fitted Kriging metamodel. ANOVA developed by Sir R. A. Fisher is a powerful tool to analyze experimental data. It is a statistical procedure to allocate the amount of variation in a process and to determine if a factor's impact on outputs is significant or caused by random noise [129]. Average Absolute Relative Error (AARE) between actual values and Kriging outputs is a chosen evaluation standard. ANOVA tests significant differences of AAREs and then determines which Kriging metamodel is more effective and should be used for the following simulation interpolation and sensitivity analysis. If ANOVA indicates multiple significant metamodels, theoretically any of them can be adopted. In the dissertation the Kriging metamodel with a smallest AARE and/or a simplest function form among these models will be chosen and used subsequently.

4.5   Conclusions

Chapter 4 used Taylor expansion to approximate a drift function, thus developing a novel Kriging model. Taylor expansion has strong functional approximating capabilities and can help Kriging better capture the mean drift behavior of data, thereby enhancing the simulation interpolation potentials of Kriging. Influence distance of covariance is an important parameter of a Kriging model. Sample standard deviation is suggested to be its metric. One significant advantage of using sample standard deviation is that different simulation problems have a consistent measurement unit for influence distance of covariance. Sensitivity analysis can investigate the variation in simulation outputs caused by inputs. The chapter developed a partial differentiation equation of Kriging and used it to analyze the sensitivity of a dependent variable to independent variables in a Kriging metamodel, thus indirectly realizing the sensitivity analysis of simulation outputs on inputs in its corresponding simulation model.

CHAPTER 5

KRIGING METAMODELING FOR DETERMINISTIC AND STOCHASTIC SIMULATION

INTERPOLATION

5.1 Introduction

Interpolation is a key problem of simulation modeling. The objective of interpolation is to significantly reduce the expense of running simulation models. Chapter 4 theoretically establishes a Kriging model based on Taylor expansion named Taylor Kriging (TK) to enhance the potentials of Kriging simulation interpolation. This chapter will conduct the empirical research to show its superiority. Some simulated test cases under deterministic and stochastic environments will be used to compare and analyze the interpolation accuracy and efficiency of Simple Kriging (SK) where the drift function is zero, Ordinary Kriging (OK) where the drift function is an unknown nonzero constant, and TK. These test cases are based on some multimodal benchmark functions.

The chapter is organized as follows. First, the generation of initial simulation inputs based on Latin Hypercube Sampling (LHS) will be introduced. Second, the interpolation accuracy and effectiveness of SK, OK and TK under deterministic environments will be compared and analyzed by using four simulated cases. Next, the stochastic simulation interpolation of SK, OK and TK will be conducted, and some problems related to stochastic simulation interpolation will be investigated. Finally, some conclusions will be provided.

5.2   Generation of Simulation Data

Latin Hypercube Sampling (LHS) is chosen to generate initial simulation input values. LHS is proposed by McKay, Beckman and Conover [90]. They point out that LHS has a smaller variance than Simple Random Sampling (SRS). Stein [139] and Owen [113] give a specific expression for the variance. Sacks, Welch, Mitchell and Wynn [124] indicate that LHS is an extension of stratified sampling which ensures that each of input variables has all portions of its range represented. The main advantages of LHS are that it is computationally less expensive to generate, and can deal with a large number of runs and input variables. The procedure of LHS is as follows [39]:

1. Let $\pi_j(1), \cdots, \pi_j(N)$ be a random permutation of the integers $1, \cdots, N$. Generate s permutations ($j = 1, \cdots, s$), thus obtaining a Latin Hypercube Design (LHD) expressed as LHD(N, s);

2. Next generate a matrix of $N \times s$ uniform variates (also called random numbers in Monte Carlo methods) $U_k^j \sim U(0, 1), k = 1, \cdots, N; j = 1, \cdots, s$, which are mutually independent. Let $\mathbf{X}_k = (x_{k1}, \cdots, x_{ks})$, where,

$$x_{kj} = \frac{\pi_j(k) - U_k^j}{N} \qquad k = 1, \cdots, N; j = 1, \cdots, s. \tag{5.1}$$

where $0 \leq x_{kj} \leq 1$. Then $D_N = \{\mathbf{X}_1, \cdots, \mathbf{X}_N\}$ is an LHS denoted by $LHS(N, s)$.

Scatter figures 5.1 and 5.2 are two sampling examples, each including 25 observations, from SRS and LHS, respectively. For Figure 5.2, corresponding to the algorithm above, N is equal to 25, s equal to 2 which means that each point has two dimensions, and the algorithm generates an LHS(25, 2) design. In Figure 5.2, each row and each column only

have a randomly representative point. However, in Figure 5.1, it is possible that there are no points or more than one point in a row or column. Clearly, LHS avoids redundant sampling and ensures that there is sufficient observations.

5.3 Kriging Interpolation for Deterministic Simulation

5.3.1 Test Functions and Procedures

Four deterministic simulation cases based on multimodal benchmark functions are used to compare the simulation interpolation capabilities of SK, OK and TK. These functions are as follows.

1. The Himmelblau Function (Hmb). The search domain of function is $-6 < x_i < 6, i = 1, 2$. The function has several local minima and one global minimum [71].

$$Hmb = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 + 0.1((x_1 - 3)^2 + (x_2 - 2)^2) \qquad (5.2)$$

2. The Six Hump Camelback Function (SHC). The search domain of function is $-3 < x_1 < 3$ and $-2 < x_2 < 2$. The function has six local minima and two global minima [8].

$$SHC = (4 - 2.1x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1 x_2 + (-4 + 4x_2^2)x_2^2 \qquad (5.3)$$

3. The B2 Function. The search domain of function is $-100 < x_i < 100, i = 1, 2$. The function has several local minima and one global minimum [38].

$$B2 = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1) - 0.4\cos(3\pi x_2) + 0.7 \qquad (5.4)$$
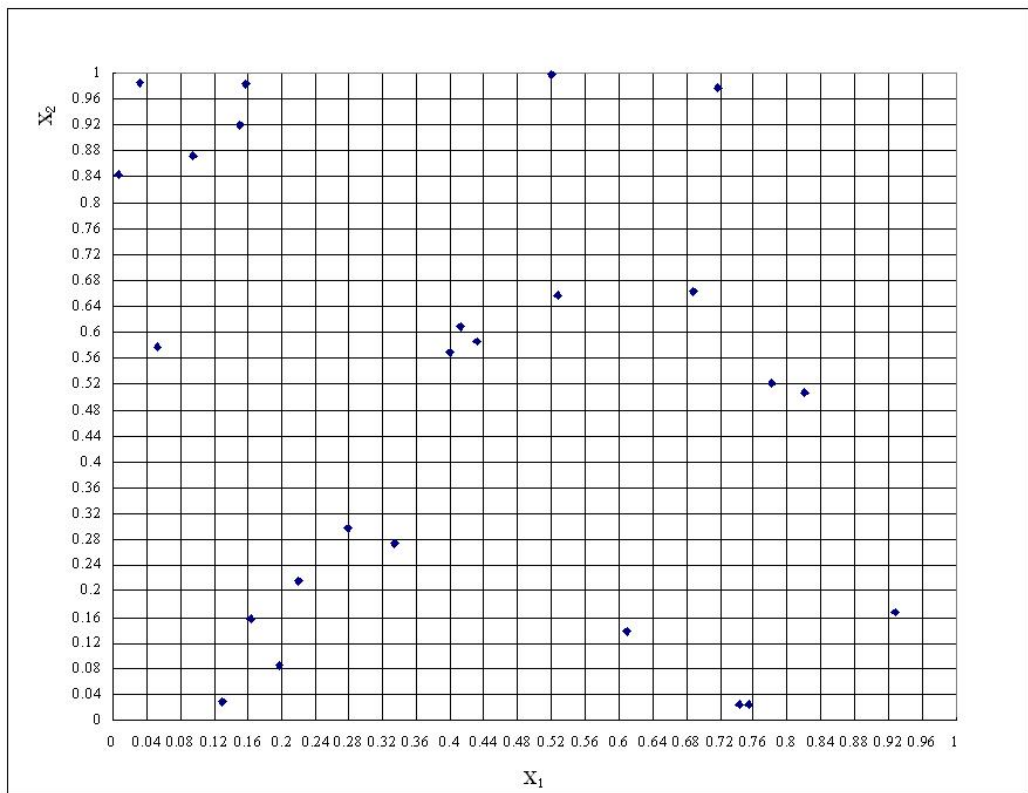
Figure 5.1: Scatter Figure of 25 Samples from SRS
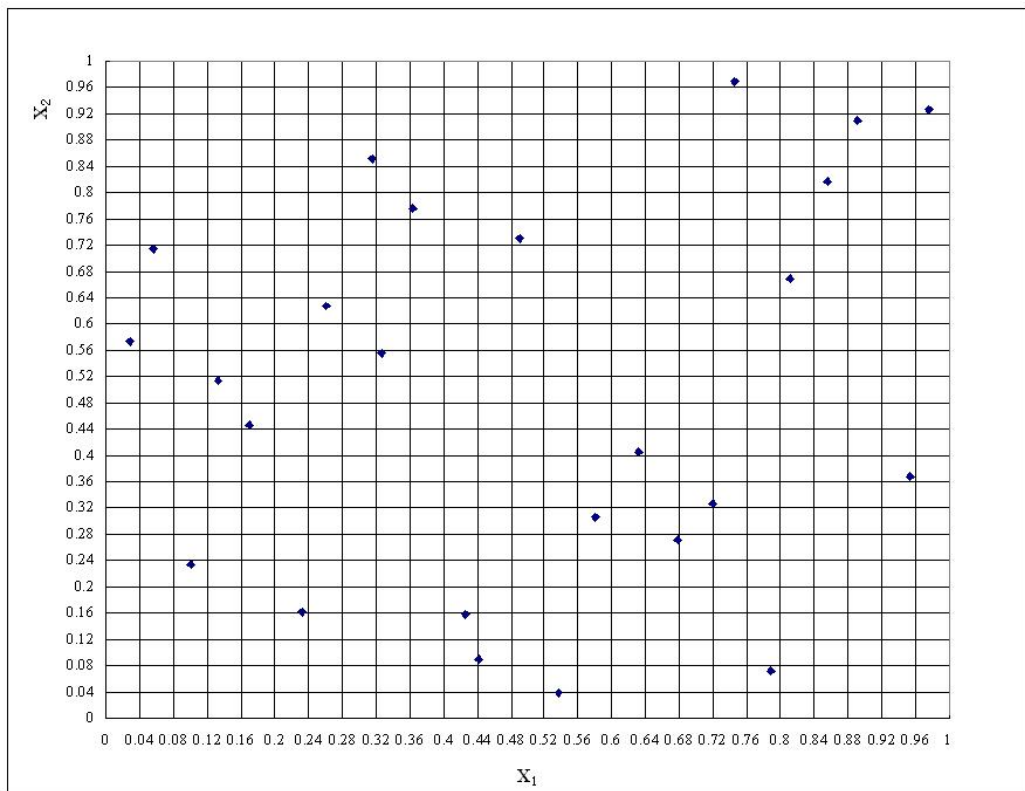
Figure 5.2: Scatter Figure of 25 Samples from LHD

4. The Rastrigin Function (Ras). The search domain of function is $-5.12 \leq x_i \leq$ 5.12, $i = 1, 2, \cdots, D$. D is the dimension number. The function has several local minima and one global minimum [38].

$$Ras = 10D + \sum_{i=1}^{D}(x_i^2 - 10\cos(2\pi x_i)) \qquad (5.5)$$

In the Ras function, D is arbitrarily set at 10.

The test process is divided into four steps. The first step is to use LHS to generate 150 simulation inputs in the feasible region of independent variables, and to obtain their corresponding simulation outputs (function values). Using LHS to generate 150 simulation inputs means that in the LHS algorithm N is set to 150, which is similar to the example of N equal to 25 in Section 5.2. The second step is to use the 150 inputs and outputs to fit Kriging models with different parameter settings. Next, 50 inputs generated uniformly and randomly in the feasible region are used to test the fitted Kriging models. Finally, the average absolute relative error (AARE) percentages of 50 predicted outputs are used as the standard of performance measurement to evaluate the performance of Kriging models. Here, Absolute Relative Error (ARE) is calculated as follows:

$$ARE = |\frac{y_p - y_a}{y_a}| \times 100\% \qquad (5.6)$$

where y_a and y_p are the actual and predicted simulation outputs, respectively, and $|y_p - y_a|$ expresses the absolute error.

5.3.2 Result Comparison

Scatter figures 5.3 and 5.4 graphically describe the simulation inputs of the first test function. The simulation inputs in Figure 5.3 are generated by using LHS in the feasible regions of the Hmb function. The simulation inputs in Figure 5.4 are uniformly and stochastically generated. The AARE percentages of 50 interpolated outputs from different Kriging models for the four test functions are shown in Tables 5.1, 5.2, 5.3 and 5.4, respectively. In these tables, SD represents standard deviation. The column of SD gives the values of the influence distance of covariance which could be 2SD, or 5SD, etc.. According to these tables, the best fitted models for different types of Kriging are listed in Table 5.5.

The Kriging models with smallest AARE percentages from SK, OK and TK, respectively, are compared in Table 5.6 which indicates the averages, minima and maxima of ARE percentages of 50 tested simulation outputs. These values consistently indicate that for the four functions, TK performs best. Since drift functions of SK, OK, and TK are zero, a nonzero constant, and a polynomial whose order is usually less than 5, respectively, TK has the strongest approximation potentials. As expected, TK has the best performance. For example, the Taylor expansion with order 3 serving as a drift function is very close to the Hmb function. The Kriging model with the Taylor expansion exhibits the best performance.

In Table 5.1, the AAREs of the OK models with cubic (2) covariance function and the TK models of order 3 with Cubic (2) covariance function are close. It is difficult to tell whether the outputs from these two types of Kriging models are significantly different. Here the paired t-test is used to test the AARE differences, and Table 5.7 gives the results. In Table 5.7, Variables 1 and 2 represent the OK models with cubic (2) covariance function
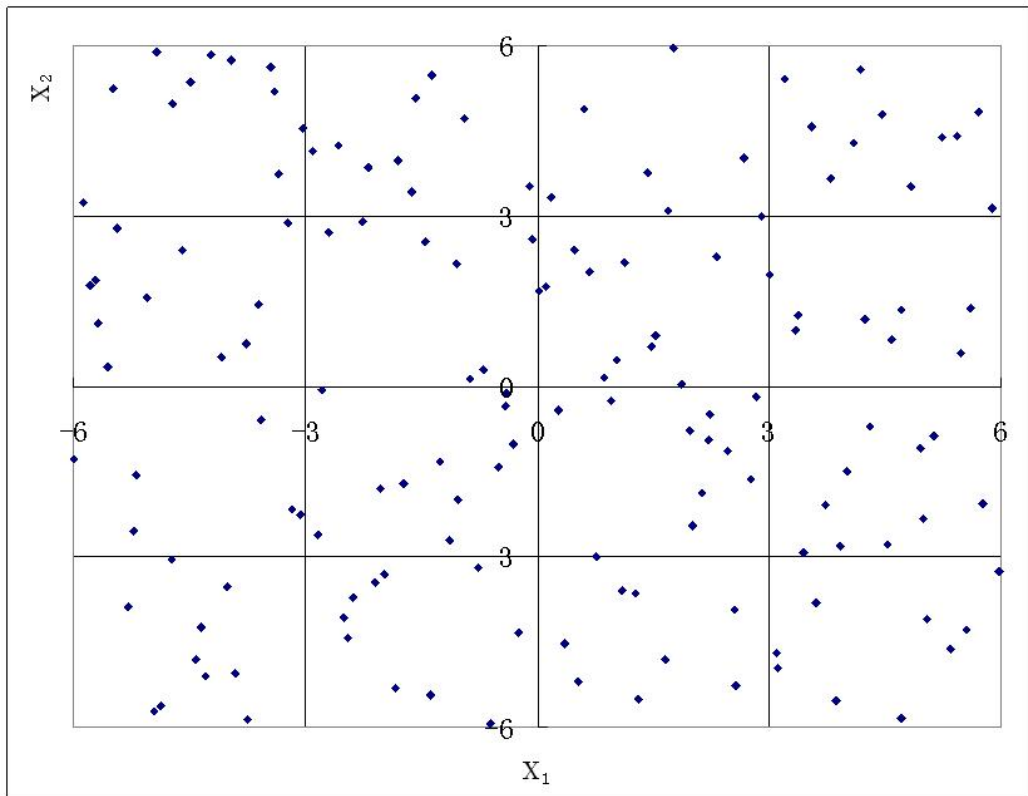
Figure 5.3: Scatter Figure of 150 Samples of First Test Function (Hmb)
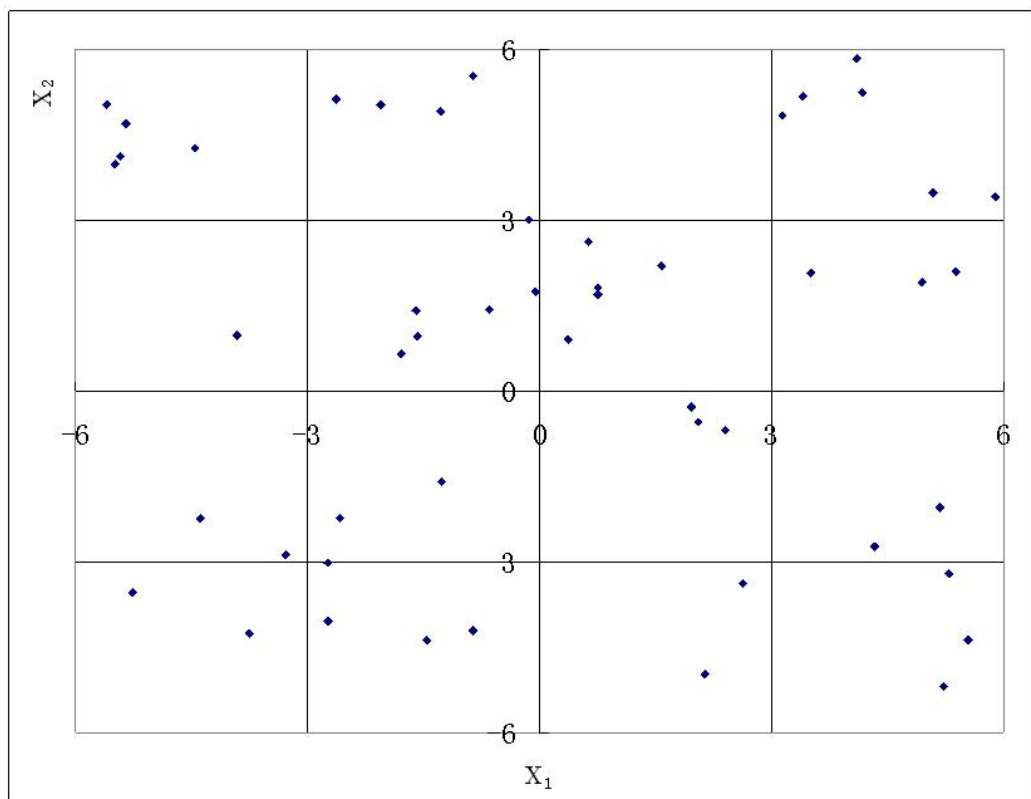
Figure 5.4: Scatter Figure of 50 Test Samples of First Test Function (Hmb)

Table 5.1: AARE of First Test Function (Hmb) (%)

Covariance Function	SD	SK	OK	Order of Taylor Kriging			
				1	2	3	4
Nugget	0	100	205.3478	205.3478	205.0371	115.8544	106.6875
Linear	1	10.7457	10.7088	10.7088	10.6995	8.3727	8.4311
	2	10.7269	10.7088	10.7088	10.6995	8.3727	8.4311
	3	10.7208	10.7088	10.7088	10.6995	8.3727	8.4311
	4	10.7177	10.7088	10.7088	10.6995	8.3727	8.4311
	5	10.7159	10.7088	10.7088	10.6995	8.3727	8.4311
Cubic (2)	1	1.3279	1.3046	1.3046	1.3048	1.1597	1.1636
	2	1.3161	1.3041	1.3041	1.3042	1.1597	1.1641
	3	1.3115	1.3031	1.3031	1.3049	1.1602	1.1636
	4	1.3086	1.3041	1.3041	1.3023	1.1583	1.1630
	5	1.3052	1.3008	1.3008	1.3010	1.1595	1.1588
Spherical	1	10.7633	10.7071	10.7071	10.6978	8.3721	8.4306
	2	10.7357	10.7084	10.7084	10.6991	8.3725	8.4310
	3	10.7267	10.7086	10.7086	10.6993	8.3726	8.4311
	4	10.7222	10.7087	10.7087	10.6994	8.3726	8.4311
	5	10.7195	10.7087	10.7087	10.6994	8.3726	8.4311
Exponential	1	10.7446	10.7099	10.7099	10.7003	8.3729	8.4314
	2	10.7267	10.7091	10.7091	10.6997	8.3727	8.4312
	3	10.7207	10.7090	10.7090	10.6996	8.3727	8.4312
	4	10.7177	10.7089	10.7089	10.6995	8.3727	8.4312
	5	10.7159	10.7089	10.7089	10.6995	8.3727	8.4311
Minimum=				1.1583			

Table 5.2: AARE of Second Test Function (SHC) (%)

Covariance Function	SD	SK	OK	Order of Taylor Kriging			
				1	2	3	4
Nugget	0	100	78.9620	78.9620	80.4530	69.6518	764.9876
Linear	1	29.3027	51.1483	51.1483	51.2805	41.9032	35.1451
	2	29.3737	51.1483	51.1483	51.2805	41.9032	35.1451
	3	29.3946	51.1483	51.1483	51.2805	41.9032	35.1451
	4	29.4047	51.1483	51.1483	51.2805	41.9032	35.1451
	5	29.4105	51.1483	51.1483	51.2805	41.9032	35.1451
Cubic (2)	1	3.7112	4.0921	4.0921	4.0936	4.0055	3.4727
	2	3.6456	4.1018	4.1018	4.1022	4.0112	3.4778
	3	3.6213	4.1035	4.1035	4.1038	4.0122	3.4786
	4	3.6087	4.1040	4.1040	4.1045	4.0124	3.4791
	5	3.6015	4.1045	4.1045	4.1042	4.0132	3.4790
Spherical	1	29.3269	51.0731	51.0731	51.2141	42.0161	35.0955
	2	29.3637	51.1230	51.1230	51.2571	41.9311	35.1327
	3	29.3838	51.1365	51.1365	51.2696	41.9155	35.1396
	4	29.3951	51.1415	51.1415	51.2743	41.9101	35.1420
	5	29.4022	51.1440	51.1440	51.2765	41.9076	35.1431
Exponential	1	29.3259	51.1874	51.1874	51.3364	41.8479	35.1699
	2	29.3796	51.1526	51.1526	51.2935	41.8892	35.1513
	3	29.3974	51.1482	51.1482	51.2863	41.8970	35.1478
	4	29.4063	51.1470	51.1470	51.2837	41.8997	35.1467
	5	29.4116	51.1468	51.1468	51.2826	41.9009	35.1461
Minimum=				3.4727			

Table 5.3: AARE of Third Test Function (B2) (%)

Covariance Function	SD	SK	OK	Order of Taylor Kriging			
				1	2	3	4
Nugget	0	100	139.1253	139.1253	140.0583	0.0049	0.0050
Linear	1	1.0818	1.0698	1.0698	1.0753	0.0051	0.0051
	2	1.0757	1.0698	1.0698	1.0753	0.0051	0.0051
	3	1.0737	1.0698	1.0698	1.0753	0.0051	0.0051
	4	1.0728	1.0698	1.0698	1.0753	0.0051	0.0051
	5	1.0722	1.0698	1.0698	1.0753	0.0051	0.0051
Cubic (2)	1	0.0248	0.0236	0.0236	0.0236	0.0073	0.0073
	2	0.0243	0.0236	0.0236	0.0238	0.0075	0.0079
	3	0.0250	0.0251	0.0251	0.0241	0.0097	0.0112
	4	0.0334	0.0357	0.0357	0.0347	0.0140	0.0135
	5	0.0457	0.0425	0.0425	0.0406	0.0269	0.0244
Spherical	1	1.0876	1.0695	1.0695	1.0749	0.0051	0.0051
	2	1.0786	1.0697	1.0697	1.0752	0.0051	0.0051
	3	1.0757	1.0698	1.0698	1.0752	0.0051	0.0051
	4	1.0742	1.0698	1.0698	1.0752	0.0051	0.0051
	5	1.0733	1.0698	1.0698	1.0752	0.0051	0.0051
Exponential	1	1.0814	1.0699	1.0699	1.0754	0.0051	0.0051
	2	1.0756	1.0698	1.0698	1.0753	0.0051	0.0051
	3	1.0737	1.0698	1.0698	1.0753	0.0051	0.0051
	4	1.0727	1.0698	1.0698	1.0753	0.0051	0.0051
	5	1.0721	1.0698	1.0698	1.0753	0.0051	0.0051
Minimum=				0.0049			

Table 5.4: AARE of Fourth Test Function (Ras) (%)

Covariance Function	SD	SK	OK	Order of Taylor Kriging			
				1	2	3	4
Nugget	0	100	15.77	15.77	16.113	14.334	1012.61
Linear	1	13.801	12.377	12.377	12.468	14.733	990.789
	2	12.614	12.377	12.377	12.468	14.733	1182.203
	3	12.468	12.377	12.377	12.468	14.733	489.747
	4	12.429	12.377	12.377	12.468	14.733	3172.592
	5	12.408	12.377	12.377	12.468	14.733	4221.687
Cubic (2)	1	16.288	15.257	15.257	14.987	16.287	1459.488
	2	15.096	15.227	15.227	15.122	16.135	11049.384
	3	15.139	15.195	15.195	15.132	16.091	1042.38
	4	15.146	15.179	15.179	15.135	16.074	9381.389
	5	15.147	15.169	15.169	15.136	16.066	65029.349
Spherical	1	15.156	12.423	12.423	12.518	14.744	859.771
	2	12.834	12.39	12.39	12.479	14.736	2257.75
	3	12.59	12.383	12.383	12.473	14.734	8957.383
	4	12.491	12.381	12.381	12.471	14.734	818.743
	5	12.449	12.379	12.379	12.47	14.734	6815.851
Exponential	1	12.705	12.328	12.328	12.45	14.728	775.231
	2	12.489	12.354	12.354	12.463	14.732	14370.155
	3	12.433	12.364	12.364	12.466	14.733	12770.434
	4	12.41	12.368	12.368	12.467	14.733	8664.75
	5	12.396	12.37	12.37	12.467	14.733	4070.494
Minimum=				12.328			

Table 5.5: Parameter Settings of Chosen Kriging Models for Four Test Functions

Problem	Kriging Type	Covariance Function	Influence Distance (SD)	Order
Hmb	Simple	Cubic (2)	5	· · · · · ·
	Ordinary	Cubic (2)	5	· · · · · ·
	Taylor	Cubic (2)	4	3
SHC	Simple	Cubic (2)	5	· · · · · ·
	Ordinary	Cubic (2)	1	· · · · · ·
	Taylor	Cubic (2)	1	4
B2	Simple	Cubic (2)	2	· · · · · ·
	Ordinary	Cubic (2)	1	· · · · · ·
	Taylor	Nugget	0	3
Ras	Simple	Exponential	5	· · · · · ·
	Ordinary	Exponential	1	· · · · · ·
	Taylor	Exponential	1	

Table 5.6: ARE Comparison of Different Kriging Models for Four Test Functions (%)

Test Tunction	ARE	Simple Kriging	Ordinary Kriging	Taylor Kriging
Hmb	Average	1.305	1.301	1.158
	Minimum	0.003	0.001	0.000
	Maximum	10.219	10.196	9.424
SHC	Average	3.601	4.092	3.473
	Minimum	0.014	0.014	0.024
	Maximum	37.904	59.654	38.853
B2	Average	0.024	0.024	0.005
	Minimum	0.000	0.000	0.000
	Maximum	0.343	0.336	0.024
Ras	Average	12.396	12.328	12.328
	Minimum	0.11	0.158	0.158
	Maximum	42.022	41.715	41.482

Table 5.7: Paired t-test for Chosen Kriging Models in Table 5.1

	Variable 1	Variable 2
Mean	1.3034	1.1595
Degree of freedom		4
t_0		180.9594
P($T \geq t_0$) one-tail		0.0000
$t_{0.05}$ Critical one-tail		2.1318

Table 5.8: Paired t-test for Chosen Kriging Models in Table 5.2

	Variable 1	Variable 2
Mean	3.6377	3.4774
Degree of freedom		4
t_0		7.6133
P($T \geq t_0$) one-tail		0.0008
$t_{0.05}$ Critical one-tail		2.1318

and the TK models of order 3 with cubic (2) covariance function, respectively. The test results show that their AAREs are significantly different, and the TK models are better.

Similarly, Table 5.8 gives the t-test result of the AAREs from the SK models with cubic (2) covariance function and the TK models of order 4 with cubic (2) covariance function in Table 5.2 where Variables 1 and 2 are the SK models with cubic (2) covariance function and the TK models of order 4 with cubic (2) covariance function, respectively. The tested AAREs show that the TK models are better. For Table 5.3, the TK models are obviously best and there are no needs to conduct a statistical test.

5.4 Kriging Interpolation for Stochastic Simulation

5.4.1 Test Functions and Procedures

The following stochastic function is designed to analyze the stochastic simulation interpolation capabilities of SK, OK and TK:

$$F(\mathbf{X}) = J(\mathbf{X}) + \varepsilon(\mathbf{X}) \qquad (5.7)$$

$J(\mathbf{X})$ is a multimodal benchmark function which is simulated to generate trend data. The chosen benchmark functions are Hmb and SHC, respectively, and $\varepsilon(\mathbf{X})$ is a normal random variable.

The interpolation procedure is divided into four steps. The first step is similar to the deterministic simulation interpolation, and LHS is used to generate 150 simulation inputs in the feasible region of independent variables. Simulating $J(\mathbf{X})$ can generate their corresponding trend data. The function "randn" in Matlab is used to generate standard NIID (normal independent and identical distribution) random values which are then transformed to the values $\varepsilon(\mathbf{X})$ with $N(0, \sigma^2)$. Thus, 150 simulation outputs of $F(\mathbf{X})$ can be obtained according to Eq. (5.7). In the second step, the 150 input and output values of $F(\mathbf{X})$ are used to fit Kriging models with different parameter settings. In the third step, 50 inputs are stochastically sampled in the feasible region of independent variables. The stochastic sampling follows a uniform distribution. Similarly, the function "randn" in Matlab is used to generate random values of $\varepsilon(\mathbf{X})$. The 50 inputs and outputs of $F(\mathbf{X})$ are then utilized to test the fitted Kriging models. In the fourth step, a number of replications are conducted since $F(\mathbf{X})$ is stochastic. For the two constructed test problems, the number of replications

are 100. The norm L_2 of the errors of 50 predicted outputs for each replication is calculated according to the following formula:

$$\|\mathbf{X}_e\| = \sqrt{\sum_{i=1}^{50} x_{ei}^2} \tag{5.8}$$

The extreme values, means, standard deviations, and quantiles $Q_{0.1}$ and $Q_{0.9}$ of L_2 of 100 replications are then used to evaluate the stochastic simulation interpolation performance of different Kriging models.

5.4.2   Result Comparison

The L_2 comparison results of the replications from different Kriging models for Hmb and SHC functions are shown in Tables 5.9, 5.10, 5.11 5.12, (5.13, 5.14, 5.15 and 5.16). In these tables, "TK-O3-Cubic(2)-3" represents a TK model of order 3 with cubic (2) covariance function and influence distance of covariance equal to 3 standard deviations (SD). For other expressions, readers can refer to the explanation.

According to the mean columns of these tables, the best fitted Kriging models for Hmb and SHC are a TK model of order 4 with cubic (2) covariance function and influence distance of covariance equal to 1, and a TK model of order 2 with cubic (2) covariance function and influence distance of covariance equal to 1, respectively. Figures 5.5 and 5.6 are drawn to show the interpolation effectiveness of these two Kriging models. In the two figures, actual Y implies the average values of function outputs of 100 replications for each of 50 test points, and predicted Y represents the average values of predicted outputs from Kriging of 100 replications for each of 50 test points. The graphical trends show that the

means of actual and predicted outputs are very close, which indicates that the two Kriging models have good prediction capabilities.

5.4.3 Analysis on Replication and Variance

In the Kriging interpolation for stochastic simulation, there are two crucial parameters which need be investigated. They are the number of simulation replications and the variance of random variable $\varepsilon(\mathbf{X})$. That is, how many simulation replications are needed so that Kriging can provide valid mean outputs by interpolation? How does the variance of $\varepsilon(\mathbf{X})$ influence the interpolation performance of Kriging? For the simulation with costly computational expense, the selection of replication numbers becomes more important. It is necessary to limit the number of replications while obtaining valid Kriging interpolation values. This part is to test different numbers of simulation replications and different variance values of random variable $\varepsilon(\mathbf{X})$ to investigate their influence on Kriging interpolation performance.

For each point of 50 test samples, actual and predicted output means of 100 replications are calculated, and then the AREs of actual and predicted output means can be obtained. The average of 50 AREs from 50 test samples, that is AARE, is used to evaluate the interpolation performance of different Kriging models. The related results are shown in Tables 5.17 and 5.18.

Figures 5.7 and 5.8 are drawn to show the change trend of the AAREs for different replications and variance values. From the two figures, it can be concluded that as variance values increase, AAREs increase, and as replications increase, AAREs decrease. That is, a larger variance value for random variable $\varepsilon(\mathbf{X})$ makes Kriging interpolation less accurate,

Table 5.9: L_2 Comparison of Different Kriging Models for Hmb Function - 1

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
SK-Nugget-0	2697.059	2703.110	2700.274	1.162	2698.793	2701.664
SK-Linear-1	225.215	231.029	228.177	1.115	226.819	229.572
SK-Linear-2	224.786	230.599	227.747	1.115	226.390	229.143
SK-Linear-3	224.647	230.459	227.608	1.115	226.251	229.004
SK-Linear-4	224.578	230.390	227.539	1.115	226.183	228.935
SK-Linear-5	224.537	230.349	227.497	1.115	226.142	228.894
SK-Cubic(1)-1	44.953	53.165	48.923	1.602	46.899	50.687
SK-Cubic(1)-2	44.556	52.774	48.530	1.603	46.504	50.294
SK-Cubic(1)-3	44.420	52.627	48.400	1.602	46.374	50.177
SK-Cubic(1)-4	44.364	52.587	48.337	1.604	46.314	50.069
SK-Cubic(1)-5	44.257	52.489	48.299	1.604	46.282	50.081
SK-Cubic(2)-1	45.122	53.331	49.094	1.601	47.072	50.857
SK-Cubic(2)-2	44.646	52.861	48.621	1.603	46.594	50.384
SK-Cubic(2)-3	44.487	52.703	48.462	1.603	46.435	50.226
SK-Cubic(2)-4	44.402	52.628	48.383	1.604	46.353	50.145
SK-Cubic(2)-5	44.337	52.571	48.336	1.603	46.313	50.094
SK-Spherical-1	225.627	231.442	228.590	1.114	227.230	229.984
SK-Spherical-2	224.990	230.804	227.952	1.115	226.594	229.348
SK-Spherical-3	224.782	230.595	227.744	1.115	226.387	229.140
SK-Spherical-4	224.679	230.492	227.641	1.115	226.284	229.037
SK-Spherical-5	224.618	230.430	227.579	1.115	226.223	228.975
OK-Nugget-0	1826.512	1831.189	1828.974	0.987	1827.633	1830.200
OK-Linear-1	224.374	230.185	227.334	1.116	225.979	228.731
OK-Linear-2	224.374	230.185	227.334	1.116	225.979	228.731
OK-Linear-3	224.374	230.185	227.334	1.116	225.979	228.731
OK-Linear-4	224.374	230.185	227.334	1.116	225.979	228.731
OK-Linear-5	224.374	230.185	227.334	1.116	225.979	228.731
OK-Cubic(1)-1	44.179	52.397	48.159	1.603	46.129	49.919
OK-Cubic(1)-2	44.176	52.400	48.153	1.604	46.121	49.916
OK-Cubic(1)-3	44.183	52.384	48.151	1.603	46.121	49.922
OK-Cubic(1)-4	44.178	52.390	48.151	1.606	46.116	49.887
OK-Cubic(1)-5	44.148	52.284	48.155	1.601	46.141	49.873
OK-Cubic(2)-1	44.210	52.427	48.194	1.602	46.164	49.952
OK-Cubic(2)-2	44.183	52.402	48.164	1.603	46.133	49.925
OK-Cubic(2)-3	44.179	52.394	48.156	1.604	46.124	49.920
OK-Cubic(2)-4	44.167	52.395	48.153	1.605	46.119	49.915
OK-Cubic(2)-5	44.159	52.384	48.152	1.602	46.129	49.907

Table 5.10: L_2 Comparison of Different Kriging Models for Hmb Function - 2

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
OK-Spherical-1	224.344	230.155	227.304	1.116	225.949	228.701
OK-Spherical-2	224.366	230.178	227.327	1.116	225.971	228.724
OK-Spherical-3	224.370	230.182	227.331	1.116	225.976	228.728
OK-Spherical-4	224.372	230.183	227.332	1.116	225.977	228.730
OK-Spherical-5	224.373	230.184	227.333	1.116	225.978	228.730
TK-O1-Nugget-0	1826.512	1831.189	1828.974	0.987	1827.633	1830.200
TK-O1-Linear-1	224.374	230.185	227.334	1.116	225.979	228.731
TK-O1-Linear-2	224.374	230.185	227.334	1.116	225.979	228.731
TK-O1-Linear-3	224.374	230.185	227.334	1.116	225.979	228.731
TK-O1-Linear-4	224.374	230.185	227.334	1.116	225.979	228.731
TK-O1-Linear-5	224.374	230.185	227.334	1.116	225.979	228.731
TK-O1-Cubic(1)-1	44.179	52.397	48.159	1.603	46.129	49.919
TK-O1-Cubic(1)-2	44.176	52.400	48.153	1.604	46.121	49.916
TK-O1-Cubic(1)-3	44.183	52.384	48.151	1.603	46.121	49.922
TK-O1-Cubic(1)-4	44.178	52.390	48.151	1.606	46.116	49.887
TK-O1-Cubic(1)-5	44.148	52.284	48.155	1.601	46.141	49.873
TK-O1-Cubic(2)-1	44.210	52.427	48.194	1.602	46.164	49.952
TK-O1-Cubic(2)-2	44.183	52.402	48.164	1.603	46.133	49.925
TK-O1-Cubic(2)-3	44.179	52.394	48.156	1.604	46.124	49.920
TK-O1-Cubic(2)-4	44.167	52.395	48.153	1.605	46.119	49.915
TK-O1-Cubic(2)-5	44.159	52.384	48.152	1.602	46.129	49.907
TK-O1-Spherical-1	224.344	230.155	227.304	1.116	225.949	228.701
TK-O1-Spherical-2	224.366	230.178	227.327	1.116	225.971	228.724
TK-O1-Spherical-3	224.370	230.182	227.331	1.116	225.976	228.728
TK-O1-Spherical-4	224.372	230.183	227.332	1.116	225.977	228.730
TK-O1-Spherical-5	224.373	230.184	227.333	1.116	225.978	228.730
TK-O2-Nugget-0	1787.354	1792.224	1790.028	1.026	1788.667	1791.174
TK-O2-Linear-1	221.807	227.581	224.740	1.117	223.398	226.147
TK-O2-Linear-2	221.807	227.581	224.740	1.117	223.398	226.147
TK-O2-Linear-3	221.807	227.581	224.740	1.117	223.398	226.147
TK-O2-Linear-4	221.807	227.581	224.740	1.117	223.398	226.147
TK-O2-Linear-5	221.807	227.581	224.740	1.117	223.398	226.147

Table 5.11: L_2 Comparison of Different Kriging Models for Hmb Function - 3

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
TK-O2-Cubic(1)-1	44.170	52.389	48.146	1.604	46.116	49.908
TK-O2-Cubic(1)-2	44.168	52.390	48.146	1.604	46.114	49.910
TK-O2-Cubic(1)-3	44.190	52.372	48.146	1.603	46.123	49.894
TK-O2-Cubic(1)-4	44.162	52.389	48.144	1.601	46.129	49.889
TK-O2-Cubic(1)-5	44.114	52.414	48.152	1.605	46.132	49.880
TK-O2-Cubic(2)-1	44.192	52.412	48.169	1.604	46.138	49.931
TK-O2-Cubic(2)-2	44.173	52.395	48.152	1.604	46.122	49.914
TK-O2-Cubic(2)-3	44.173	52.391	48.149	1.604	46.120	49.918
TK-O2-Cubic(2)-4	44.175	52.394	48.148	1.604	46.118	49.903
TK-O2-Cubic(2)-5	44.168	52.390	48.147	1.603	46.122	49.902
TK-O2-Spherical-1	221.771	227.545	224.704	1.117	223.362	226.111
TK-O2-Spherical-2	221.798	227.572	224.731	1.117	223.389	226.138
TK-O2-Spherical-3	221.803	227.577	224.736	1.117	223.394	226.143
TK-O2-Spherical-4	221.805	227.578	224.738	1.117	223.396	226.145
TK-O2-Spherical-5	221.806	227.579	224.739	1.117	223.397	226.146
TK-O3-Nugget-0	972.240	975.865	974.217	0.875	973.060	975.299
TK-O3-Linear-1	165.315	171.507	168.544	1.210	167.073	170.082
TK-O3-Linear-2	165.315	171.507	168.544	1.210	167.073	170.082
TK-O3-Linear-3	165.315	171.507	168.544	1.210	167.073	170.082
TK-O3-Linear-4	165.315	171.507	168.544	1.210	167.073	170.082
TK-O3-Linear-5	165.315	171.507	168.544	1.210	167.073	170.082
TK-O3-Cubic(1)-1	38.465	46.799	42.486	1.622	40.398	44.188
TK-O3-Cubic(1)-2	38.458	46.807	42.486	1.623	40.396	44.184
TK-O3-Cubic(1)-3	38.474	46.795	42.486	1.624	40.387	44.184
TK-O3-Cubic(1)-4	38.518	46.790	42.485	1.621	40.391	44.170
TK-O3-Cubic(1)-5	38.487	46.831	42.491	1.623	40.337	44.228
TK-O3-Cubic(2)-1	38.448	46.783	42.471	1.622	40.382	44.172
TK-O3-Cubic(2)-2	38.461	46.793	42.482	1.622	40.393	44.183
TK-O3-Cubic(2)-3	38.464	46.794	42.484	1.622	40.395	44.182
TK-O3-Cubic(2)-4	38.461	46.802	42.486	1.623	40.397	44.181
TK-O3-Cubic(2)-5	38.484	46.800	42.486	1.623	40.397	44.191

Table 5.12: L_2 Comparison of Different Kriging Models for Hmb Function - 4

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
TK-O3-Spherical-1	165.308	171.500	168.537	1.210	167.066	170.074
TK-O3-Spherical-2	165.313	171.505	168.543	1.210	167.071	170.080
TK-O3-Spherical-3	165.314	171.506	168.544	1.210	167.072	170.081
TK-O3-Spherical-4	165.315	171.507	168.544	1.210	167.073	170.081
TK-O3-Spherical-5	165.315	171.507	168.544	1.210	167.073	170.081
TK-O4-Nugget-0	812.664	816.649	814.652	0.875	813.591	815.811
TK-O4-Linear-1	163.578	169.955	167.048	1.283	165.486	168.720
TK-O4-Linear-2	163.578	169.955	167.048	1.283	165.486	168.720
TK-O4-Linear-3	163.578	169.955	167.048	1.283	165.486	168.720
TK-O4-Linear-4	163.578	169.955	167.048	1.283	165.486	168.720
TK-O4-Linear-5	163.578	169.955	167.048	1.283	165.486	168.720
TK-O4-Cubic(1)-1	35.518	44.362	39.800	1.557	37.716	41.569
TK-O4-Cubic(1)-2	35.533	44.366	39.800	1.556	37.716	41.570
TK-O4-Cubic(1)-3	35.531	44.380	39.800	1.557	37.711	41.564
TK-O4-Cubic(1)-4	35.445	44.342	39.806	1.555	37.745	41.589
TK-O4-Cubic(1)-5	35.456	44.362	39.799	1.548	37.787	41.576
TK-O4-Cubic(2)-1	35.510	44.353	39.790	1.557	37.707	41.560
TK-O4-Cubic(2)-2	35.518	44.359	39.797	1.556	37.715	41.567
TK-O4-Cubic(2)-3	35.521	44.361	39.799	1.556	37.716	41.567
TK-O4-Cubic(2)-4	35.524	44.345	39.799	1.556	37.713	41.567
TK-O4-Cubic(2)-5	35.497	44.360	39.800	1.556	37.740	41.562
TK-O4-Spherical-1	163.571	169.948	167.041	1.283	165.479	168.713
TK-O4-Spherical-2	163.577	169.953	167.046	1.283	165.484	168.718
TK-O4-Spherical-3	163.578	169.954	167.047	1.283	165.485	168.719
TK-O4-Spherical-4	163.578	169.954	167.047	1.283	165.486	168.719
TK-O4-Spherical-5	163.578	169.954	167.047	1.283	165.486	168.719
Minimum=	35.445	44.342	39.790	0.875	37.707	41.560

Table 5.13: L_2 Comparison of Different Kriging Models for SHC Function - 1

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
SK-Nugget-0	170.537	175.568	172.799	1.030	171.755	174.322
SK-Linear-1	16.120	21.293	19.191	1.143	17.833	20.599
SK-Linear-2	15.940	21.122	19.023	1.142	17.658	20.430
SK-Linear-3	15.888	21.073	18.974	1.142	17.609	20.380
SK-Linear-4	15.862	21.050	18.951	1.142	17.586	20.357
SK-Linear-5	15.848	21.036	18.937	1.142	17.572	20.343
SK-Cubic(1)-1	8.649	19.362	13.436	2.126	10.811	16.181
SK-Cubic(1)-2	8.659	19.323	13.414	2.130	10.770	16.179
SK-Cubic(1)-3	8.664	19.307	13.409	2.132	10.764	16.182
SK-Cubic(1)-4	8.667	19.299	13.407	2.133	10.762	16.183
SK-Cubic(1)-5	8.668	19.294	13.406	2.134	10.761	16.185
SK-Cubic(2)-1	8.647	19.400	13.426	2.123	10.810	16.171
SK-Cubic(2)-2	8.658	19.337	13.414	2.129	10.774	16.179
SK-Cubic(2)-3	8.663	19.316	13.410	2.131	10.764	16.182
SK-Cubic(2)-4	8.665	19.305	13.408	2.132	10.762	16.183
SK-Cubic(2)-5	8.667	19.299	13.407	2.133	10.761	16.184
SK-Spherical-1	16.244	21.415	19.303	1.144	17.954	20.713
SK-Spherical-2	16.007	21.183	19.085	1.142	17.721	20.492
SK-Spherical-3	15.932	21.115	19.016	1.142	17.650	20.422
SK-Spherical-4	15.896	21.081	18.982	1.142	17.617	20.388
SK-Spherical-5	15.875	21.061	18.962	1.142	17.597	20.368
OK-Nugget-0	139.172	143.610	141.195	0.961	139.741	142.287
OK-Linear-1	15.791	20.984	18.884	1.141	17.520	20.290
OK-Linear-2	15.791	20.984	18.884	1.141	17.520	20.290
OK-Linear-3	15.791	20.984	18.884	1.141	17.520	20.290
OK-Linear-4	15.791	20.984	18.884	1.141	17.520	20.290
OK-Linear-5	15.791	20.984	18.884	1.141	17.520	20.290
OK-Cubic(1)-1	8.676	19.269	13.404	2.138	10.759	16.200
OK-Cubic(1)-2	8.675	19.272	13.403	2.137	10.758	16.194
OK-Cubic(1)-3	8.675	19.272	13.403	2.137	10.758	16.192
OK-Cubic(1)-4	8.675	19.272	13.403	2.137	10.758	16.192
OK-Cubic(1)-5	8.675	19.273	13.403	2.137	10.758	16.191
OK-Cubic(2)-1	8.672	19.324	13.386	2.132	10.746	16.187
OK-Cubic(2)-2	8.675	19.284	13.399	2.136	10.755	16.193
OK-Cubic(2)-3	8.675	19.278	13.401	2.136	10.757	16.193
OK-Cubic(2)-4	8.675	19.275	13.402	2.137	10.757	16.192
OK-Cubic(2)-5	8.675	19.274	13.402	2.137	10.757	16.192

Table 5.14: L_2 Comparison of Different Kriging Models for SHC Function - 2

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
OK-Spherical-1	15.731	20.927	18.826	1.141	17.460	20.233
OK-Spherical-2	15.776	20.969	18.870	1.141	17.505	20.275
OK-Spherical-3	15.784	20.977	18.878	1.141	17.513	20.283
OK-Spherical-4	15.787	20.980	18.881	1.141	17.516	20.286
OK-Spherical-5	15.789	20.982	18.882	1.141	17.517	20.288
TK-O1-Nugget-0	139.172	143.610	141.195	0.961	139.741	142.287
TK-O1-Linear-1	15.791	20.984	18.884	1.141	17.520	20.290
TK-O1-Linear-2	15.791	20.984	18.884	1.141	17.520	20.290
TK-O1-Linear-3	15.791	20.984	18.884	1.141	17.520	20.290
TK-O1-Linear-4	15.791	20.984	18.884	1.141	17.520	20.290
TK-O1-Linear-5	15.791	20.984	18.884	1.141	17.520	20.290
TK-O1-Cubic(1)-1	8.676	19.269	13.404	2.138	10.759	16.200
TK-O1-Cubic(1)-2	8.675	19.272	13.403	2.137	10.758	16.194
TK-O1-Cubic(1)-3	8.675	19.272	13.403	2.137	10.758	16.192
TK-O1-Cubic(1)-4	8.675	19.272	13.403	2.137	10.758	16.192
TK-O1-Cubic(1)-5	8.675	19.273	13.403	2.137	10.758	16.191
TK-O1-Cubic(2)-1	8.672	19.324	13.386	2.132	10.746	16.187
TK-O1-Cubic(2)-2	8.675	19.284	13.399	2.136	10.755	16.193
TK-O1-Cubic(2)-3	8.675	19.278	13.401	2.136	10.757	16.193
TK-O1-Cubic(2)-4	8.675	19.275	13.402	2.137	10.757	16.192
TK-O1-Cubic(2)-5	8.675	19.274	13.402	2.137	10.757	16.192
TK-O1-Spherical-1	15.731	20.927	18.826	1.141	17.460	20.233
TK-O1-Spherical-2	15.776	20.969	18.870	1.141	17.505	20.275
TK-O1-Spherical-3	15.784	20.977	18.878	1.141	17.513	20.283
TK-O1-Spherical-4	15.787	20.980	18.881	1.141	17.516	20.286
TK-O1-Spherical-5	15.789	20.982	18.882	1.141	17.517	20.288
TK-O2-Nugget-0	140.294	145.073	142.598	0.962	141.147	143.698
TK-O2-Linear-1	15.752	20.964	18.860	1.142	17.490	20.267
TK-O2-Linear-2	15.752	20.964	18.860	1.142	17.490	20.267
TK-O2-Linear-3	15.752	20.964	18.860	1.142	17.490	20.267
TK-O2-Linear-4	15.752	20.964	18.860	1.142	17.490	20.267
TK-O2-Linear-5	15.752	20.964	18.860	1.142	17.490	20.267

Table 5.15: L_2 Comparison of Different Kriging Models for SHC Function - 3

Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
TK-O2-Cubic(1)-1	8.675	19.273	13.402	2.136	10.757	16.190
TK-O2-Cubic(1)-2	8.675	19.273	13.402	2.136	10.757	16.190
TK-O2-Cubic(1)-3	8.675	19.273	13.402	2.136	10.757	16.190
TK-O2-Cubic(1)-4	8.675	19.273	13.402	2.136	10.757	16.190
TK-O2-Cubic(1)-5	8.675	19.273	13.402	2.136	10.757	16.190
TK-O2-Cubic(2)-1	8.668	19.332	13.383	2.129	10.743	16.168
TK-O2-Cubic(2)-2	8.673	19.288	13.398	2.135	10.754	16.184
TK-O2-Cubic(2)-3	8.674	19.280	13.400	2.136	10.756	16.187
TK-O2-Cubic(2)-4	8.674	19.277	13.401	2.136	10.756	16.188
TK-O2-Cubic(2)-5	8.674	19.275	13.402	2.136	10.757	16.189
TK-O2-Spherical-1	15.690	20.906	18.801	1.142	17.429	20.207
TK-O2-Spherical-2	15.736	20.949	18.845	1.142	17.475	20.252
TK-O2-Spherical-3	15.745	20.958	18.854	1.142	17.483	20.260
TK-O2-Spherical-4	15.748	20.960	18.857	1.142	17.486	20.263
TK-O2-Spherical-5	15.749	20.962	18.858	1.142	17.488	20.264
TK-O3-Nugget-0	79.911	85.980	82.950	1.038	81.727	84.230
TK-O3-Linear-1	12.421	17.847	15.666	1.137	14.118	17.119
TK-O3-Linear-2	12.421	17.847	15.666	1.137	14.118	17.119
TK-O3-Linear-3	12.421	17.847	15.666	1.137	14.118	17.119
TK-O3-Linear-4	12.421	17.847	15.666	1.137	14.118	17.119
TK-O3-Linear-5	12.421	17.847	15.666	1.137	14.118	17.119
TK-O3-Cubic(1)-1	8.741	19.004	13.496	2.193	10.812	16.340
TK-O3-Cubic(1)-2	8.741	19.004	13.496	2.193	10.812	16.340
TK-O3-Cubic(1)-3	8.741	19.004	13.496	2.193	10.812	16.340
TK-O3-Cubic(1)-4	8.741	19.004	13.496	2.193	10.812	16.340
TK-O3-Cubic(1)-5	8.741	19.004	13.496	2.193	10.812	16.340
TK-O3-Cubic(2)-1	8.742	19.012	13.497	2.195	10.812	16.347
TK-O3-Cubic(2)-2	8.741	19.006	13.497	2.193	10.812	16.342
TK-O3-Cubic(2)-3	8.741	19.005	13.496	2.193	10.812	16.341
TK-O3-Cubic(2)-4	8.741	19.005	13.496	2.193	10.812	16.341
TK-O3-Cubic(2)-5	8.741	19.004	13.496	2.193	10.812	16.340

Table 5.16: L_2 Comparison of Different Kriging Models for SHC Function - 4

| Kriging Models	Min	Max	Mean	Stdev	Percentile 0.1	Percentile 0.9
TK-O3-Spherical-1	12.413	17.840	15.659	1.137	14.111	17.112
TK-O3-Spherical-2	12.419	17.845	15.664	1.137	14.116	17.118
TK-O3-Spherical-3	12.420	17.846	15.665	1.137	14.117	17.119
TK-O3-Spherical-4	12.420	17.846	15.666	1.137	14.118	17.119
TK-O3-Spherical-5	12.421	17.846	15.666	1.137	14.118	17.119
TK-O4-Nugget-0	80.460	86.396	83.350	1.013	82.126	84.631
TK-O4-Linear-1	12.172	17.751	15.517	1.154	14.028	17.070
TK-O4-Linear-2	12.172	17.751	15.517	1.154	14.028	17.070
TK-O4-Linear-3	12.172	17.751	15.517	1.154	14.028	17.070
TK-O4-Linear-4	12.172	17.751	15.517	1.154	14.028	17.070
TK-O4-Linear-5	12.172	17.751	15.517	1.154	14.028	17.070
TK-O4-Cubic(1)-1	8.788	19.083	13.501	2.209	10.810	16.465
TK-O4-Cubic(1)-2	8.788	19.083	13.501	2.209	10.810	16.465
TK-O4-Cubic(1)-3	8.788	19.083	13.501	2.209	10.810	16.465
TK-O4-Cubic(1)-4	8.788	19.083	13.501	2.209	10.810	16.465
TK-O4-Cubic(1)-5	8.788	19.083	13.501	2.209	10.810	16.465
TK-O4-Cubic(2)-1	8.789	19.089	13.502	2.211	10.810	16.472
TK-O4-Cubic(2)-2	8.788	19.085	13.501	2.209	10.810	16.467
TK-O4-Cubic(2)-3	8.788	19.084	13.501	2.209	10.810	16.466
TK-O4-Cubic(2)-4	8.788	19.084	13.501	2.209	10.810	16.465
TK-O4-Cubic(2)-5	8.788	19.084	13.501	2.209	10.810	16.465
TK-O4-Spherical-1	12.166	17.745	15.510	1.154	14.023	17.063
TK-O4-Spherical-2	12.171	17.749	15.515	1.154	14.027	17.068
TK-O4-Spherical-3	12.172	17.750	15.516	1.154	14.028	17.069
TK-O4-Spherical-4	12.172	17.751	15.517	1.154	14.028	17.069
TK-O4-Spherical-5	12.172	17.751	15.517	1.154	14.028	17.069
Minimum=	8.647	17.745	13.383	0.961	10.743	16.168
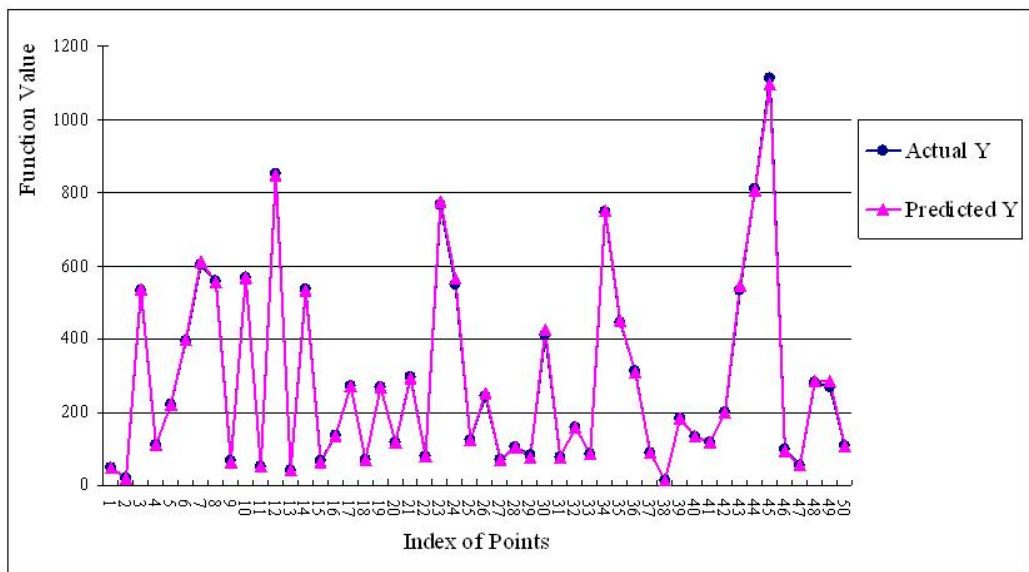
Figure 5.5: Actual and Predicted Output Means of 100 Replications of 50 Test Samples (Hmb)
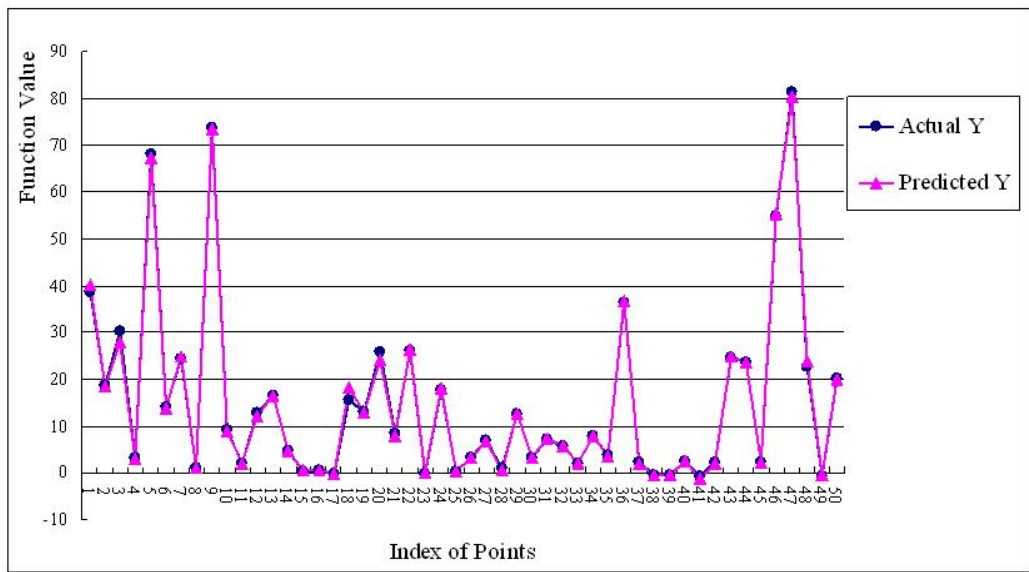
Figure 5.6: Actual and Predicted Output Means of 100 Replications of 50 Test Samples (SHC)

Table 5.17: AARE Comparison of Different Replications and Variances for Hmb

Replications	AARE (%)			
Number	$N(0, 1^2)$	$N(0, 5^2)$	$N(0, 10^2)$	$N(0, 50^2)$
10	1.55	3.11	4.97	17.59
20	1.37	2.39	3.65	13.26
30	1.31	2.27	3.59	14.00
40	1.26	1.97	3.12	12.84
50	1.25	1.68	2.42	9.04
60	1.22	1.60	2.40	9.48
70	1.22	1.57	2.29	9.01
80	1.21	1.64	2.42	10.05
90	1.21	1.59	2.31	9.90
100	1.19	1.49	2.11	8.82

Table 5.18: AARE Comparison of Different Replications and Variances for SHC

Replications	AARE (%)			
Number	$N(0, 0.2^2)$	$N(0, 1^2)$	$N(0, 5^2)$	$N(0, 10^2)$
10	7.74	75.57	394.41	120.92
20	7.37	908.90	67.46	131.10
30	5.34	35.51	280.64	106.84
40	4.75	54.85	47.96	100.37
50	5.01	87.33	84.20	219.21
60	4.95	31.06	93.50	228.74
70	5.08	22.24	116.89	177.93
80	4.24	14.18	38.66	100.37
90	3.96	12.00	53.06	299.97
100	3.90	11.11	57.19	108.86

and more replications result in more accurate interpolation, as expected. Whether a variance value is too large depends on $J(\mathbf{X})$. For example, in Figure 5.5 the output means from Hmb are around 200, and a variance value equal to 10^2 in Figure 5.7 does not result in a severe volatile AAREs; however, in Figure 5.6, the output means from SHC are around 10, and Figure 5.8 shows a variance value equal to 10 makes AAREs very unstable, which indicates that a variance equaling 10^2 is unacceptable for SHC. In Figure 5.7, it seems that a variance value equal to 50^2 is too large for Hmb. Regarding replications, these two figures show that when the number of replications is greater than 50, the corresponding AAREs become stable, which implies that 50 may be a good threshold for the selection of replication numbers.

5.5 Conclusions

The chapter empirically investigated the deterministic and stochastic simulation interpolation capabilities of SK, OK and TK. The results are that TK has the best interpolation performance for both deterministic and stochastic simulations. For stochastic simulations, it is noted that 50 may be a good threshold for the selection of simulation replication numbers, and running a number of replications greater than 50 can improve the Kriging interpolation accuracy for simulation output means. The selection of replication numbers is important for stochastic simulation with costly computational expense because it is necessary to limit the number of replications to reduce computational expenses. In addition, the variance of random variable $\varepsilon(\mathbf{X})$ in a stochastic simulation can influence the interpolation performance for output means. The empirical result shows that a large variance makes interpolation less accurate, as expected.

Figure 5.7: Actual and Predicted Output Means of Different Replications of 50 Test Samples under Different Variance Levels of Random Variable $\varepsilon(\mathbf{X})$ (Hmb)
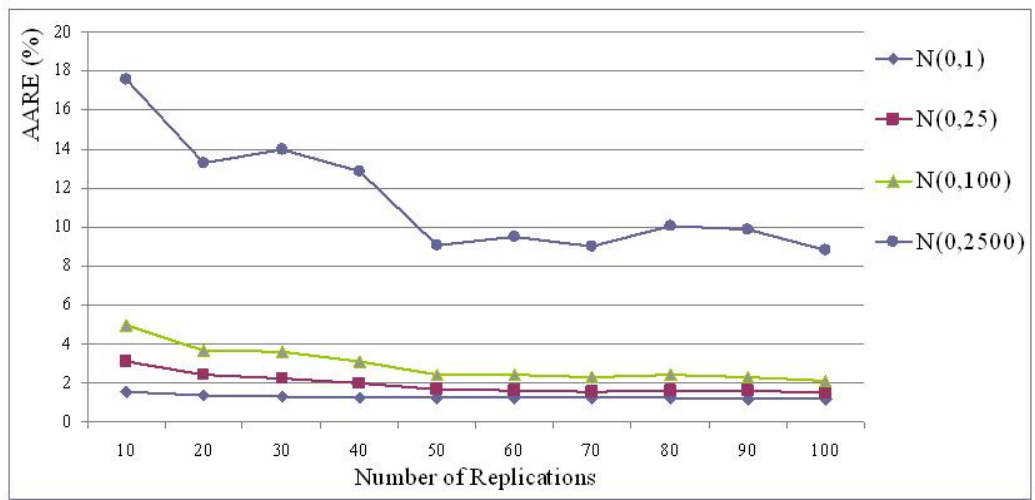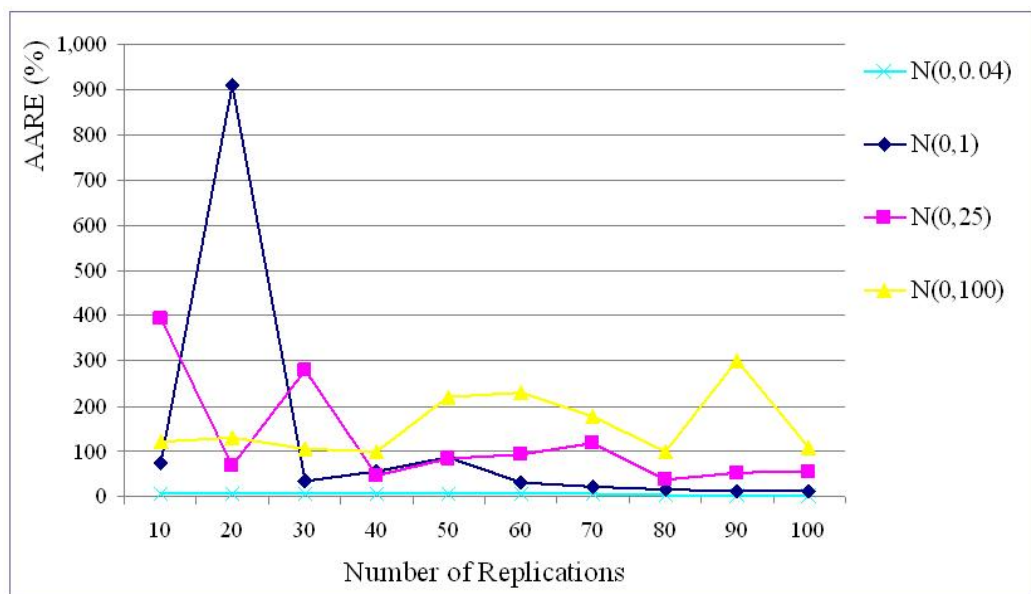
Figure 5.8: Actual and Predicted Output Means of Different Replications of 50 Test Samples under Different Variance Levels of Random Variable $\varepsilon(\mathbf{X})$ (SHC)

Kriging Metamodeling For Sensitivity Analysis On Simulation Models

6.1 Introduction

Sensitivity analysis is a key problem of simulation modeling. Sensitivity analysis is to quantify the influence of input factors on outputs, and to find the input factors in a simulation model which have critical influence on outputs. The results of sensitivity analysis can help decision makers to select optimal settings for a simulation model. One method of conducting sensitivity analysis is an analytical approach, that is, using the analytical feature of a model to perform sensitivity analysis. However, a simulation model is usually not explicit or invisible and thus not differentiable. Using the analytical approach to directly perform sensitivity analysis on a simulation model is difficult. An alternative is to use the analytical feature of its corresponding fitted metamodel to assist in sensitivity analysis of a simulation model. The fitted metamodel must be an effective replacement of the corresponding simulation model.

Kriging is a precise interpolator and it is explicit and differentiable. Using a Kriging metamodel to assist in sensitivity analysis on a simulation model is methodically feasible, which, however, has not been seen in the literature. This chapter will empirically investigate this problem. Besides Kriging, Regression Analysis (RA) is a simple and effective interpolation tool. RA is straightforward and differentiable, and thus a regression metamodel can be used to assist in sensitivity analysis on its corresponding simulation model. The chapter will effectively compare the potentials of Kriging with those of RA in this aspect.

A physical simulation case based on cost estimation is chosen for the empirical investigation. Physical simulation means that its simulator is a scaled replica of a real system, and a simulation is the actual operation of the system [37]. The main goal of physical simulation is to evaluate inputs and find critical input factors. Physical simulation has wide applications. For example, it was used to simulate a wastewater treatment system to evaluate its features and economic feasibilities [103], and in industrial production it can be used to simulate manufacturing systems to determine their product qualities and costs [19, 40]. However, similar to computer simulation with costly computational expense, physical simulation has the same difficulties in obtaining data of adequate sample size due to economic and time constraints. It has the realistic meaning to use a metamodel to perform sensitivity analysis of a physical simulation model to find critical input factors.

In order to conduct sensitivity analysis, the first step is to fit an effective metamodel. Based on this procedure, the chapter is organized as follows. First, a fitted Taylor Kriging (TK) is chosen for the above physical simulation. Next, the interpolation accuracy and effectiveness of Kriging, RA and artificial neural networks (ANN) are compared. While establishing regression metamodels, multicollinearity, heteroscedasticity and specification errors are the focus other than the regression metamodels themselves. The goal is to illuminate how to develop an effective regression metamodel so that the comparison between Kriging and RA will be effective. The results of ANN are cited from a reference. Second, after the fitted Kriging and regression metamodels are chosen, the sensitivity analysis assisted by them on the physical simulation model is formulated, and the results are analyzed. Third, according to the empirical results, several important theoretical properties of

Kriging, RA and ANN for simulation problems are discussed. These properties include the applicability of model, accuracy of interpolation, and feasibility of sensitivity analysis.

6.2 The Chosen Simulation Problem

The chosen simulation problem is a physical simulation case related to cost estimation (interpolation). Cost estimation is used to establish an initial budget that assures an expected profit. For a product, estimating its cost requires the production of some alternatives in order to collect cost data. However, because of budget constraint, it is difficult to obtain cost-driven data of adequate sample sizes for evaluation and comparison. Thus, using some mathematical methods to estimate the costs of some alternatives is necessary. Especially, in order to manage and control cost, recognizing key cost factors is important, which requires that the chosen method should have the capability to perform sensitivity analysis. Kriging has the potentials in this aspect.

The case is the pressure vessel cost as a function of the height, diameter and wall thickness obtained from a manufacturer. The sample data are listed in Table 6.1. Scatter diagrams of the height, diameter and thickness of products are created to show the distribution features of the sample data. In the scatter plots, the height and actual cost of sample 1, the diameters of samples 6 and 19, and the height, thickness and actual cost of sample 20 are the extreme values. These extreme values imply that the estimation methods have to extrapolate the data rather than interpolate them. However, extrapolation tends to yield a less precise and more erratic estimate and should be avoided.

Table 6.1: Table of Original Cost Data

Sample	Height	Diameter	Thickness	Actual Cost
1	1200	1066	10	$10,754
2	4500	1526	15	$18,172
3	6500	1500	16	$23,605
4	12250	1200	12	$23,956
5	21800	1050	12	$28,400
6	23300	900	14	$31,400
7	26700	1500	15	$42,200
8	12100	3000	11	$47,970
9	17500	2400	12	$48,000
10	26500	1348	14	$51,000
11	28300	1800	14	$53,900
12	14700	2400	10	$54,600
13	26600	1500	15	$58,040
14	24800	2500	13	$61,790
15	25000	2100	14	$61,800
16	24700	2000	16	$67,460
17	29500	2250	13	$80,400
18	21900	3150	12	$85,750
19	32300	5100	17	$207,800
20	53500	3000	29	$240,000
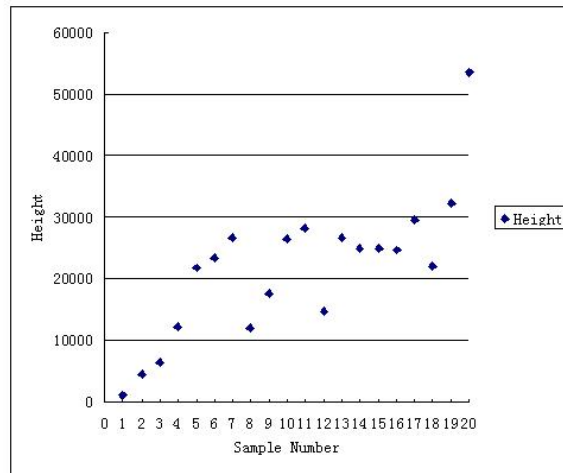
Figure 6.1: Scatter Figure of Sample Height
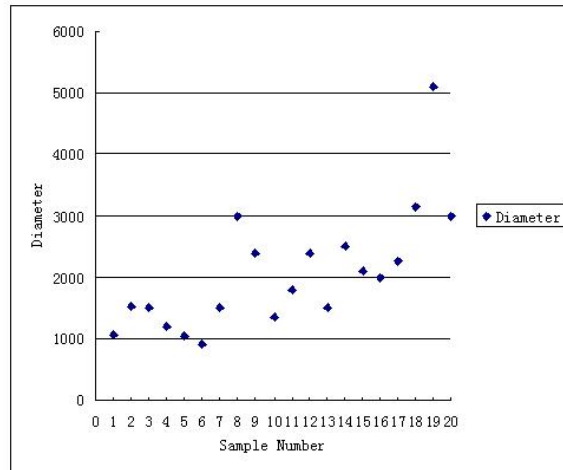
Figure 6.2: Scatter Figure of Sample Diameter
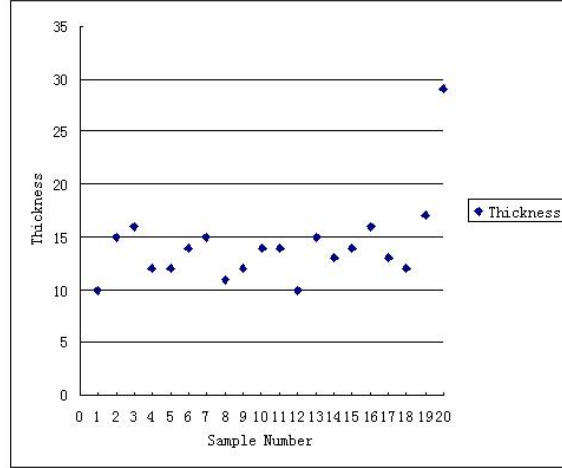
Figure 6.3: Scatter Figure of Sample Thickness

Since the data of height, diameter and thickness have different magnitudes, they are normalized to the interval [0, 1]. The normalization formula is as follows:

$$x_i^{'} = \frac{x_i - min(x_i)}{max(x_i) - min(x_i)} \qquad (6.1)$$

where $max(x_i)$ and $min(x_i)$ mean the maximum and minimum values that x_i can be, respectively. Thus, $x_i^{'}$ lies within the interval [0, 1]. The inverse process of (6.1) is used to get actual costs. Certainly, the output of metamodels may occur outside the interval [0, 1]. If it happens, the actual value is set to 0 if the output is less than 0, and adjusted to 1 if greater than 1. The normalized data are shown in Table 6.2.

The cross validation method using jackknife is adopted to avoid data re-substitution while metamodels are fitted. Twenty samples from [10] are used to establish twenty groups,

Table 6.2: Table of Normalized Cost Data

Sample	Height	Diameter	Thickness	Normalized Cost
1	0.0000	0.0395	0.0000	0.0000
2	0.0631	0.1490	0.2632	0.0324
3	0.1013	0.1429	0.3158	0.0561
4	0.2113	0.0714	0.1053	0.0576
5	0.3939	0.0357	0.1053	0.0770
6	0.4226	0.0000	0.2105	0.0901
7	0.4876	0.1429	0.2632	0.1372
8	0.2084	0.5000	0.0526	0.1623
9	0.3117	0.3571	0.1053	0.1625
10	0.4837	0.1067	0.2105	0.1756
11	0.5182	0.2143	0.2105	0.1882
12	0.2581	0.3571	0.0000	0.1913
13	0.4857	0.1429	0.2632	0.2063
14	0.4512	0.3810	0.1579	0.2226
15	0.4551	0.2857	0.2105	0.2227
16	0.4493	0.2619	0.3158	0.2474
17	0.5411	0.3214	0.1579	0.3038
18	0.3958	0.5357	0.1053	0.3271
19	0.5946	1.0000	0.3684	0.8595
20	1.0000	0.5000	1.0000	1.0000

and each group only has a sample. Any nineteen groups are adopted to estimate the cost of the remaining one-sample group.

6.3   Development of Metamodels

6.3.1   Kriging Metamodeling

Each product design (Ht, Dr, Ts) is regarded as a point in a three-dimensional space, where Ht, Dr and Ts imply the height, diameter and thickness of product, respectively. The covariance function of Kriging is supposed to only depend on the distance of two points relative to each other, and is not influenced by their specific locations, i.e., the covariance structure is homogeneous and isotropic. Simple Kriging (SK), Ordinary Kriging (OK) and TK are used to estimate the costs, respectively. For these Kriging models, the covariance functions (3.34, 3.35, 3.36, 3.37, 3.38), respectively, are adopted to construct Kriging coefficient matrices. In these covariance functions, five different levels of covariance influence distance are tested. The average absolute relative error (AARE) of the twenty estimated values is used to evaluate the performance of different metamodels.

The results of the Kriging models with different parameter settings are shown in Table 6.3. Table 6.3 shows that a TK model of order 3 with the nugget covariance function has the smallest AARE percentage, which indicates that this Kriging model better describes the relationship between the cost and the height, diameter and thickness and is chosen to estimate product cost. The estimation results are shown in Table 6.4 where the minimal and maximal AREs are 0.00% and 30.99%, respectively.

Figure 6.4 is drawn to describe the graphical trend change of actual and estimated costs of twenty samples in Table 6.4. Figure 6.4 indicates that the estimation accuracy

Table 6.3: AAREs of Cost Estimation from Different Kriging Models (%)

Covariance Function	SD	SK	OK	Order of Taylor Kriging		
				1	2	3
Nugget	0	72.12	91.65	91.65	26.39	13.31
Linear	1	33.20	75.44	75.44	19.51	16.69
	2	27.81	55.60	55.60	14.95	14.71
	3	25.74	44.16	44.16	15.65	15.90
	4	29.46	35.85	35.85	15.56	16.00
	5	23.36	29.89	29.89	15.21	15.93
Cubic(1)	1	98.91	120.91	120.91	111.00	241.11
	2	236.59	227.83	227.83	125.70	200.68
	3	117.46	121.31	121.31	120.31	158.10
	4	113.05	116.13	116.13	123.78	159.59
	5	115.59	117.82	117.82	121.96	160.98
Cubic(2)	1	50.74	91.47	91.47	45.85	95.99
	2	88.09	98.67	98.67	83.07	241.20
	3	110.70	112.42	112.42	103.91	180.84
	4	110.05	112.15	112.15	110.99	168.44
	5	113.87	116.54	116.54	115.65	165.84
Spherical	1	36.66	81.74	81.74	20.13	16.53
	2	28.66	61.59	61.59	15.28	16.32
	3	24.36	51.21	51.21	15.59	16.06
	4	26.09	43.65	43.65	15.54	16.03
	5	24.10	37.35	37.35	15.50	16.03

Table 6.4: AREs from TK of Order 3 with Nugget Covariance Function

Sample	Actual Cost	Estimated Cost	ARE
1	10754	10754	0.00%
2	18172	17267	4.98%
3	23605	26047	10.35%
4	23956	24293	1.41%
5	28400	35749	25.88%
6	31400	37056	18.01%
7	42200	53200	26.07%
8	47970	56316	17.40%
9	48000	51949	8.23%
10	51000	42845	15.99%
11	53900	59936	11.20%
12	54600	43069	21.12%
13	58040	48614	16.24%
14	61790	74810	21.07%
15	61800	60472	2.15%
16	67460	55505	17.72%
17	80400	68385	14.94%
18	85750	87777	2.36%
19	207800	143405	30.99%
20	240000	240000	0.00%
AARE=			13.31%

Table 6.5: ANOVA of Actual and Kriging-estimated Costs

Source	DF	SeqSS	$A_{dj}SS$	$A_{dj}MS$	F	P
Treatment (Cost)	1	88645058	88645058	88645058	0.67	0.424
Block	19	1.11E+11	1.11E+11	5.83E+09	43.81	0
Error	19	2.53E+09	2.53E+09	1.33E+08		
Total	39	1.13E+11				
R-Sq=97.77%			R-Sq(adj)= 95.42%			

of the chosen Taylor Kriging is adequate. Besides the graphic trend comparison, analysis of variance (ANOVA) is used to test whether there is a significant difference between the actual and estimated costs in Table 6.4. The ANOVA results are in Table 6.5. The *P-value* in this table shows that statistically there is no significant difference between the actual and estimated costs, which confirms the conclusion from Figure 6.4. The R-square and adjusted R-square in Table 6.5 are more than 90%, indicating that the results from the chosen TK model are quite satisfactory. In this table, DF, $A_{dj}SS$ and $A_{dj}MS$ indicate degree of freedom, adjusted sum of square and adjusted mean square, respectively.

6.3.2   Regression Metamodeling

Gerrard *et al.* [10, 45] and Smith *et al.*[134] use Multivariable Linear Regression (MLR) to estimate the product cost by interpolation. However, their analysis does not consider multicollinearity, heteroscedasticity and specification errors, and does not discuss the possibility of adopting other regression models. Although this part uses MLR to estimate the product cost, multicollinearity, heteroscedasticity and specification errors, instead of MLR,
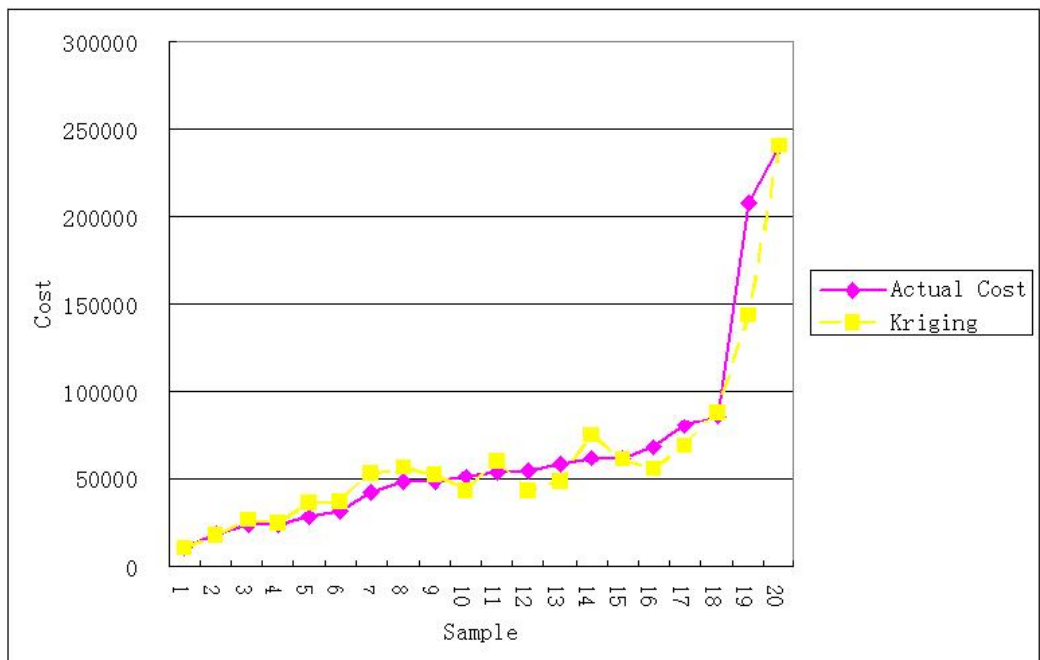
Figure 6.4: Comparison of Actual and Kriging-estimated Costs

are mainly considered. And how to choose an effective linear or nonlinear regression model is emphasized.

Multicollinearity

The Gauss-Markov theorem states that among all of the linear unbiased estimators, the least squares estimator has the smallest variance. However, it does not assure that the least squares estimator has a small variance in any absolute sense. If two regressors are perfectly correlated, the variance of the least squares estimator is infinite. The case of an exact or highly linear relationship among regressors is a serious failure of the model assumptions, not of the data. The problems caused by highly correlated regressors include the following symptoms:

1. Small changes in the data produce wide swings in parameter estimates.

2. Coefficients may have very high standard errors and low significance levels even though they are jointly significant, and the R^2 for the regression is quite high.

3. Coefficients may have implausible magnitudes.

The Fisher transformation is used to test the existence of multicollinearity in the three regressors which are the Height, Diameter and Thickness of product. The constructed hypothesis is: $H_0 : \rho_{ij} = 0$; $H_1 :$ *At least a ρ_{ij} is not 0*, where ρ_{ij} means the correlation coefficient between two explanatory variables x_i and x_j. The statistic adopted is

$$t = \frac{r_{ij}\sqrt{N-2}}{\sqrt{1-r_{ij}^2}} \sim t(N-2) \tag{6.2}$$

Table 6.6: Multicollinearity Results

Explanatory Variables		Fisher Transformation	
Ht and Dr	Ht and Ts Dr and Ts	Ht and Dr	Ht and Ts Dr and Ts
Correlation Coefficients		t_0	
0.3965	0.7217 0.2714	1.8324	4.4235 1.1961
$t_{0.05,18} = 1.7341$			

where r_{ij} is the estimator of ρ_{ij}. The specific test results are given in Table 6.6. Table 6.6 shows that any two explanatory variables except for the Diameter and Thickness are statistically significantly correlated under the significant level of 5% for the two-tailed t test. For the Diameter and Thickness, there is not sufficient evidence to support their significant correlation.

Heteroscedasticity

The presence of heteroscedascity often means that a misspecification in RA has taken place, which may be in terms of a missing interaction effect. In this part, the White's general test is formulated to test the heteroscedasticity of the cost data. First, the regression of Actual Cost (AC) on Ht, Dr and Ts is performed to obtain the estimation errors from the regression equation. Then, the regression of the square of estimation errors on Ht, Dr, Ts, Height Square of the Product (HS), Diameter Square of the Product (DS), Thickness Square of the Product (TS), Height $\times$ Diameter of the Product (HD), Height $\times$ Thickness of the Product (HT) and Diameter $\times$ Thickness of the Product (DT) is conducted, and the corresponding R^2 is 0.8388. A hypothesis is constructed: $H_0 : \sigma_i^2 = \sigma^2$ for all i; $H_1 :$

Not all σ_i^2 equal to σ^2. The statistic below is adopted to test the significance of hypothesis.

$$NR^2 \sim \chi^2(K-1) \tag{6.3}$$

where K is the number of explanatory variables (including the constant term) in the second regression equation. The test result is $\chi_{0.95}^2(9) = 16.92 > NR^2 = 16.7755$. According to the White's general test, the null hypothesis can not be rejected under the significant level of 5% and statistically there is insufficient evidence to support the significant existence of heteroscedasticity.

Specification Errors

Specification errors are caused by omitted variables, an incorrect functional form, or errors in measurement. Here the Ramsey regression equation specification error test (RESET) is used to examine whether the fitted MLR model has misspecification errors. For a detailed introduction to RESET, readers can refer to [57, 58, 114]. The RESET is performed as follows.

According to the regression of AC on Ht, Dr and Ts, the variables $\hat{AC}^2$, $\hat{AC}^3$ and $\hat{AC}^4$ are first created, where $\hat{AC}$ is the estimator of AC. Then, the regression of AC on Ht, Dr, Ts and $\hat{AC}^2$ is conducted. The hypothesis is:

$H_0:$ *The coefficient of $\hat{AC}^2$ is equal to zero;*

$H_1:$ *The coefficient of $\hat{AC}^2$ is not equal to zero.*

The F statistic is used to test the hypothesis:

$$F = \frac{(ESS_1 - ESS_2)/1}{ESS_2/(N-K-1)} \sim F(1, N-K-1) \tag{6.4}$$

where ESS_1 and ESS_2 are the Error Sum of Squares (ESS) from the first and second regression equations, respectively. The test result is its $P-value$ equal to $7.45E-07$. Similarly, the regression of AC on Ht, Dr, Ts, $\hat{AC}^2$ and $\hat{AC}^3$, and the regression of AC on Ht, Dr, Ts, $\hat{AC}^2$, $\hat{AC}^3$ and $\hat{AC}^4$ are performed. And their $P-values$ of corresponding F statistic are $7.18E-06$ and $6.44E-05$, respectively. All of these values show that there are statistically significant specification errors in the fitted MLR model.

An Alternative Regression Model

The statistically significant existences of multicollinearity and specification errors show that it is necessary to establish a better alternative for the MLR. The alternative meta-model developed is named Logarithmic Linear Regression (LLR) and mathematically it is expressed as (6.5).

$$Ln(AC) = \eta_0 + \eta_1 Ln(Ht) + \eta_2 Ln(Dr) + \eta_3 Ts \tag{6.5}$$

The J test is used to determine which of two regression metamodels, the MLR and LLR, is more valid. The test procedure consists of first regressing AC on Ht, Dr and Ts, and then Ln(AC) being regressed on Ln(Ht), Ln(Dr), Ts and Predicted Cost by MLR (PC). Suppose the regression coefficient of PC is ϕ. The hypothesis is: $H_0 : \phi = 0$; $H_1 : \phi \neq 0$, and the corresponding statistic is:

$$t = \frac{\hat{\phi}}{se(\hat{\phi})} \sim t(N-K) \tag{6.6}$$

91

where $se(\hat{\phi})$ is the standard error of the estimator $\hat{\phi}$. The test result is $t_0 = 3.23$ and the degree of freedom of t distribution is 15. Note $P\text{-}value = P(|t| > 3.23) = 0.0056$, which indicates that the J test is significant and statistically the first model is rejected.

6.3.3 Analysis on Different Metamodels

This section analyzes the results of the different metamodels and discusses their estimation capabilities. The results of the ANN are from [134]. The estimated values from the metamodels are in Table 6.7 where the estimated value of the first sample from the MLR is truncated to 0 from -30603 because the cost can not be a negative number.

Figure 6.5 shows the change in trends of the costs estimated by TK, LLR, MLR and ANN. It is clear that TK, LLR and ANN give the better results. Table 6.8 provides the ARE comparison to show the performance of different metamodels, and the means of AREs of 20 and 16 samples for each metamodel are given. The sample size 16 implies that samples 1, 6, 19 and 20 are excluded. These means investigate the average estimation effects of models and the possible invalidity of extrapolation. The AREs of 16 and 20 samples show TK is better than LLR and MLR, but worse than ANN. However, the means of TK, LLR and ANN are close, which, to some extent, demonstrates that these three metamodels have similar estimation capabilities. Figure 6.6 give the graphic trend changes of AREs from these models.

Furthermore, ANOVA is used to test whether there exist significant differences among AREs of TK, LLR, MLR(Truncated) and ANN. The results are listed in Table 6.9 where the $P\text{-}value$ shows that statistically the ARE differences of different metamodels are highly significant. Then the AREs from MLR (Truncated) are removed and ANOVA is conducted

Table 6.7: Comparison of Estimated Costs from TK, LLR, MLR and ANN

Sample	Actual Cost	TK	LLR	MLR	MLR (Truncated)	ANN
1	10754	10754	5003	-30603	0	10904
2	18172	17267	25566	33014	33014	22691
3	23605	26047	29554	42548	42548	23725
4	23956	24293	26210	9059	9059	22941
5	28400	35749	29873	17674	17674	29665
6	31400	37056	27894	27916	27916	33913
7	42200	53200	51216	60241	60241	52673
8	47970	56316	58555	65921	65921	53942
9	48000	51949	57371	57478	57478	49440
10	51000	42845	42781	46901	46901	47959
11	53900	59936	58677	65798	65798	60063
12	54600	43069	47150	38868	38868	40081
13	58040	48614	49430	58396	58396	53263
14	61790	74810	71941	77578	77578	72069
15	61800	60472	63528	70023	70023	64632
16	67460	55505	65403	78381	78381	63756
17	80400	68385	68523	73872	73872	69240
18	85750	87777	78130	87376	87376	81911
19	207800	143405	164557	177543	177543	208266
20	240000	240000	176939	185351	185351	217750

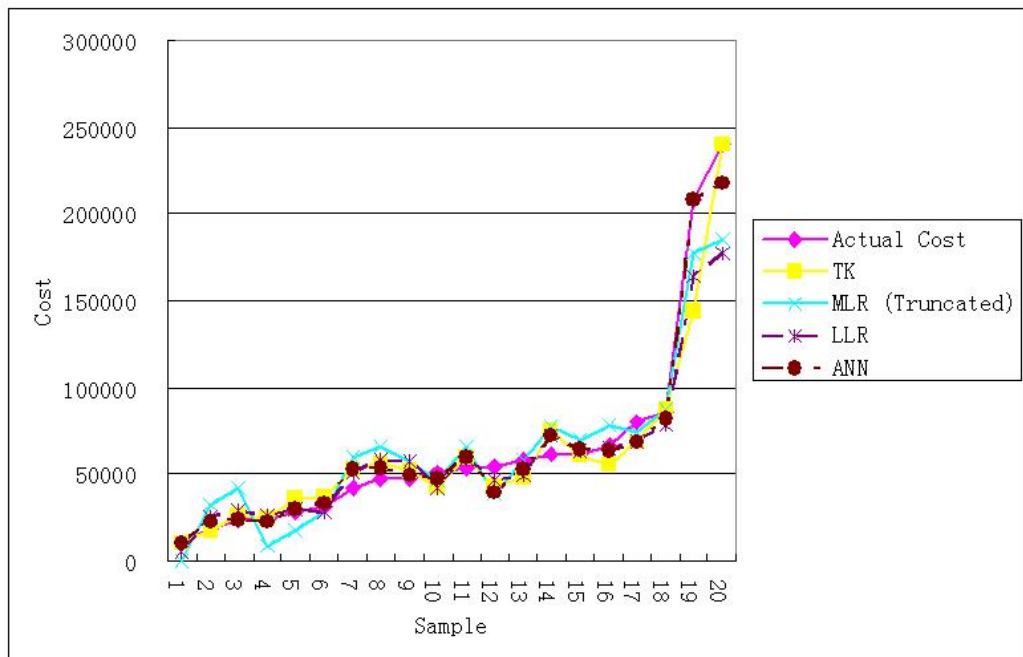Figure 6.5: Comparison of Estimated Costs from Different Models

94

Table 6.8: Comparison of AREs from TK, MLR, LLR and ANN (%)

| Sample	TK	LLR	MLR	MLR (Truncated)	ANN
1	0.00	53.48	384.57	100.00	1.39
2	4.98	40.69	81.64	81.64	24.87
3	10.35	25.20	80.23	80.23	0.51
4	1.41	9.41	62.18	62.18	4.24
5	25.88	5.19	37.78	37.78	4.45
6	18.01	11.17	11.11	11.11	8.00
7	26.07	21.36	42.75	42.75	24.82
8	17.40	22.07	37.42	37.42	12.45
9	8.23	19.52	19.74	19.74	3.00
10	15.99	16.12	8.04	8.04	5.96
11	11.20	8.86	22.07	22.07	11.43
12	21.12	13.64	28.82	28.82	26.59
13	16.24	14.83	0.61	0.61	8.23
14	21.07	16.43	25.55	25.55	16.64
15	2.15	2.80	13.30	13.30	4.58
16	17.72	3.05	16.19	16.19	5.49
17	14.94	14.77	8.12	8.12	13.88
18	2.36	8.89	1.90	1.90	4.48
19	30.99	20.81	14.56	14.56	0.22
20	0.00	26.28	22.77	22.77	9.27
Mean (20 Samples)	13.31	17.73	45.97	31.74	9.53
Mean (16 Samples)	13.57	15.18	30.40	30.40	10.73
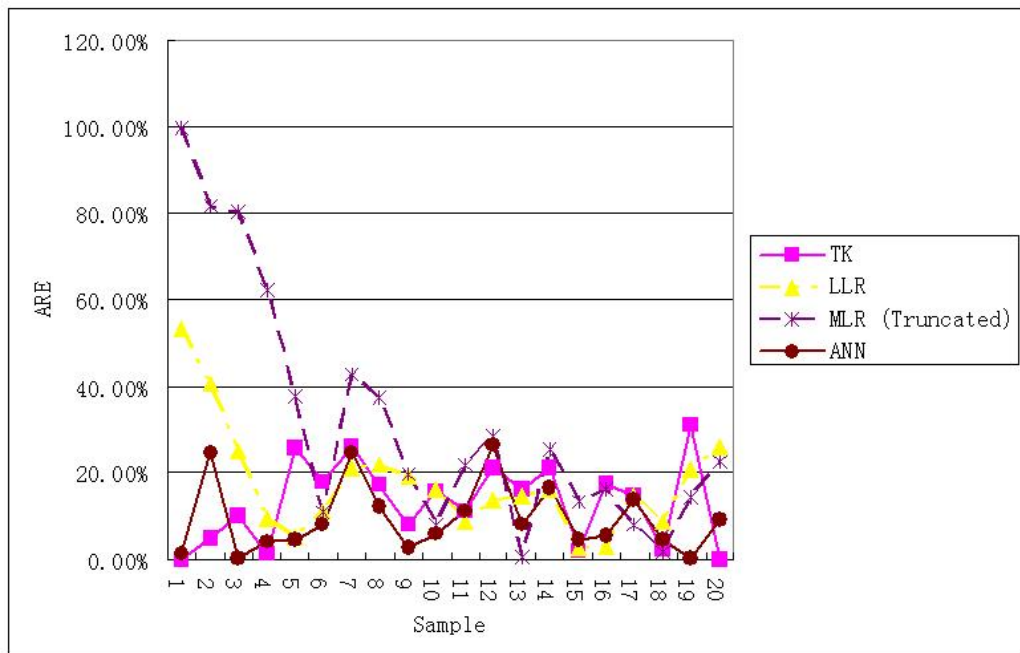
Figure 6.6: Comparison of AREs from Different Models

Table 6.9: ANOVA of AREs from TK, LLR, MLR and ANN

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Models	3	5654.7	5654.7	1884.9	6.79	0.000
Error	76	21087.4	21087.4	277.5		
Total	79	26742.1				

Table 6.10: ANOVA of AREs from TK, LLR and ANN

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Models	2	674.3	674.3	337.2	3.3	0.044
Error	57	5817.8	5817.8	102.1		
Total	59	6492.1				

again. The related results are in Table 6.10 where the *P-value* indicates that the significance of ARE differences from the rest of three metamodels are reduced, which confirms our conclusions from the graphical trend change that the estimation results from TK, LLR and ANN are close.

6.4 Sensitivity Analysis

After the fitted Kriging and regression metamodels were chosen, this section will use them to conduct sensitivity analysis on cost factors. The sensitivity analysis is based on the analytical (differentiable) method developed in the fourth chapter. Since the function relationship in the ANN is completely hidden, the sensitivity analysis based on the ANN becomes difficult and impractical, and therefore is not investigated.

6.4.1   Kriging Metamodel

For the investigated case, the best fitted Kriging model is TK of order 3 with the nugget covariance function. For a Kriging model with a nugget covariance function, C_i is equal to 0. The derivative of $Z(\mathbf{X})$ with respect to x_j becomes

$$\frac{\partial}{\partial x_j}\hat{Z}(\mathbf{X}) = \begin{bmatrix} \frac{\partial}{\partial x_j} f_1(\mathbf{X}) \\ \vdots \\ \frac{\partial}{\partial x_j} f_M(\mathbf{X}) \end{bmatrix} \tag{6.7}$$

Table 6.11 is designed to calculate the derivative of TK with respect to x_j. In this table, MC represents model coefficient. In Table 6.11,

$DBF(x_i)^*$: represents the derivative of base functions with respect to x_i;

MC_i: represents the model coefficient of the ith base function in the Kriging model;

$\mathbf{X}_0 = (x_{01}, x_{02}, x_{03})$: represents the center point of Taylor expansion.

Since the case has 20 samples, there are 20 Taylor Kriging models, that is, a Taylor Kriging model per sample. For each model, the partial derivatives with respect to three independent variables (Ht, Dr and Ts) are calculated, respectively. The average of the partial derivatives of 20 Taylor Kriging models with respect to an independent variable is used to describe the sensitivity of cost function to the variable. Since the data are normalized by use of Eq. (6.1), the sensitivity of the actual cost (AC) to an independent variable (x_i) has to be computed by using the following formula:

$$\frac{\partial AC}{\partial x_i} = \frac{\partial AC}{\partial AC'} \times \frac{\partial AC'}{\partial x_i'} \times \frac{\partial x_i'}{\partial x_i} = \frac{max(AC) - min(AC)}{max(x_i) - min(x_i)} \times \frac{\partial AC'}{\partial x_i'} \tag{6.8}$$

Table 6.11: Sensitivity Analysis of TK with Nugget Covariance Function and Order 3

MC	Base Function	DBF(x_1)	DBF (x_2)	DBF (x_3)
MC_1	C_1	0	0	0
MC_2	C_2	0	0	0
MC_3	C_3	0	0	0
MC_4	C_4	0	0	0
MC_5	C_5	0	0	0
MC_6	C_6	0	0	0
MC_7	C_7	0	0	0
MC_8	C_8	0	0	0
MC_9	C_9	0	0	0
MC_{10}	C_{10}	0	0	0
MC_{11}	C_{11}	0	0	0
MC_{12}	C_{12}	0	0	0
MC_{13}	C_{13}	0	0	0
MC_{14}	C_{14}	0	0	0
MC_{15}	C_{15}	0	0	0
MC_{16}	C_{16}	0	0	0
MC_{17}	C_{17}	0	0	0
MC_{18}	C_{18}	0	0	0
MC_{19}	C_{19}	0	0	0
MC_{20}	1	0	0	0
MC_{21}	$x_3 - x_{03}$	0	0	MC_{21}
MC_{22}	$x_2 - x_{02}$	0	MC_{22}	0
MC_{23}	$x_1 - x_{01}$	MC_{23}	0	0
MC_{24}	$(x_3 - x_{03})^2$	0	0	$2MC_{24}(x_3 - x_{03})$
MC_{25}	$(x_2 - x_{02})(x_3 - x_{03})$	0	$MC_{25}(x_3 - x_{03})$	$MC_{25}(x_2 - x_{02})$
MC_{26}	$(x_2 - x_{02})^2$	0	$2MC_{26}(x_2 - x_{02})$	0
MC_{27}	$(x_1 - x_{01})(x_3 - x_{03})$	$MC_{27}(x_3 - x_{03})$	0	$MC_{27}(x_1 - x_{01})$
MC_{28}	$(x_1 - x_{01})(x_2 - x_{02})$	$MC_{28}(x_2 - x_{02})$	$MC_{28}(x_1 - x_{01})$	0
MC_{29}	$(x_1 - x_{01})^2$	$2MC_{29}(x_1 - x_{01})$	0	0
Sum				

The results based on Eq. (6.8) and Table 6.11 are shown in Table 6.12 where it is noted that the cost function is extremely sensitive to the vessel thickness but insensitive to vessel height.

6.4.2 Regression Metamodel

Similarly, for the MLR and LLR, the method of partial derivative is used to analyze the sensitivity of the product cost to the different cost factors. Suppose a MLR is:

$$F(x_1, \cdots, x_D) = \beta_0 + \beta_1 x_1 + \cdots + \beta_D x_D + \varepsilon_i \tag{6.9}$$

Eq. (6.10) gives its one-order partial derivative with respect to variable x_i.

$$\frac{\partial}{\partial x_i} F(x_1, \cdots, x_D) = \beta_i \tag{6.10}$$

It can be seen that the coefficient β_i can be used to describe the sensitivity of $F(x_1, \cdots, x_D)$ to x_i. Note that the investigated case has 20 MLR models. For each cost factor, the average of the corresponding cost factor coefficients in the 20 MLR models is used to express the sensitivity of the product cost to this factor. The obtained results are in Table 6.13 which shows that in the MLR model the product cost is most sensitive to the vessel thickness and insensitive to vessel height.

For the LLR model, the partial derivatives of the product cost with respect to the Height, Diameter and Thickness are calculated according to Eq. (6.11, 6.12, 6.13), respectively. The numerical results are in Table 6.13. These results are similar to those of TK and MLR. However, if the means in three tables are scrutinized, it can be found that the

Table 6.12: Sensitivity Analysis for TK of Order 3 with Nugget Covariance Function

Number	$DBF(x_1)$	$DBF(x_2)$	$DBF(x_3)$
1	0.3092	-0.1245	-0.0054
2	0.4271	-0.2009	0.2100
3	0.3810	-0.2033	0.3286
4	0.2515	0.0962	0.0988
5	0.2464	0.2745	-0.0881
6	0.0657	0.1826	0.3808
7	0.2882	0.3641	0.3117
8	0.6312	0.4995	-0.2220
9	0.6044	0.4083	-0.1075
10	0.1480	0.4525	0.2280
11	0.4179	0.5257	0.0934
12	0.5515	0.4152	-0.0719
13	0.2555	0.4305	0.1704
14	0.6413	0.5944	-0.0986
15	0.4823	0.4823	0.0516
16	0.5215	0.3163	0.1278
17	0.2630	0.6594	0.1377
18	0.8838	0.6622	-0.3355
19	1.6363	0.3873	-0.8832
20	0.6508	0.6657	3.0254
Mean (Normalization)	0.4828	0.3444	0.1676
Mean (Actual)	2.1164	18.7984	2022.2842

Table 6.13: Sensitivity Analysis for the MLR Metamodel

Regression Model	Intercept	Height Coefficient	Diameter Coefficient	Thickness Coefficient
1	-124528.99	1.79	33.00	5659.62
2	-118154.40	1.18	32.30	6436.68
3	-119383.71	1.12	32.21	6645.07
4	-118932.45	1.46	32.74	5903.32
5	-119364.47	1.35	32.93	6085.58
6	-117036.19	1.40	32.49	5932.13
7	-114592.89	1.50	31.37	5852.63
8	-115166.67	1.32	33.66	5827.14
9	-115323.73	1.41	32.51	5842.40
10	-117111.85	1.39	32.46	5956.96
11	-114615.13	1.50	31.84	5760.76
12	-119687.64	1.41	31.66	6177.10
13	-116367.14	1.42	32.21	5914.37
14	-114589.90	1.48	32.54	5702.88
15	-115651.46	1.44	32.18	5856.95
16	-116674.42	1.41	32.11	5993.31
17	-117596.26	1.36	32.27	6057.62
18	-116256.68	1.42	32.33	5891.16
19	-107666.81	1.48	27.04	5845.56
20	-82059.01	1.36	32.55	3342.52
Average	-115037.99	1.41	32.12	5834.19

Table 6.14: Sensitivity Analysis for the LLR Metamodel

Sample	$\frac{\partial AC}{\partial Ht}$	$\frac{\partial AC}{\partial Dr}$	$\frac{\partial AC}{\partial Ts}$
1	5.3188	9.1275	391.6208
2	1.5709	10.7099	835.2653
3	1.4941	14.0267	1051.4364
4	0.8507	17.6891	970.5041
5	0.5707	23.9846	1149.9118
6	0.5733	32.3473	1284.9085
7	0.7021	24.6664	1734.1056
8	1.6905	14.9382	1843.7481
9	1.1930	18.2511	1876.3872
10	0.8140	34.8059	2111.3825
11	0.8359	26.7422	2184.9641
12	1.6124	20.0371	2371.7608
13	0.9298	35.3284	2366.1904
14	1.0931	22.4540	2443.8109
15	1.0758	26.3991	2521.6337
16	1.1846	30.2508	2751.0212
17	1.1605	31.8493	3417.1017
18	1.6967	23.9886	3613.6524
19	2.8197	33.8747	8379.5383
20	1.9866	71.1230	6111.5578
Mean	1.4587	26.1297	2470.5251

means from TK and LLR are closer, which probably implies that TK and LLR are better

in this case.

$$\frac{\partial AC}{\partial Ht} = \frac{\partial AC}{\partial Ln(AC)} \frac{\partial Ln(AC)}{\partial Ln(Ht)} \frac{\partial Ln(Ht)}{\partial Ht} = \eta_1 \frac{AC}{Ht} \qquad (6.11)$$

$$\frac{\partial AC}{\partial Dr} = \frac{\partial AC}{\partial Ln(AC)} \frac{\partial Ln(AC)}{\partial Ln(Dr)} \frac{\partial Ln(Dr)}{\partial Dr} = \eta_2 \frac{AC}{Dr} \qquad (6.12)$$

$$\frac{\partial AC}{\partial Ts} = \frac{\partial AC}{\partial Ln(AC)} \frac{\partial Ln(AC)}{\partial Ts} = \eta_3 AC \qquad (6.13)$$

103

6.5 Property Analysis

6.5.1 Applicability of Models

Kriging, RA and ANN all depend on reliable driver data from simulation models. Any of them does not mitigate the difficulties associated with collecting simulation data. Kriging requires that a Kriging coefficient matrix be nonsingular. The singularity requirement makes Kriging applications limited. RA faces the problems related to multicollinearity and heteroscedasticity. ANN can readily accommodate multicollinearity and heteroscedasticity. Three methods can encounter potential specification errors. However, the specification errors from ANN are most difficult to be recognized.

It is expected that Kriging will be considered a viable alternative to regression if the underlying data relationship from a simulation model has significant spatial correlation and nonlinearities. If the data behavior is functional discontinuities and significant nonlinearities and has large independent variable dimensionality, ANN will have better potential. However, the applicability concerns of ANN, such as its network architecture, training methods and stopping criteria, should not be ignored. These concerns represent formidable hurdles to widespread use and acceptance of ANN. Moreover, the selection of ANN structure and parametric settings more depends on the experience with the less supportive theories [134]. However, the parametric settings of RA and Kriging are easier.

6.5.2 Accuracy of Interpolation

What is interesting in Table 6.8 is that the average value of twenty AREs of the chosen TK metamodel is much less than that of the MLR metamodel, less than that of the LLR metamodel, greater than that of the ANN metamodel, but close to those of the LLR and

ANN metamodel. These results may reflect the simulation interpolation or estimation capabilities of Kriging. When the function relationship between dependent and independent variables in a simulation model is unknown, ANN can be a good method to capture the function relationship, and its simulation interpolation accuracy is better. If a suitable regression metamodel can be established, the simulation interpolation accuracy from regression can be acceptable. For example, for the investigated case, the LLR metamodel is acceptable. However, the diversities of regression metamodeling make it difficult to select a good regression metamodel. The drift function of Kriging has the features of regression. However, Kriging further has the capabilities to capture the spatial correlation of observation values, which makes it have stronger interpolation potentials.

6.5.3 Feasibility of Sensitivity Analysis

Although RA, especially MLR, has more difficulties in capturing a function relationship among simulation data because of the limitation of model structures, regression metamodeling can give a specific and simple function form, which can benefit sensitivity analysis of a simulation model. For ANN, the fitted function form for a simulation model is completely hidden and invisible and resembles a "black box". Essentially, it will be a complicated nonlinear function. Because of the function invisibility, sensitivity analysis on simulation input variables is difficult. In comparing with RA and ANN, Kriging can provide an explicit function as specific as that given by regression metamodeling, and sensitivity analysis based on a Kriging metamodel is feasible. However, because Kriging must consider the spatial correlation of observation values, its metamodel form and thus its sensitivity analysis process are usually more complicated.

6.6 Conclusions

The chapter empirically investigated Kriging metamodeling for sensitivity analysis by using a physical simulation case. In order to obtain an effective metamodel, the interpolation accuracy of different Kriging metamodels, regression metamodels and ANN was first compared. Among different Kriging metamodels, a TK metamodel provides the best interpolation results which are consistent with the empirical results from deterministic and random simulations in the previous chapter. While compared with regression metamodels (MLR and LLR) and ANN, the chosen Kriging metamodel works better than regression metamodels but worse than ANN. Regression metamodeling are effectively validated by considering multicollinearity, heteroscedasticity, and specification errors. Therefore, the comparison is credible.

Chapter 7

Simulation Optimization Based On Kriging And Evolutionary Algorithms

7.1 Introduction

Simulation can be used to model real complex systems. These simulation models can be used to calculate approximate outputs of real systems, perform sensitivity analysis on inputs, and search for optimal inputs. Due to the complexity of systems, simulation models are often computer expensive, which makes the above three operations difficult. In the previous chapters, Kriging was used to build metamodels for inputs and outputs of simulation models. These metamodels were utilized to perform the interpolation operation to reduce the computational expense of running simulation models. Furthermore, Kriging metamodels were used to assist sensitivity analysis on simulation models. The objective of this chapter is to discuss how to optimize a simulation model. A novel simulation optimization algorithm based on Kriging and Evolutionary Algorithm (EA) will be developed.

The chapter is organized as follows. First, EAs are briefly introduced, and then a novel simulation optimization algorithm is created which integrates Kriging with EAs. Second, the properties and parameter settings of the algorithm are analyzed, and some computational cases are presented to investigate the effectiveness of the algorithm. Finally, summary and conclusions are given.

7.2 A Novel Simulation Optimization Algorithm

7.2.1 Evolutionary Algorithms

The typical evolutionary algorithms (EAs) include Evolutionary Programming [41], Genetic Algorithms (GA) [55], Evolutionary Strategy (ES)[117, 130], Genetic Evolution of Data Structures [102], Genetic Evolution of Programs [80], Tabu Search [46, 47], Artificial Immune Systems [54], Cultural Algorithms [121], DNA Computing [1], Ant Colony Optimization [32], Particle Swarm Optimization (PSO)[69], Memetic Algorithms [48], Estimation of Distribution Algorithms [84], etc.. For the specific introductions to these algorithms, a reader can refer to the related references. Figure 7.1 gives the basic optimization process of these algorithms. EAs use fitness functions to calculate fitness values of candidate solutions, and use these values as a standard to evaluate the qualities of solutions. According to fitness values, EAs perform evolutionary operations such as crossover and mutation to generate some new evolved solutions. The evolved solutions have to be evaluated by fitness functions. Then, part of the evolved solutions and current solutions are chosen to become the next generation solutions. The algorithms continue this optimization loop until a stopping criterion is satisfied.

7.2.2 Development of Simulation Optimization Algorithm (SOAKEA)

EAs are an effective simulation optimization tool and have been used to optimize simulation models. However, EAs frequently need to evaluate fitness functions. A simulation evaluation means to run a simulation model only once. When running a simulation model is computer expensive, EAs for simulation optimization will encounter essential difficulties because of computer limitations such as memory capacity and processor speed. The dash lines

108

in Figure 7.1 show that evolutionary operations in EAs will be stopped due to insufficient fitness values.

One way to solve this problem is to adopt known observation values of a simulation model to fit a surrogate fitness function which is computationally inexpensive, and to use it to evaluate inputs. Thus, simulation evaluations can be reduced. Kriging is an effective interpolation tool for a simulation model, which makes it feasible to use a fitted Kriging metamodel to replace its corresponding simulation model to evaluate inputs. With the assistance of Kriging metamodels, EAs can continuously perform evolution optimization to search for optimal solutions of simulation inputs.

With this idea in mind, the chapter integrates Kriging with EAs to develop a novel simulation optimization algorithm based on Kriging and EAs where Kriging serves as a surrogate fitness function. The new algorithm is named Simulation Optimization Algorithm based on Kriging and EA (SOAKEA). Figure 7.2 is its flow diagram. The specific details of SOAKEA are in Table 7.1.

7.2.3 Analysis on SOAKEA

Fitting an initial Kriging model: In the initialization process, fitting a Kriging model is important. A method of computer experimental design such as Latin Hypercube Sampling (LHS) can be used to generate $3E_0/4$ simulation inputs (individuals). Running a simulation model can provide their corresponding outputs. For EAs, these outputs are actual fitness values. The inputs and outputs are used to fit different Kriging metamodels. To select a best fitted metamodel among them, a validation process is needed. $E_0/4$ simulation inputs are randomly sampled from a uniform distribution of the feasible domain, and used to
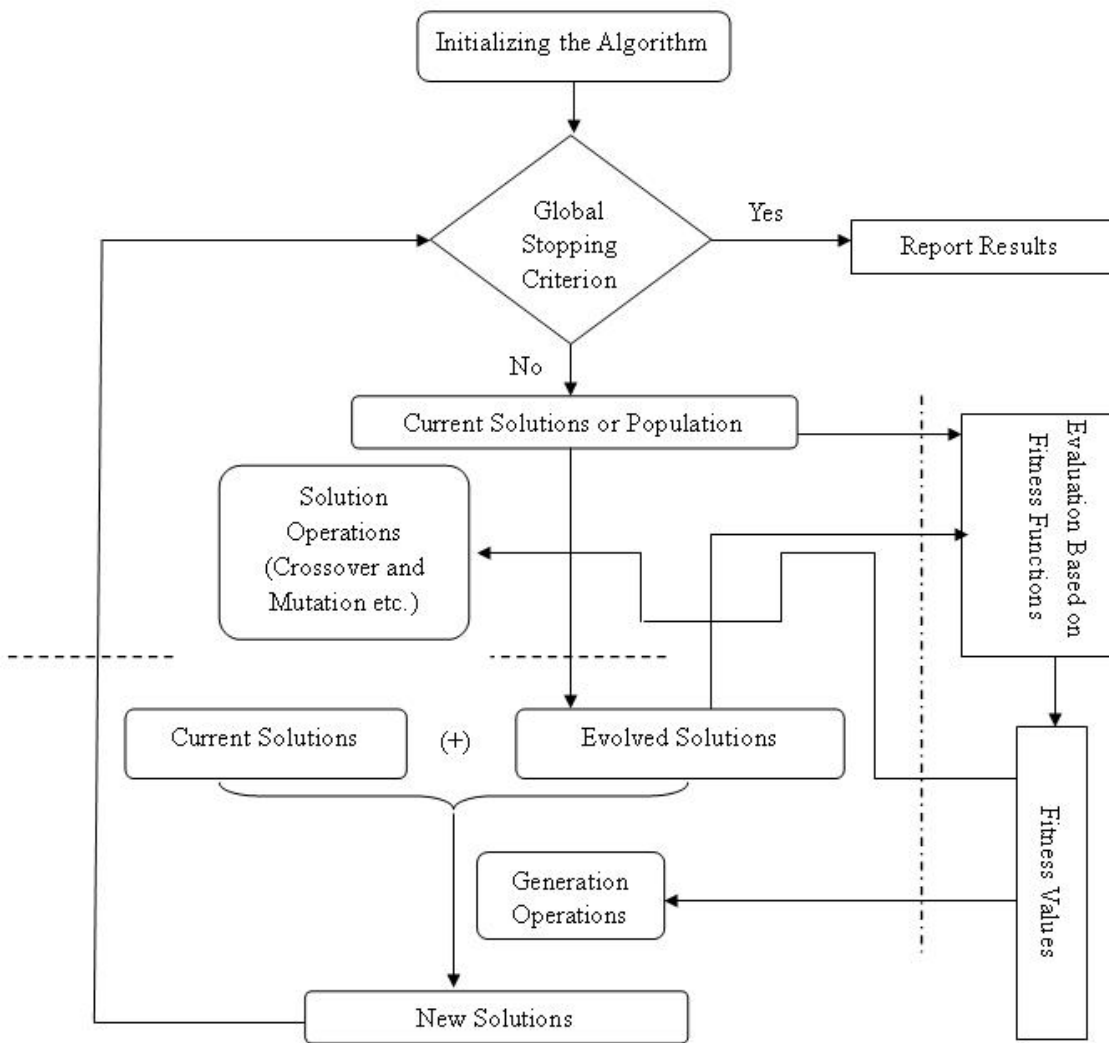
109

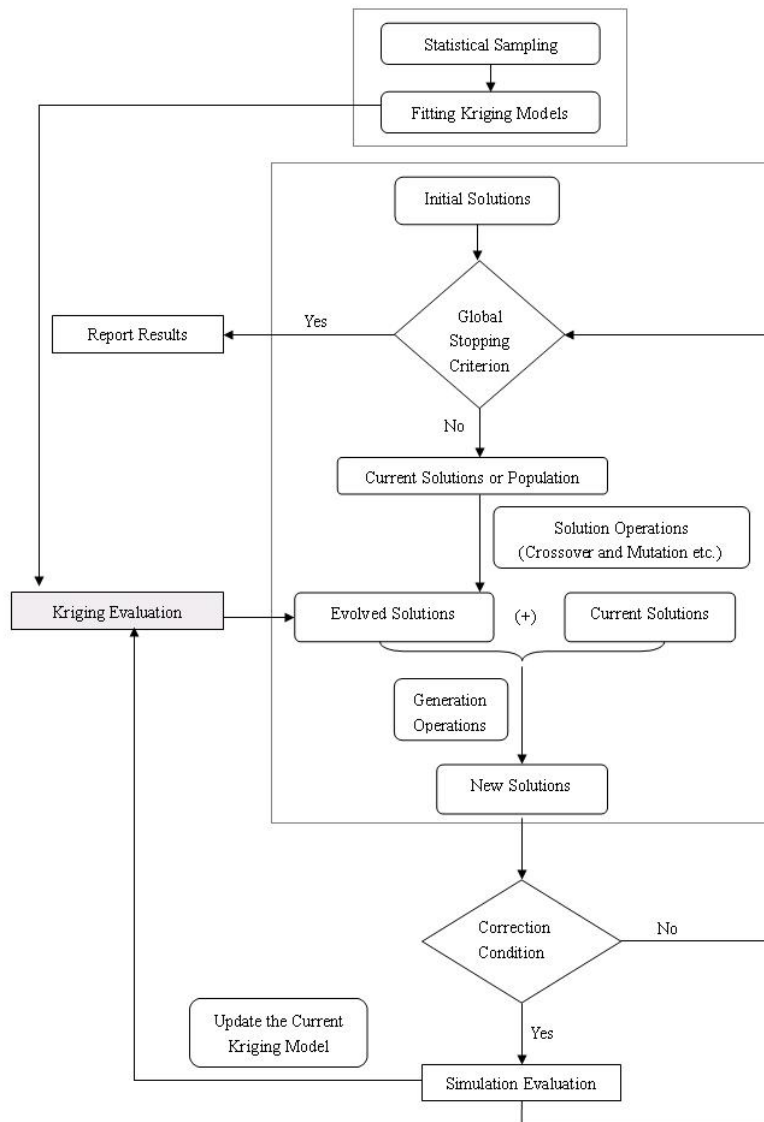Figure 7.1: Operations of Evolutionary Algorithms

Figure 7.2: The Flow Diagram of SOAKEA

Table 7.1: Simulation Optimization Algorithm Based on Kriging and EA

BEGIN

Initialize:

(1)   Use a method of computer experimental design to generate E_0 simulation inputs (individuals) in the feasible region of independent variables, and run a corresponding simulation model to obtain their outputs called fitness values for an EA;

(2)   Name the E_0 inputs and outputs an observation set;

(3)   Use the E_0 inputs and outputs to fit different Kriging metamodels according to a cross-validation method and set the best one to be a current Kriging metamodel;

(4)   Record the input with the best output in the observation set, and name it *bestIndividual*;

(5)   Randomly generate an initial population of size M_0 for a chosen evolutionary algorithm and name it the current population;

(6)   Use the current Kriging metamodel to evaluate the individuals in the current population;

While (Global stopping criterion) {

(1)   Perform evolutionary operations to generate some evolved individuals;

(2)   Use the current Kriging metamodel to evaluate them;

(3)   Use generation operations to generate a new population from the evolved individuals and the individuals in the current population;

(4)   Name the new population the current population;

(5)   If (corrective condition) {

(5.1) Correct the fitness values of m_0 individuals in the current population by running the simulation model to evaluate them;

(5.2) Add the simulation inputs and corrected outputs to the observation set;

(5.3) Update *bestIndividual*

(5.4) Fit a new Kriging model by using the individuals in the observation set and name it the current Kriging model;

}

}

End   While

Report Results:

(1)   Evaluate the optimal individual in the current population by running the simulation model;

(2)   Add the optimal individual to the observation set, and update *bestIndividual*;
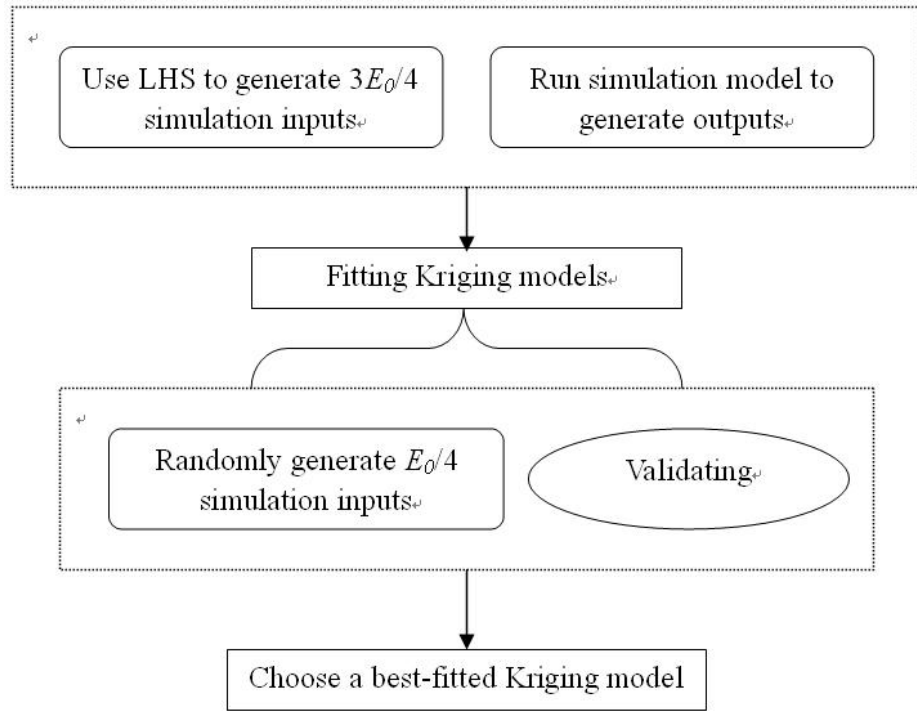
(3)   Report the results.

END

Figure 7.3: Flow Diagram of Fitting an Initial Kriging Model

test the fitted Kriging metamodels. The random sampling from a uniform distribution can better show the robustness of chosen metamodel. Figure 7.3 gives the flow diagram to fit an initial Kriging model.

Optimization capability of SOAKEA: SOAKEA uses an EA to search for optimal simulation inputs. In the EA, a simulation model is essentially an implicit fitness function. Because the prohibitive computational expense of running the simulation model blocks the

EA optimization operations, a fitted Kriging model temporarily replaces the simulation model to assist in the EA. The fitted Kriging model serves as a surrogate fitness function. If each simulation input is allowed to be evaluated by the simulation model, SOAKEA becomes the EA itself. Therefore, the upper bound of optimization capability of SOAKEA is the optimization capability of EA, which makes it necessary to choose an EA with strong optimization capability.

Corrective operation: Since Kriging models in SOAKEA are surrogate fitness functions, fitness values from Kriging will have some errors. Thus, a corrective operation is needed in order to avoid errors misguiding evolutionary search. In the corrective operation, the corrective condition can be a given number of function evaluations or generations or an improvement rate, etc.. A trial-and-error method can be applied in order to find a suitable value. Frequent corrections can increase computational expenses caused by running a simulation model. However, insufficient corrections can cause error accumulation. Achieving a good tradeoff between these two aspects is a standard to choose a corrective procedure.

Potential corrective operations can be the corrections of partial and complete fitness values in a generation. Suppose the number of corrected fitness values in a generation is m_0. If $m_0 < M_0$ (see Table 7.1), the operation is a partial correction. Otherwise, it is a complete correction. A disadvantage of correcting a complete generation is that its computational cost is expensive. However, a partial correction will result in individuals with uncorrected outputs continuing to misguide evolutionary search.

Correction operations make fitting Kriging models become a dynamic process. When SOAKEA corrects fitness values by running a simulation model, more actual outputs can be obtained. With more actual outputs, SOAKEA will update a current Kriging model,

resulting in a Kriging model which can better approximate its simulation model. If a correction can only provide a small number of simulation outputs, the update may be delayed to accumulate actual outputs. Thus, the expense related to fitting Kriging models can be reduced. Usually, an update adopts the same Kriging structure and only changes its coefficients.

As optimization and corrective operations continue, a current Kriging model will be frequently updated and become more and more accurate. Thus, a possible improved correction strategy is that it has a high correction frequency in the initial optimization stage and then little by little decreases the frequency as the optimization operation of SOAKEA continues. In the corrective operation, the distance between two points may be considered. If the points needed to be corrected are close to those already simulated, it may not be needed to correct them or use them to update a current Kriging model because they can not add much information. However, a disadvantage is that frequently calculating and comparing distances of points will tremendously increase the computational complexity of SOAKEA.

Stopping criterion: In SOAKEA, selecting a global stopping criterion deserves discussion. Most references of EAs use a certain number of function evaluations or generations to terminate optimization search. A drawback is that the number of function evaluations necessary for convergence is unknown. Thus, a trial-and-error method may be needed to search for a suitable number. Zielinski *et al.* [162, 163] investigate some other stopping criteria that include reference criteria, exhaustion-based criteria, improvement-based criteria, movement-based criteria, distribution-based criteria, and combined criteria. For the detailed introductions to these criteria, a reader can refer to [162, 163].

In summary, the idea behind SOAKEA is to integrate the precise interpolation features of Kriging with the optimization capabilities of an EA to search for optimal simulation inputs according to limited simulation observations. If SOAKEA, under a tolerated error level, can efficiently find optimal inputs, it can be concluded that it is effective for the optimization of a simulation model whose run-time is computationally expensive.

7.3   Computational Experience

Several simulation cases based on benchmark functions will be used to empirically investigate the effectiveness of SOAKEA. Particle Swarm Optimization (PSO) is the chosen evolutionary algorithm used in SOAKEA, and will be introduced in the next section. Then the settings of some important parameters in SOAKEA will be empirically analyzed. These parameters include the size of initial sample used to fit an initial Kriging model, the correction strategy, and the number of corrected points used to update a current Kriging model.

7.3.1   Particle Swarm Optimization (PSO)

PSO is a population-based stochastic optimization method, and was developed by Eberhart and Kennedy [69]. PSO uses a fitness function to calculate fitness values of particles. According to fitness values, PSO performs evolution operations to update position vectors. PSO continues this loop until stopping criteria are satisfied. In PSO, each particle has several common attributes: $xVector$, $pVector$, $vVector$, $xFitness$ and $pFitness$. $xVector$ records the current position of a particle. If the investigated space is D dimensions, $xVector$ can be denoted by $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$. $pVector$ is the best position found so far by

a particle and can be represented by $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$. $vVector$ is a velocity vector, represented as $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$. $xFitness$ and $pFitness$ record the fitness values of $xVector$ and $pVector$, respectively.

The evolutionary operation of each particle mainly includes two steps which can be described as follows ([69]):

1. Step 1: The algorithm finds particle g with the best $pFitness$ in the neighborhood of particle i, and then combines the $pVector$ of particle g and the $pVector$ of particle i to generate a new $vVector$ for particle i. Meanwhile, the algorithm introduces the random disturbance $rand()$ to increase the search pressure. The following is the formula that updates the dth dimension of $vVector$:

$$v_{id} = v_{id} + \varphi_1 * rand() * (p_{id} - x_{id}) + \varphi_2 * rand() * (p_{gd} - x_{id}) \qquad (7.1)$$

2. Step 2: The algorithm uses the updated $vVector$ to obtain a new $xVector$ for particle i. The related formula is:

$$x_{id} = x_{id} + v_{id} \qquad (7.2)$$

In the above formulae, the two constants φ_1 and φ_2 are the cognition and social learning components, respectively. In 1999, Maurice Clerc introduced a constriction factor K to control the growth speed of $vVector$, and it is calculated as follows:

$$K = \frac{2}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|} \qquad (7.3)$$

117

where $\varphi = \varphi_1 + \varphi_2$ and $\varphi > 4$. The update operation of the dth dimension of $vVector$ then becomes

$$v_{id} = K[v_{id} + \varphi_1 * rand() * (p_{id} - x_{id}) + \varphi_2 * rand() * (p_{gd} - x_{id})] \qquad (7.4)$$

The typical structures of PSO include ring and star topologies.

7.3.2 The Determination of Initial Sample Size

This section will discuss how to select a size of initial samples used to fit an initial Kriging model. Too large a size of samples will increase the computational expenses of running a simulation model. An insufficient sample size will result in an inaccurate fitted Kriging interpolation model. One simulated case based on a benchmark function is used to empirically investigate the selection problem of sample sizes. The chosen benchmark function is SHC. The different scenarios of sample sizes investigated are in Table 7.2. The table shows that the numbers of fitting and testing samples are about 3/4 and 1/4 of total initial samples, respectively. For these scenarios, the best fitted Kriging models are shown in Table 7.3. Table 7.4 uses Scenario 4 to illustrate how to use AAREs as the performance standard to choose a best fitted Kriging model.

Table 7.3 indicates that different scenarios of sample sizes result in different best-fitted Kriging models. However, in the four scenarios, TK is always chosen. As the sample size increases, a cubic (2) covariance function is further considered to be adequate. Four scenarios show that in the chosen TK models, the orders of Taylor expansion are greater than or equal to 3, and the influence distances of covariance are 1 or 5 standard deviations. AAREs

Table 7.2: Sizes of Initial Samples

Fitting Samples	Testing Samples
15	5
60	20
100	30
150	50

Table 7.3: The Chosen Kriging Models and Their AAREs under Different Scenarios

Scenario	Size of Samples		Kriging Models and Parameters				AARE
	Fitting	Validating	Model	Covariance Function	Order	SD	(%)
1	15	5	TK	Gaussian	4	1	28.940
2	60	20	TK	Cubic (2)	3	5	13.906
3	100	30	TK	Cubic (2)	3	1	12.839
4	150	50	TK	Cubic (2)	4	1	3.473

Table 7.4: AAREs of Different Kriging Models (The Size of Initial Samples=200)(%)

Covariance Function	SD	SK	OK	Order of Taylor Kriging			
				1	2	3	4
Nugget	0	100.0000	78.9620	78.9620	80.4530	69.6518	764.9876
Linear	1	29.3027	51.1483	51.1483	51.2805	41.9032	35.1451
	2	29.3737	51.1483	51.1483	51.2805	41.9032	35.1451
	3	29.3946	51.1483	51.1483	51.2805	41.9032	35.1451
	4	29.4047	51.1483	51.1483	51.2805	41.9032	35.1451
	5	29.4105	51.1483	51.1483	51.2805	41.9032	35.1451
Cubic (2)	1	3.7112	4.0921	4.0921	4.0936	4.0055	3.4727
	2	3.6456	4.1018	4.1018	4.1022	4.0112	3.4778
	3	3.6213	4.1035	4.1035	4.1038	4.0122	3.4786
	4	3.6087	4.1040	4.1040	4.1045	4.0124	3.4791
	5	3.6015	4.1045	4.1045	4.1042	4.0132	3.4790
Spherical	1	29.3269	51.0731	51.0731	51.2141	42.0161	35.0955
	2	29.3637	51.1230	51.1230	51.2571	41.9311	35.1327
	3	29.3838	51.1365	51.1365	51.2696	41.9155	35.1396
	4	29.3951	51.1415	51.1415	51.2743	41.9101	35.1420
	5	29.4022	51.1440	51.1440	51.2765	41.9076	35.1431
Exponential	1	29.3259	51.1874	51.1874	51.3364	41.8479	35.1699
	2	29.3796	51.1526	51.1526	51.2935	41.8892	35.1513
	3	29.3974	51.1482	51.1482	51.2863	41.8970	35.1478
	4	29.4063	51.1470	51.1470	51.2837	41.8997	35.1467
	5	29.4116	51.1468	51.1468	51.2826	41.9009	35.1461
Minimum=				3.4727			

state that Scenario 4 gives the best fitted Kriging model. When an initial sample size increases, more information of a simulation model can be obtained, and a corresponding fitted Kriging model will become more accurate. With a more accurate Kriging model, corrective operations in later optimization can be reduced, which can decrease the computational expenses related to simulation correction. However, a large number of initial observations will tremendously increase the initial computational expense of simulation sampling, which should be avoided. When two important parameters of a Kriging model, such as model type and covariance function, can be determined, their corresponding initial sample size can be considered to be sufficient. For TK, the order of Taylor expansion also need be considered. In the initialization process, the main goal is to find an effective Kriging model structure. The specific model coefficients can be improved in later optimization. Based on this consideration, 80 may be a good initial sample size.

7.3.3 Corrective Operations and Update of Kriging Models

Corrective operations in SOAKEA are a crucial component. Two important problems related to such operations need be considered. First, when should a corrective operation be conducted, or what correction strategy is adopted? Second, after a corrective operation, how are corrected points used to update a current Kriging model? One simulated case based on the benchmark function SHC will be used to empirically analyze these two problems.

In SOAKEA, PSO is used to perform optimization operations. The population size of PSO is set at 10, and the cognition and social learning rates are both set at 2.05, consistent with [38, 69, 162]. The global stopping criterion is the number of Kriging function evaluations, and it is set at 1000. The corrective condition is the number of updated

generations. For every G_0 updated generations, a corrective operation will be conducted. A correction per G_0 generations is called a correction strategy. In order to investigate the first problem, different correction strategies are tested. These strategies are that G_0 are set to 10, 20, 40, and 80, respectively. G_0 equal to 10 implies that for every 10 generations or 100 (10×10) Kriging function evaluations, a correction will be conducted. A correction means that $xVector$ and $vVector$ of all 10 particles in a current population will be corrected. If an $xVector$ or $vVector$ is the same as one of the vectors which were already corrected, its simulation output can be directly obtained and running a simulation model is not needed. After one correction, all of the corrected vectors are added to an observation set and utilized to update a current Kriging model. The size of initial samples is set to 80. The algorithm is run 30 times for each correction strategy.

Table 7.5 provides the test results. In this table, the total number of simulation evaluations is decreased to 96 from 220 while G_0 is increased to 80 from 10. When a simulation evaluation is computationally expensive, the significant decrease of simulation evaluations has evident advantages for simulation optimization. In the table, Kriging points are defined as the known observation points used to fit a Kriging model, and the number is finally decreased to 95 from 219. Decreasing the number of Kriging points can effectively reduce the computational complexity of fitting a Kriging model. However, as G_0 is increased to 80 from 10, a disadvantage of SOAKEA is that the success rate of searching optimal solutions is decreased to 6.67% from 100%, which is unacceptable.

Although a strategy with frequent corrections, or G_0 equal to a small value, ensures that optimal solutions of a simulation model can be found more easily, frequent corrections will significantly increase the correction expenses caused by running a simulation model.

122

Table 7.5: The Test Results of Different Correction Strategies

Correction Strategy	Kriging Evaluation	Kriging Points	Simulation Evaluation	Success Rate (%)
10	0	80	80	
	110	80	80	
	210	99	99	
	310	109	109	
	410	123	123	
	510	137	137	100
	610	153	153	
	710	167	167	
	810	183	183	
	910	201	201	
	1000	219	220	
20	0	80	80	
	210	80	80	
	410	98	98	100
	610	113	113	
	810	129	129	
	1000	145	145	
40	0	80	80	
	410	80	80	83.33
	810	97	97	
	1000	111	112	
80	0	80	80	
	810	80	80	6.67
	1000	95	96	

However, if G_0 is too large, the corresponding correction strategy may not be able to provide enough corrections. It is then difficult for the algorithm to find optimal solutions. Figure 7.4 provides an intuitive explanation for this problem. Suppose the current position of a particle is Circle 1. The optimal solution is Circle 2. The current search direction of particle is shown in the figure. In the search process, assume there are three possible correction points which are $Corr_1$, $Corr_2$ and $Corr_3$. According to the distance of points, it can be concluded that point $Corr_1$ corresponds to a frequent correction. Point $Corr_3$ represents an insufficient correction. Figure 7.4 indicates that the correction in point $Corr_1$ may be unnecessary because the particle can move closer to the optimal solution. This correction will waste some simulation evaluations but may not be able to improve the search direction. However, point $Corr_3$ is far away from the current position of particle and the optimal solution. The correction at point $Corr_3$ will waste some Kriging evaluations but can not provide more useful information. The search direction can not be surely improved. An effective correction point may be $Corr_2$. Point $Corr_2$ is closer to the optimal solution. The simulation output in this point can provide more information of the optimal solution from the distance perspective. The search direction of particle can be more effectively improved because when the corrected output of point $Corr_2$ is added to the observation set, a new Kriging model can be fitted, and it can capture more information about the optimal solution or Circle 2.

In the initial stage of optimization operations, there are less known observation points and a fitted Kriging model can not provide precise interpolation values. In order to avoid errors misguiding evolutionary search, frequent corrections are needed. As fitted Kriging
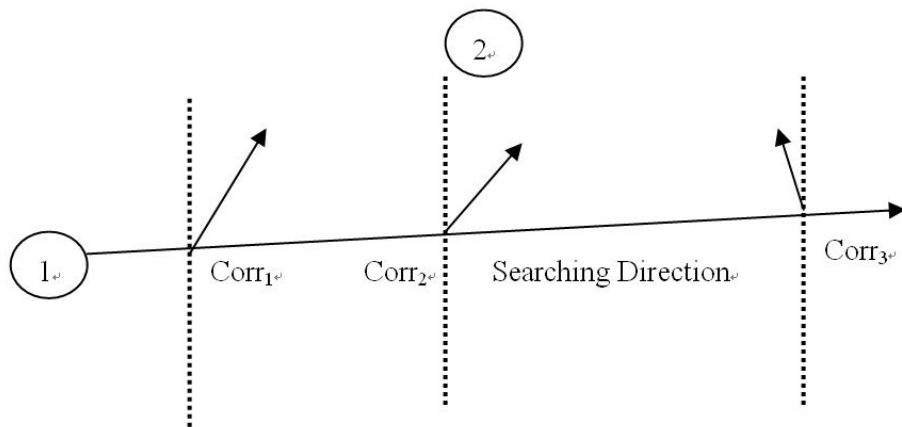
124

Figure 7.4: Diagrammatic Explanation of Correction Operations

Table 7.6: Numbers of Simulation Evaluations for Different Correction Strategies

		Correction Strategy in the Second Half Stage			
		10	20	40	80
Correction	10	220	173	156	156
Strategy	20		145	132	132
in the	40			112	112
First Half	80				96
Stage					

models become more accurate, Kriging evaluations can better approximate simulation evaluations. Thus, the correction frequency can be reduced. Based on this idea, a combination correction strategy is created. $[N_1, N_2]$ is used to express a combination correction strategy. It implies that in the first half of optimization stage, a correction per N_1 generations is conducted, and in the second half a correction per N_2 generations is used.

The effectiveness of combination correction strategies is empirically investigated. Tables 7.6 and 7.7 provide the results. The numbers of simulation evaluations and success rates for different strategies show that combination correction strategies have significant advantages. They can substantially reduce simulation evaluations while achieving comparable success rates. For example, if [10, 10] is replaced by the combination correction strategy [10, 80], the number of simulation evaluations is reduced to 156 from 220 but the 100% success rate is maintained.

Combination correction strategies can have different mutations. A combination strategy can consist of more than two strategies. These strategies could be combined during any stage of optimization process. In the above empirical research, only a simple half-half combination of two strategies is adopted.

Table 7.7: Success Rates of Optimization Search for Different Correction Strategies (%)

		Correction Strategy in the Second Half Stage			
Correction		10	20	40	80
Strategy	10	100	100	100	100
in the	20		100	100	100
First Half	40			83.33	83.33
Stage	80				6.67

After a corrective operation, more simulation outputs are available. These outputs could be used to update a current Kriging model. During the fitting process, as more simulation outputs are adopted, a more accurate fitted Kriging model can be generated. However, a negative influence is that adopting more simulation outputs will increase the computational complexity of fitting a Kriging model. How to balance the interpolation accuracy and computational complexity of Kriging will be investigated next. In the investigation, the correction strategy G_0 is set to 10. The parameter settings of PSO are the same as before. Two scenarios are investigated. Scenario 1 is that all simulation outputs in the observation set are used to update a current Kriging model; in scenario 2, as the number of observation points in the observation set increases to 100, the algorithm continues to use the current Kriging model fitted by the initial 100 simulation observations and stops updating it, although corrective operations will be continuously conducted and the number of points in the observation set will increase.

The empirical results are shown in Table 7.8. These results demonstrate that in the initial optimization stage, increasing Kriging points can significantly improve current optimal values. Kriging points imply the known simulation observation points in the observation set used to fit Kriging models. However, in the following optimization search, the increase

Table 7.8: Numbers of Simulation Evaluations Used in Different Correction Strategies

Kriging Evaluation	Scenario 1		Scenario 2		Increase of Rate Kriging Points (%)	Improvement Rate of Optimal Values (%)
	Kriging Points	Current Optimal Value	Kriging Points	Current Optimal Value		
0	80	-0.4739	80	-0.4739	0.0	0.000
110	80	-0.9850	80	-0.9850	0.0	0.000
210	99	-1.0306	99	-1.0306	0.0	0.000
310	100	-1.0306	109	-1.0315	9.0	0.093
410	100	-1.0310	123	-1.0316	23.0	0.058
510	100	-1.0314	137	-1.0316	37.0	0.024
610	100	-1.0315	153	-1.0316	53.0	0.015
710	100	-1.0316	167	-1.0316	67.0	0.007
810	100	-1.0316	183	-1.0316	83.0	0.000
910	100	-1.0316	201	-1.0316	101.0	0.000
1000	100	-1.0316	219	-1.0316	119.0	0.000

of Kriging points can not significantly improve current optimal values. For example, as Kriging points in Scenario 1 increase to 219 from 137, Kriging points in Scenario 2 are always 100 and the current optimal values in two scenarios are very close. It is evident that fitting a Kriging model in Scenario 1 has higher computational complexity and thus wastes some computational resources. Furthermore, the two scenarios can both eventually reach the optimal value of a test function. Figure 7.5 shows that the trend changes of current optimal values in the optimization search of the two scenarios are close.

7.3.4 The Empirical Investigation on SOAKEA

After some important parameter settings of SOAKEA are analyzed, this section uses four simulated cases based on benchmark functions to investigate its optimization capabilities. The four benchmark functions are SHC, B2, Hmb, and Ras whose dimension of
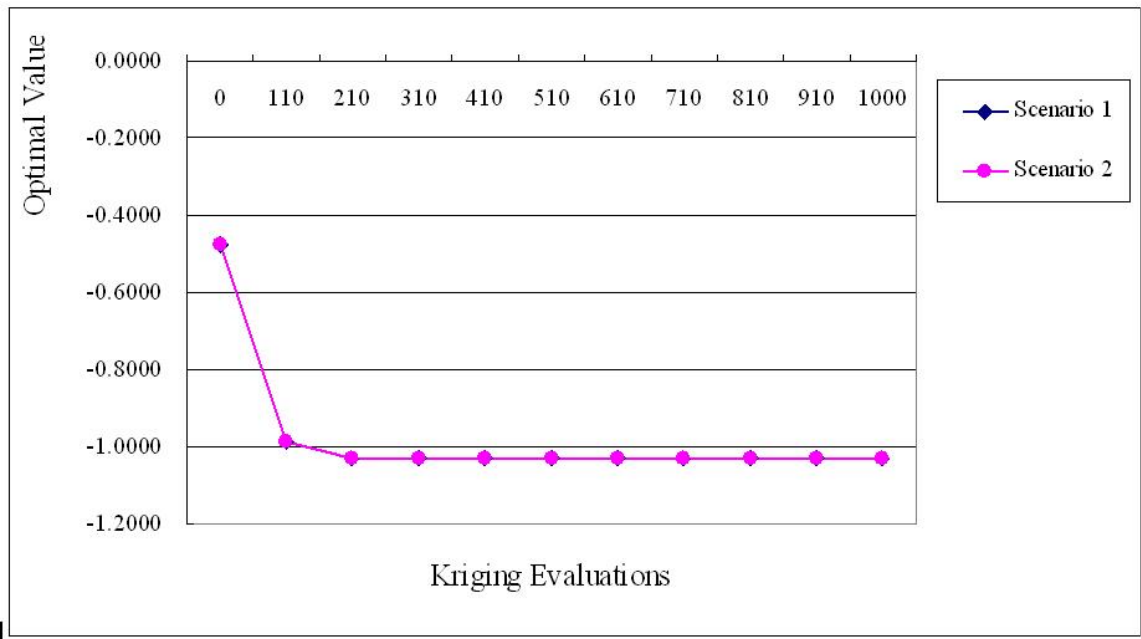
Figure 7.5: Optimal Result Comparison of Different Scenarios

Table 7.9: Chosen Kriging Models for Different Cases

Benchmark Function	Chosen Kriging Type	Covariance Function	Order	SD
SHC	TK	Cubic (2)	4	5
B2	TK	Nugget	3	0
Hmb	TK	Cubic (2)	3	4
Ras	TK	Exponential	1	1

independent variables is arbitrarily set to 5. The chosen evolutionary algorithm in SOAKEA is still PSO. The population size of PSO is set at 10, and the cognitional and social learning rates are both 2.05. The size of initial samples of each case is 80. Specifically, 60 simulation inputs are generated by LHS, and their outputs are obtained by running the corresponding simulation model. These inputs and outputs are used to fit different Kriging models. The rest of 20 simulation inputs are randomly sampled from a uniform distribution and utilized to validate the fitted Kriging models. The best fitted model is used as a surrogate fitness function in later optimization operations. According to this procedure, for each case, 10 different initial sample sets are generated and used to optimize a constructed simulation model to test the robustness of SOAKEA. For four test cases, the best fitted Kriging models are shown in Table 7.9. These models are used as surrogate fitness functions while PSO searches for optimal solutions. In order to further examine the robustness of SOAKEA, 10 different initial populations for each initial sample set are randomly generated for PSO, and thus SOAKEA is run 10 times for each initial sample set. Totally, the algorithm is run for 100 times for each case.

For the four test cases, the crucial parameter settings of SOAKEA are shown in Table 7.10. The parameter settings of test case based on the benchmark function SHC are used to

explain Table 7.10. The global optimization condition (stopping criterion) of SOAKEA is the total number of Kriging evaluations. A Kriging evaluation denotes that SOAKEA uses a fitted Kriging model to evaluate a solution only once. The global optimization condition is 1000 Kriging evaluations, which means that when the number of Kriging evaluations reaches 1000, SOAKEA will stop optimization search. Kriging points imply the known simulation observation points used to fit Kriging models. The column of Kriging points in Table 7.10 provides the upper bound of numbers of Kriging points. For instance, a number in the column being 100 implies that as the number of points in the observation set increases to and then exceed 100, the following optimization operation in SOAKEA will still use the Kriging model fitted by the initial 100 points. Based on the structure of dual Kriging model, it means that SOAKEA will not need to calculate inverse Kriging coefficient matrices later and the computational complexity caused by Kriging can be effectively controlled.

SOAKEA uses combination correction strategies to correct fitness values from Kriging in the four test cases. The general form of a combination correction strategy is defined as $[(N_{11}, N_{12}), (N_{21}, N_{22}), \cdots, (N_{m1}, N_{m2})]$. That is, the combination strategy is combined by m simple correction strategies. When the ratio of current and total numbers of Kriging evaluations is less than $\frac{N_{11}+\cdots+N_{i1}}{N_{11}+\cdots+N_{m1}}$ and greater than or equal to $\frac{N_{11}+\cdots+N_{i-1,1}}{N_{11}+\cdots+N_{m1}}$, the current correction strategy is one correction per N_{i2} generations. The ratio $(N_{11} : N_{21} : \cdots : N_{m1})$ determines when SOAKEA needs to adjust a current correction strategy, and therefore is defined as the correction adjustment ratio. For example, in the first test case, SOAKEA adopts a combination correction strategy $[(1, 20), (1, 80)]$. The correction adjustment ratio is 1:1. Since the total Kriging evaluations are 1000, the ratio 1:1 indicates that the correction adjustment point is the number equal to 500 which is given by $1000 \times \frac{1}{2} = 500$. That is, for

Table 7.10: Parameter Settings of Correction Operations in SOAKEA for Different Cases

Benchmark Function	Total Kriging Evaluations	Kriging Points	Correction Point	Correction Frequency 1	2
SHC	1000	100	1 : 1	20	80
B2	2000	150	1 : 2	3	10
Hmb	2000	100	1 : 2	10	40
Ras	2500	150	1 : 2	5	20

the first 500 Kriging evaluations, the corresponding correction strategy is one correction per 20 generations, and after that one correction per 80 generations. In a combination correction strategy, initial simple correction strategies should provide frequent corrections so that SOAKEA can frequently update current Kriging models and make them more precise. The first row of Table 7.10 shows that for the benchmark function SHC, the correction frequency of the first correction strategy is four times that of the second correction strategy.

The results from SOAKEA are compared with PSO without Kriging and two other well known metaheuristics Simulated Annealing (SA) and Evolutionary Strategy (ES). For SA, ES and PSO without Kriging, actual fitness functions or simulation evaluations are used. Thus, it can be seen how many simulation evaluations are needed for optimization algorithms without Kriging. The comparison results are in Table 7.11. These algorithms record the number of simulation evaluations as soon as the defined optimal solution is reached. The average value of used simulation evaluations in successful optimization search of running an algorithm for 100 times is given in the column of simulation evaluations of Table 7.11. These values are used to compare the optimization performance of different algorithms. Table 7.11 indicates that SOAKEA uses far fewer simulation evaluations but achieves better or comparable success rates. It must be mentioned that SOAKEA has

Table 7.11: Performance Comparison of Different Algorithms

Benchmark	Simulation Evaluations				Success Rate (%)			
Function	SA	ES	PSO	Krig-PSO	SA	ES	PSO	Krig-PSO
SHC	1203	460	250	119	95	97	99	99
B2	4247	1184	650	221	95	99	100	100
Hmb	2440	843	511	108	42	78	82	80
Ras	300145	3456	2365	141	2	7	21	40

some additional computational burden of building Kriging models frequently. But since the number of Kriging points is limited and small in SOAKEA, the incurred additional computational burden can be controlled, and not have critical influences on optimization search.

7.4 Conclusions

This chapter developed an optimization algorithm based on Kriging and evolutionary algorithms for simulation problems with costly computational expenses. The developed algorithm is named SOAKEA. Kriging is its surrogate fitness function and evolutionary algorithms are used to search optimal solutions. Building a surrogate fitness function is an efficient approach to reducing computational cost of running a simulation model when it is computationally prohibitive for standard metaheuristics. The accurate interpolation feature of Kriging is an important factor of selecting it for SOAKEA. Taylor Kriging developed can provide more effective interpolation and thus benefit SOAKEA. The optimization capabilities of SOAKEA are determined by the chosen evolutionary algorithm, and it is necessary to choose an evolutionary algorithm with strong optimization capabilities.

The empirical investigation on some important parameters in SOAKEA, the size of initial samples, correction strategies and Kriging points, provides some critical insights for the settings of these parameters. The empirical comparison of SOAKEA with SA, ES and PSO without Kriging indicates that SOAKEA has significant advantages in optimizing complex simulation models.

CHAPTER 8

THE KRIGING APPLICATION SOFTWARE

8.1 Introduction

In this chapter, the Kriging software named Kriging Modeling is developed. The introduction to the software is divided into four parts, which are the overview, Kriging Models, Data Analysis, and Graphics. Kriging Models, Data Analysis, and Graphics are the three tool packages of the Kriging Modeling software.

8.2 The Overview

The Kriging software makes use of Visual C++ 6.0 and Visual Basic 6.0. The software includes four files which are KrigModels.dll, KrigModeling.exe, KrigInformation.ini and KrigHelp.pdf. KrigModels.dll is a dynamic-link library and developed in Visual C++ 6.0. KrigModeling.exe is an executive file coded in Visual Basic 6.0. KrigInformation.ini is an information file and KrigHelp.pdf is a help file in PDF format which provides the instructions about how to manipulate the software.

The installation of software is simple, and the specific procedure is as follows:

1. Create a file fold: C:\ KrigModels;

2. Save the following files to this designated file folder:

   KrigModels.dll

   KrigModeling.exe

KrigInformation.ini

KrigHelp.pdf

When a user double clicks KrigModeling.exe, two windows like Figure 8.1 will pop up. The small window is the welcome interface of software. It could be displayed while other functions are used. A simple way to remove it is to click the welcome window. The welcome window gives a brief introduction to the software. More introduction information can be obtained by clicking on "About" in the Help submenu under the main menu. Correspondingly, a smaller window like Figure 8.2 will pop up.

In the main menu, the software provides the functions of document editing and file manipulation. Since these functions are very similar to those in Microsoft Word, they are self-explanatory. The software has a help file which is located under the Help submenu in the main menu. If Content in the Help submenu is chosen, a PDF file will open and specific help information will be available. The software has a Windows menu under the main menu. This submenu provides the facility to arrange several working windows in the homepage of Kriging Modeling. The Tools submenu under the main menu gives the three tool packages which are Kriging Models, Data Analysis, and Graphics.

8.3 Kriging Models

By clicking Kriging Models in the Tools submenu under the main menu, a user can open the tool package named Kriging Models. The interface of this tool package is like Figure 8.3. The interface mainly includes two types of parameter settings. One type is the input settings, the other is the output settings. The input settings include the selection of Kriging models, covariance functions and covariance ranges (influence distances). When
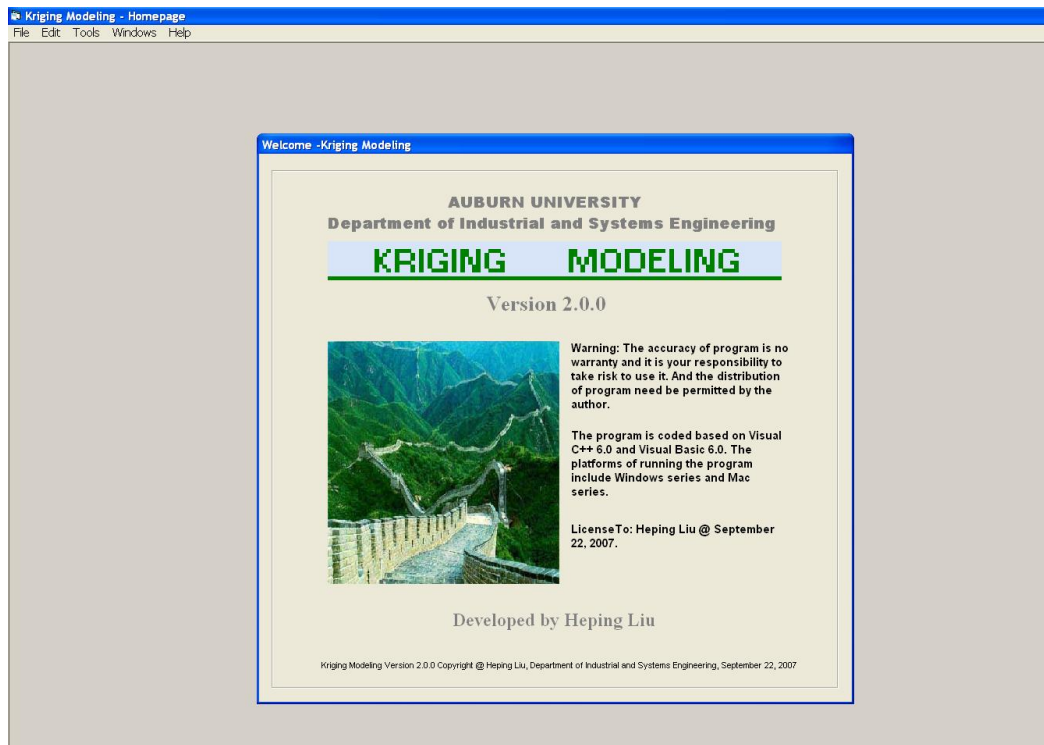
Figure 8.1: The Welcome Interface
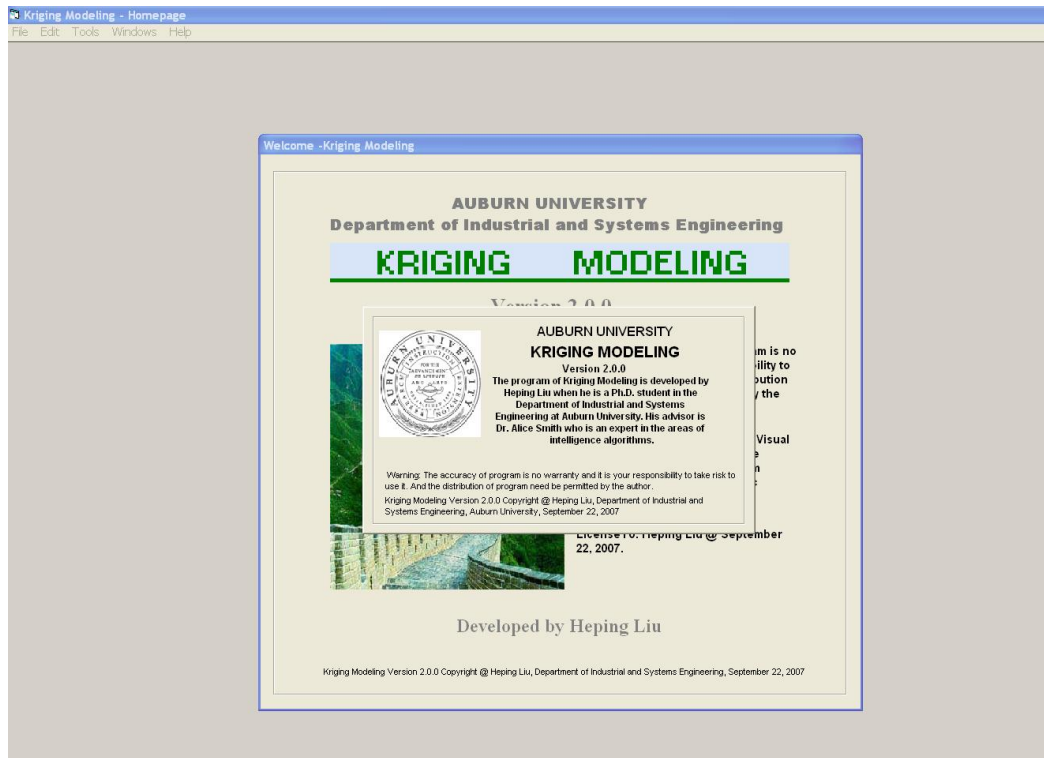
Figure 8.2: The About Interface

the chosen Kriging model is Taylor Kriging, one more parameter, namely order, needs to be specified. Similarly, when the chosen covariance function is the Power-function, a value for powLambda needs to be set. When the chosen covariance range is "Define the Range", a value for influence distance of covariance needs to be provided. The requirements of these parameters are that the highest order of Taylor expansion should be an integer among 1, 2, 3, and 4, the power of the power-function covariance function, that is b in Eq. (3.41), must be less than 2 but greater than 0, and the chosen value of covariance range must be greater than 0. The interface output settings define the output data files.

8.3.1   Data Input File

The data layout in the data input file needs to follow specific rules. Figure 8.4 gives an example of data layout. In this file, the first row of the data file needs to list the number of known points, the number of predicted points and the dimension number of independent variables of each point. In the following rows, the file needs to give known observed points and predicted points. Specifically, in the first set of rows, the known observed points are given with the last column as the dependent variable and the preceding columns representing the independent variables. After the known observed points, the predicted points are given. For these predicted points, the file gives only their independent variable values and Kriging will predict the values of the dependent variable. The dependent variable can only be one dimension, but the Kriging software can simultaneously predict the dependent variable values of multiple points by only running a Kriging Model once.
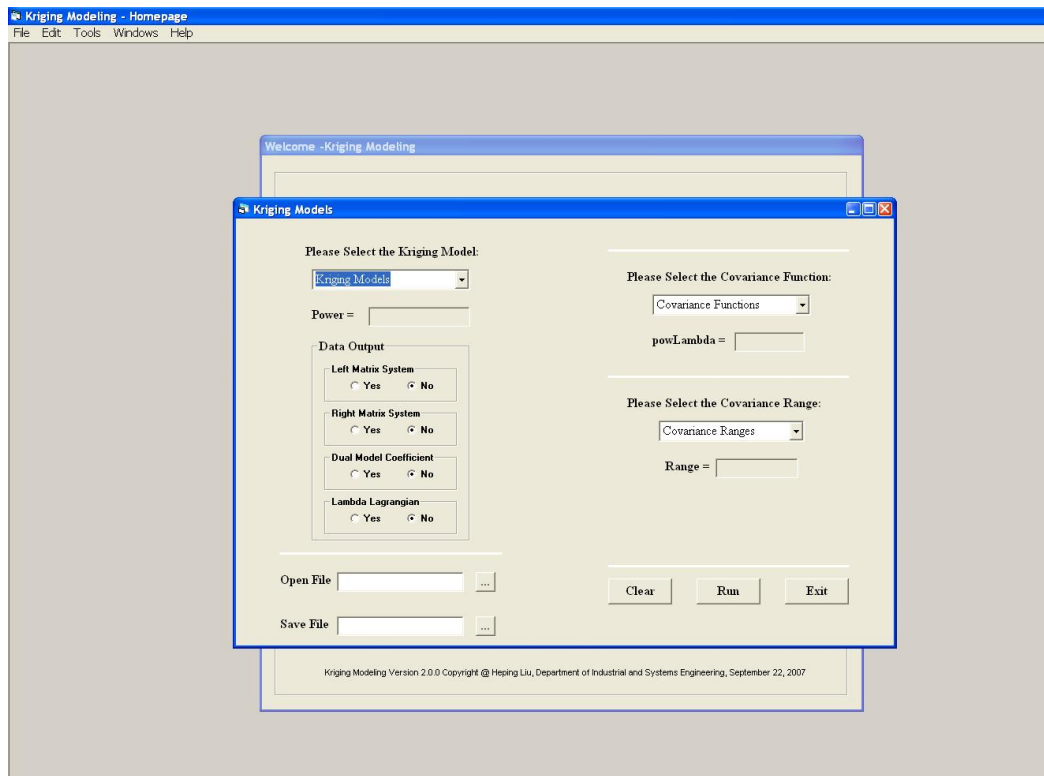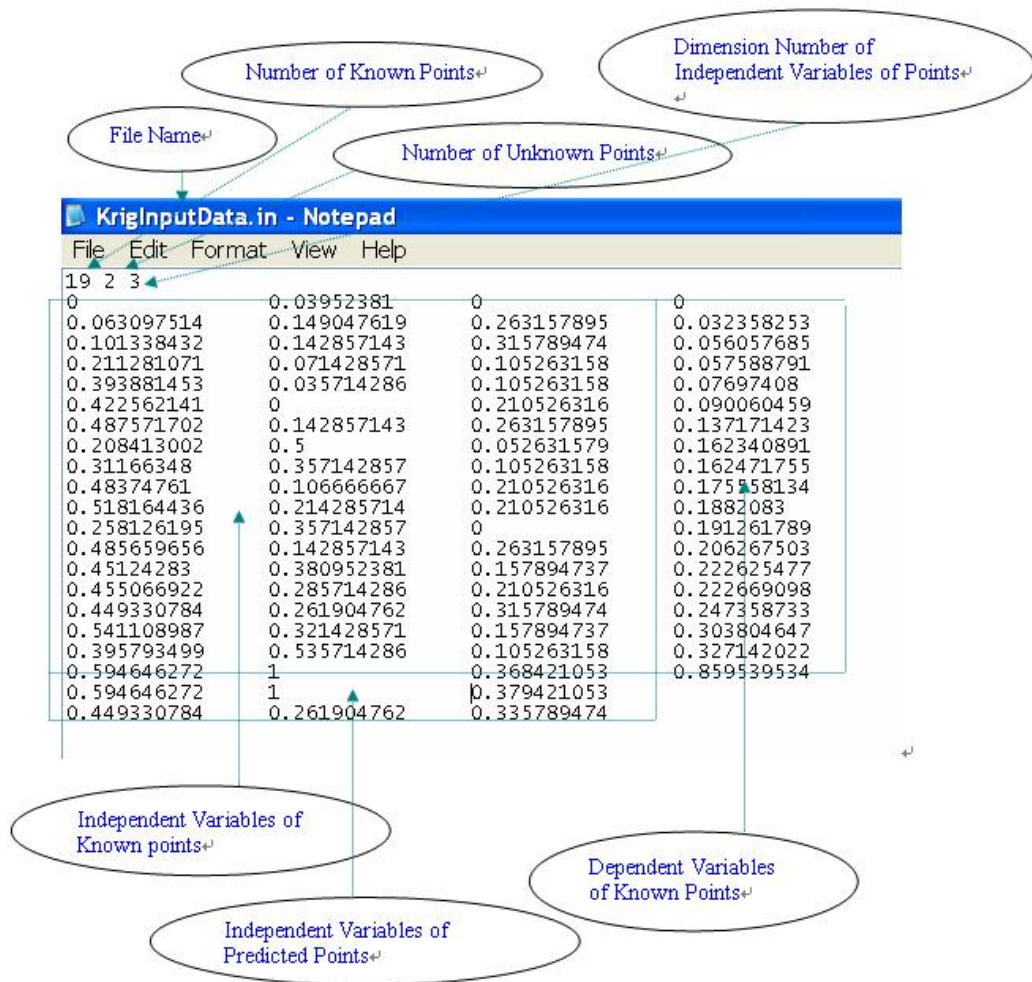
Figure 8.3: The Kriging Models Interface

Figure 8.4: An Example of the Data Layout of the Input File

8.3.2 Running the Software

The interface of Kriging Models provides three buttons which are named Clear, Run and Exit. If a user wants to clear current settings, s/he clicks the Clear button to reset the program in the default status. If a user wants to exit this tool package, s/he clicks the Exit button. If the Run button is clicked, the program will call a related Kriging model to perform required estimation operations. However, if any parameter settings do not satisfy the requirements of the program, the program will report parameter selection mistakes and the failure of calling the program, and at the same time identify the sources of mistakes by changing the color of incorrect parameter settings like Figure 8.5. If the process of calling a Kriging model is successful, the program will give a corresponding report, and ask a user whether s/he needs to see output results. If the user wants to check output results, the program will automatically open the file KrigOutputData.out.

8.3.3 Data Output File

Kriging Models can output some data files which are KrigOutputData.out, Simp-KrigOutputData.out, MatrixLeft.out, MatrixRight.out, DualModelCoefficients.out, and Lamb-daLagrangian.out. The interface in Figure 8.6 provides some choices to decide which files need be output. Notice that KrigOutputData.out and SimpKrigOutputData.out will be automatically output, and SimpKrigOutputData.out is the simplified version of KrigOut-putData.out.

Figure 8.6 gives an example of KrigOutputData.out. In this file, the first several rows log the following information:

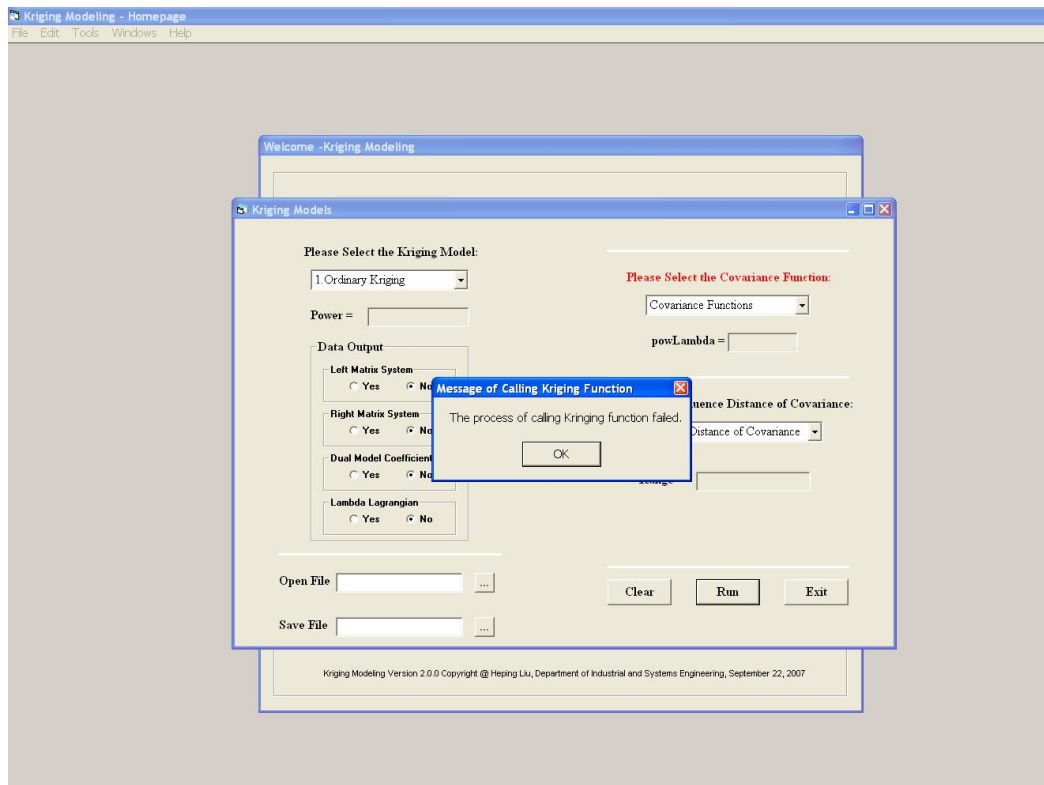1. The date and time of running the Kriging model;

Figure 8.5: The Results of Calling a Kriging Model

2. The path of input data file;

3. The index of chosen Kriging model;

4. The index of chosen covariance function;

5. The order of Power covariance function;

6. The index of chosen covariance range;

7. The chosen covariance range value;

8. The highest order in Taylor Expansion;

9. Input data;

10. Output data;

The index of a chosen Kriging model can equal to -1, 0, 1, or 2. These numbers correspond to the indices under "Please Select the Kriging Model" in the interface of Kriging Models. The matching relationship is as follows:

-1 → Kriging Models (Or, not identifying any Kriging model)

0 → Simple Kriging

1 → Ordinary Kriging

2 → Taylor Kriging

When Taylor Kriging is chosen, the software will ask a user to input the highest order of Taylor expansion. The highest order must be equal to 1, 2, 3 or 4.

The index of the chosen covariance function can be equal to -1, 0, 1, 2, 3, 4, 5, 6, 7, or 8. These numbers correspond to the indices under "Please Select the Covariance Function"

in the interface of Kriging Models. The matching relationship is:

-1 → Covariance Functions (Or, not identifying any covariance function)

0 → Pure Nugget Covariance

1 → Linear Covariance

2 → Cubic-1 Covariance

3 → Cubic-2 Covariance

4 → Spherical Covariance

5 → Exponential Covariance

6 → Gaussian Covariance

7 → Power-function Covariance

8 → Logarithmic (de Wijsian) Covariance

The index of the chosen covariance range can be equal to -1, 0, 1, 2, 3, 4, 5, or 6. These numbers correspond to the indices under "Please Select the Covariance Range" in the interface of Kriging Models. And the matching relationship is:

-1 → Covariance Range (Or, not identifying any covariance range)

0 → Default (Five Standard Deviations)

1 → One Standard Deviation

2 → Two Standard Deviations

3 → Three Standard Deviations

4 → Four Standard Deviations

5 → Five Standard Deviations

By checking output parameter settings, a user can know whether s/he chose the desired settings. Besides these parameter values, KrigOutputData.out also outputs input and output data.

SimpKrigOutputData.out is the simplified version of KrigOutputData.out. In SimpKrigOutputData.out, there are only the data of interpolated points, which makes it easier to manipulate output data.

Additionally, in the frame "Data Output", the software provides four options to choose if more data will be output. These four choices are four data output files MatrixLeft.out, MatrixRight.out, DualModelCoefficients.out, and LambdaLagrangian.out. MatrixLeft.out and MatrixRight.out output the left coefficient matrix and the right matrix in Kriging matrix system, respectively. DualModelCoefficients.out gives the coefficients of the corresponding dual Kriging model. And LambdaLagrangian.out gives the coefficients $\lambda_i(i = 1, \cdots, n)$ and Lagrangian multipliers $u_l(l = 1, \cdots, M)$. The data from these output files can help a user conduct more thorough analysis. However, in the default status, to increase the run-time efficiency of the program, the software will not output these files.

8.4   Data Analysis

The software provides a tool package named Data Analysis which is in the Tools submenu under the main menu. To perform the interpolation using Kriging, normalizing the input data or finding the extreme or mean values in the input data are often done. This tool performs the above operations. Its interface is like Figure 8.7. Similar to Kriging
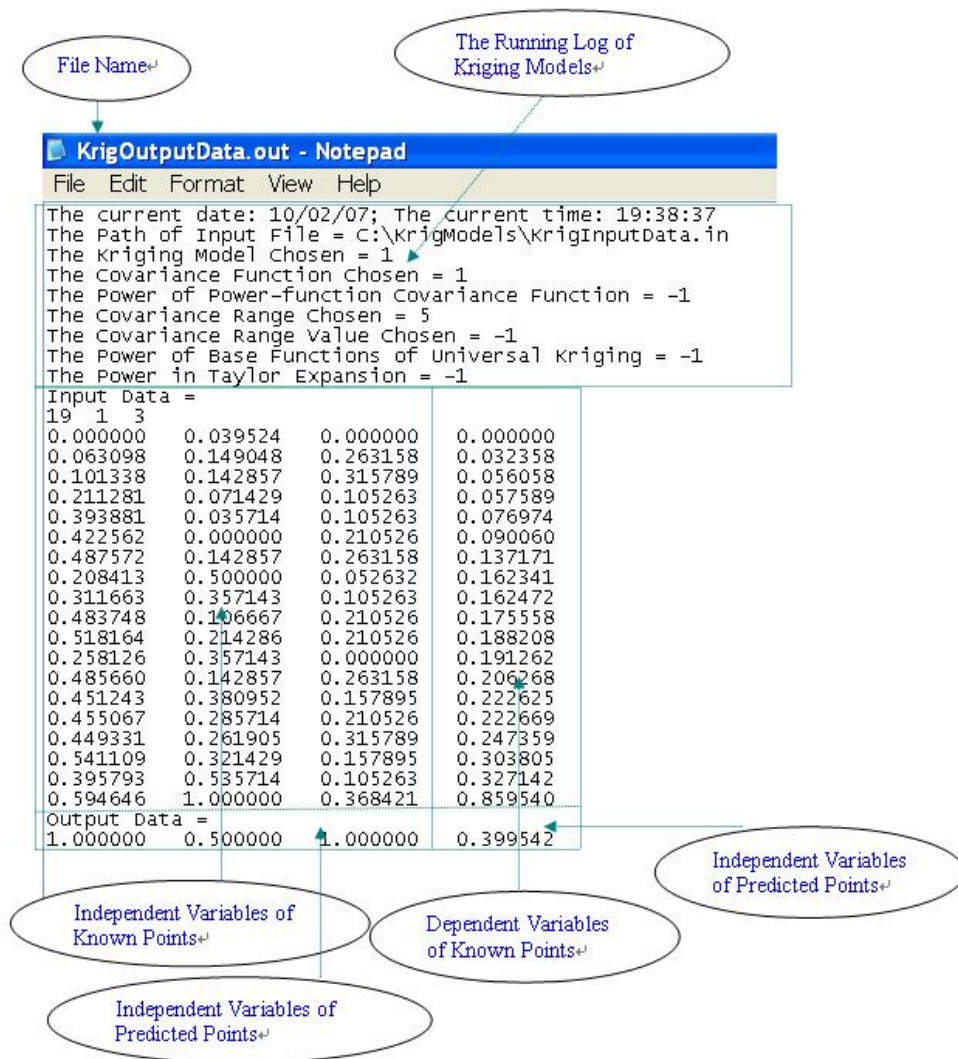
The Running Log of Kriging Models

File Name

KrigOutputData.out - Notepad

File   Edit   Format   View   Help

```
The current date: 10/02/07; The current time: 19:38:37
The Path of Input File = C:\KrigModels\KrigInputData.in
The Kriging Model Chosen = 1
The Covariance Function Chosen = 1
The Power of Power-function Covariance Function = -1
The Covariance Range Chosen = 5
The Covariance Range Value Chosen = -1
The Power of Base Functions of Universal Kriging = -1
The Power in Taylor Expansion = -1
Input Data =
19  1  3
0.000000    0.039524    0.000000    0.000000
0.063098    0.149048    0.263158    0.032358
0.101338    0.142857    0.315789    0.056058
0.211281    0.071429    0.105263    0.057589
0.393881    0.035714    0.105263    0.076974
0.422562    0.000000    0.210526    0.090060
0.487572    0.142857    0.263158    0.137171
0.208413    0.500000    0.052632    0.162341
0.311663    0.357143    0.105263    0.162472
0.483748    0.106667    0.210526    0.175558
0.518164    0.214286    0.210526    0.188208
0.258126    0.357143    0.000000    0.191262
0.485660    0.142857    0.263158    0.206268
0.451243    0.380952    0.157895    0.222625
0.455067    0.285714    0.210526    0.222669
0.449331    0.261905    0.315789    0.247359
0.541109    0.321429    0.157895    0.303805
0.395793    0.535714    0.105263    0.327142
0.594646    1.000000    0.368421    0.859540
Output Data =
1.000000    0.500000    1.000000    0.399542
```

Independent Variables of Predicted Points

Independent Variables of Known Points

Dependent Variables of Known Points

Independent Variables of Predicted Points

Figure 8.6: An Example of the Data Layout of the Output File (KrigOutputData.out)

147

Models, Data Analysis needs an input file. In the input file, the data layout has to follow the same rules as those of the input file of Kriging Models. That is, the number of points and the dimensions of each point must be placed in the first row of the input file. From the second row, the file then gives the data that will be normalized, or the point vectors. If the points have multiple dimensions, the tool package will normalize the data by dimension. After finishing the required operations, Data Analysis creates a data output file named DataAnalysis.out. In the data output file, the following parameters are output:

1. The date and time of running data analysis;

2. The path of the input data file;

3. The normalized data (by dimension);

4. The mean (by dimension);

5. The mean of all of the data;

6. The maximum and minimum values (by dimension);

However, if the option of any operation is set to "No", the corresponding data results will not be output. For example, for the operation of Data Mean, if its operation option is "No", the data output file will not output any mean values.

The interface of Data Analysis provides three buttons which are Clear, Run and Exit. A user clicks the Clear button to set the program in the default status. If the Exit button is clicked, the program will exit this tool package. If the Run button is clicked, the program will perform required operations, such as normalizing input data. However, if all operation options are "No", the program will report parameter selection mistakes. After finishing
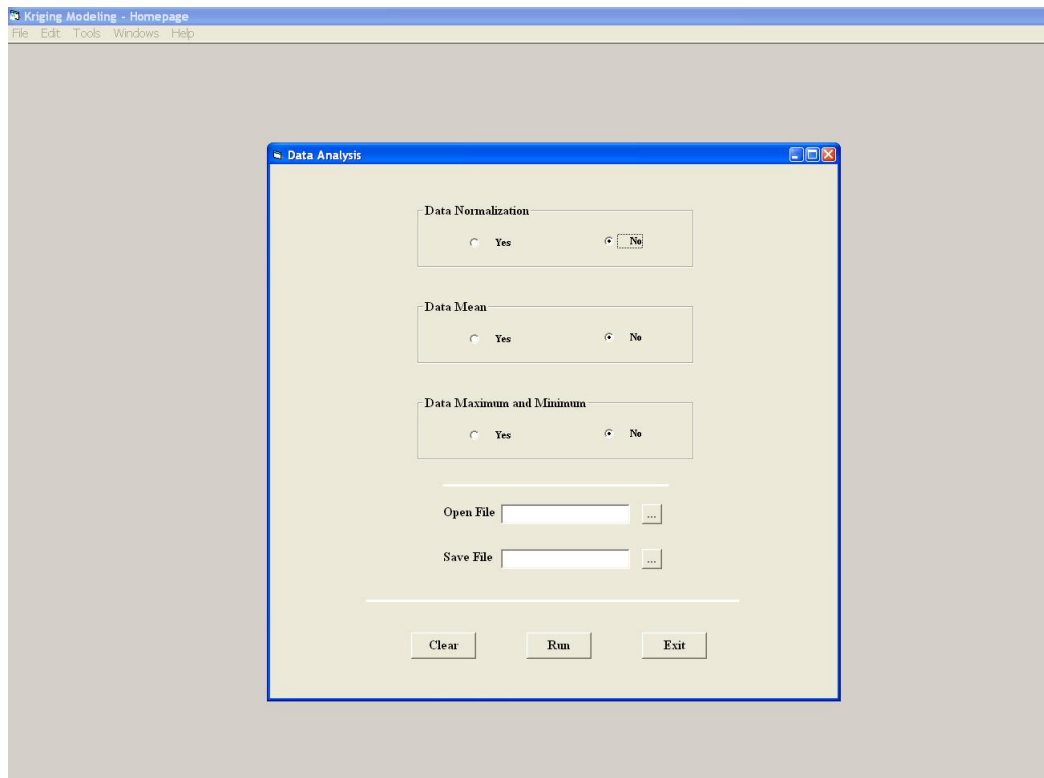
Figure 8.7: The Data Analysis Interface

required operations, the program will report if the process of data analysis is successful. If the calling process fails, the program will identify the mistake source by changing the color of the mistaken parameter setting as shown in Figure 8.8. If the calling process is successful, the program will ask a user whether s/he needs to see output results which are in a file named DataAnalysis.out.

8.5   Graphics

Kriging Modeling also provides a tool package named Graphics to do graphic analysis on data. The Graphics tool package is in the Tools submenu under the main menu in the window of Kriging Modeling. Figure 8.9 gives the interface of the Graphics tool package. In order to start drawing a figure, a user needs to double click the area of the figure in the interface.

After double clicking the interface figure in the Graphics window, the user will see a working interface like Figure 8.10. For the convenience of later operations, the user could make this window full-screen by double clicking the blue bar of the Graphics window. The Graphics window itself provides a main menu with some functions which can help users perform needed graphic operations.

Data input is the first step in the procedure for drawing a figure. One way to input data is that a user could click Datasheet under the View submenu in the Graphics window. The datasheet window will then pop up. By following the rules of data layout in this sheet, the user can input data. The other way to input data is to directly identify the input data file by clicking "Import File ···" under the Edit submenu in the Graphics window. The identified input file must be an excel file.
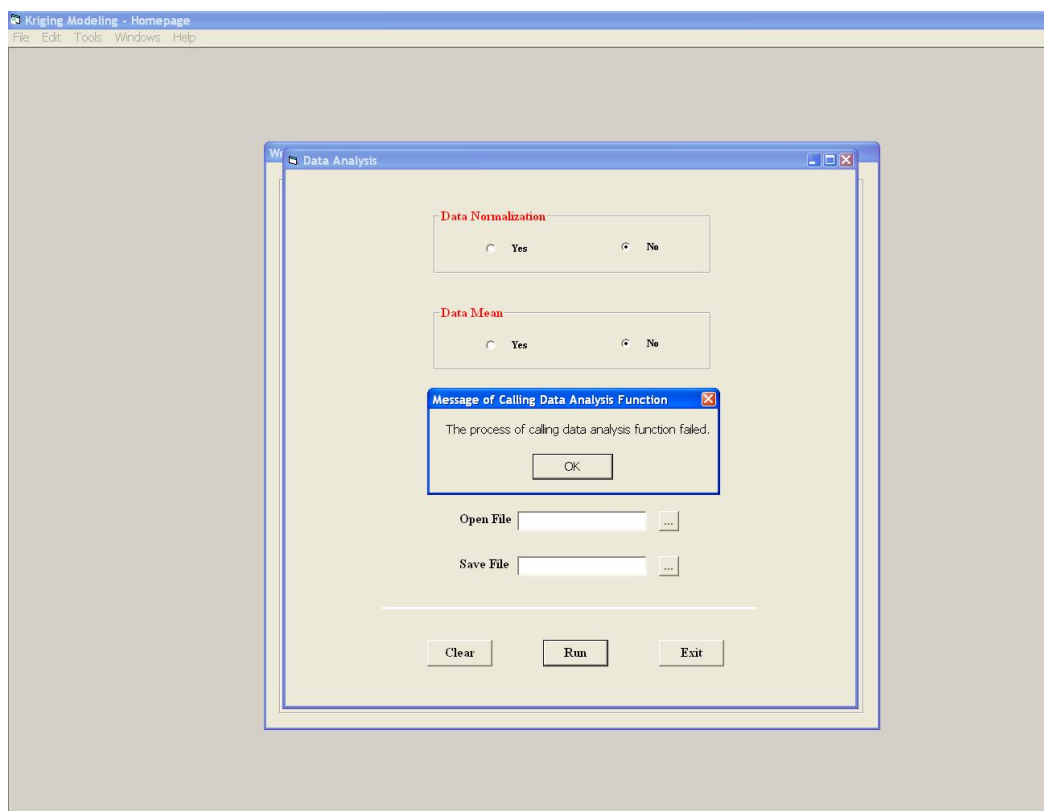
Figure 8.8: The Results of Calling Data Analysis
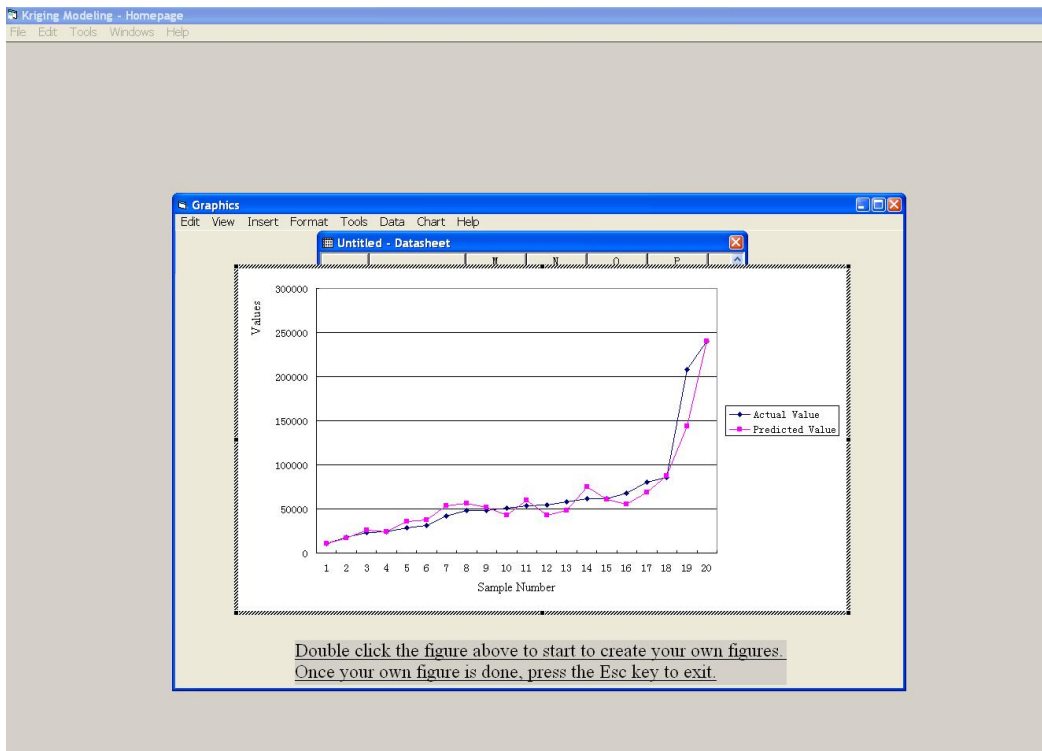
Figure 8.9: The Graphics Interface
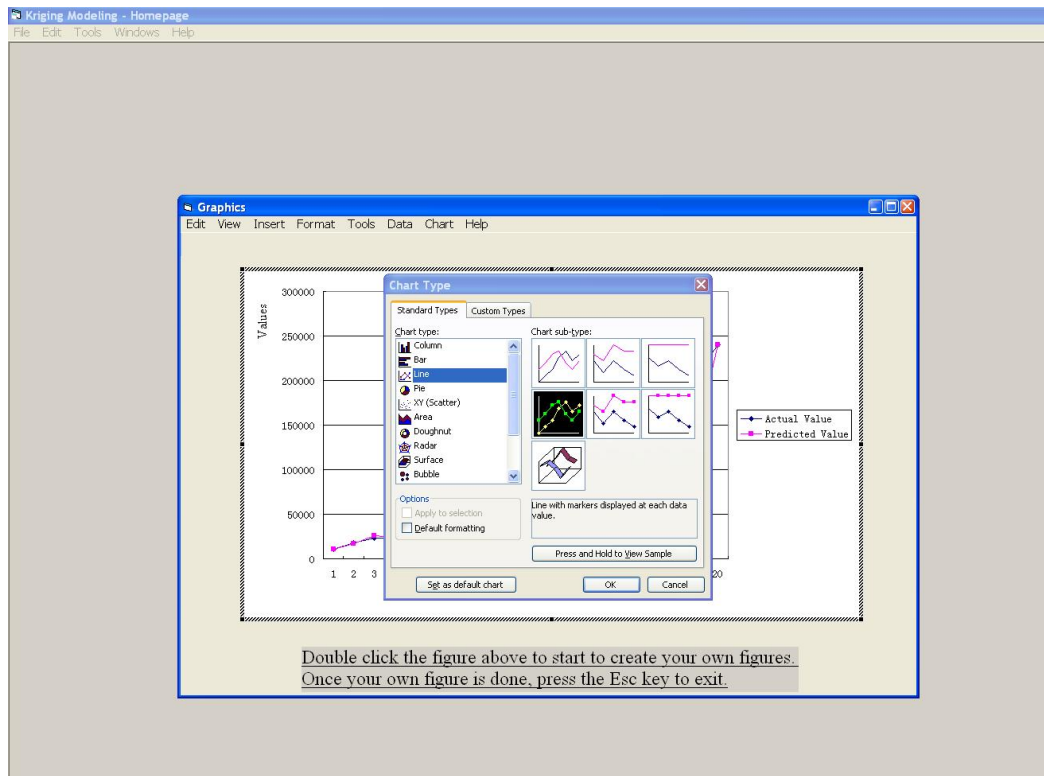
Figure 8.10: The Working Interface of the Graphics Tool Package

The Chart submenu in the Graphics window provides some functions to set the graphic parameters. After clicking Chart Type under the Chart submenu, a user can see an interface like Figure 8.10. Here, a user can choose which type of figure to draw, and can decide which parameters to set or make some adjustments as would be usually done for the figures in an excel file. For example, a user can change legends, the title of a figure and the titles of axes, and so on.

One more important submenu in the Graphics window is the Data submenu, which provides two methods to set data series in the drawn figure, that is, Series in Rows and Series in Columns.

After a user finishes drawing a figure, the Copy command under the Edit submenu in the Graphics window is used to copy the drawn figure to any identified file. To exit the operation of drawing a figure, press the Esc key.

8.6   Conclusions

The chapter developed the Kriging software named Kriging Modeling which includes three tool packages called Kriging Models, Data Analysis and Graphics, respectively. The application instructions of these tool packages were presented in detail. The installation of software is simple. The simulation cases in the previous chapters show that the software is a powerful interpolation and data analysis tool. In addition, it is not only effective for simulation interpolation but also for any other areas needing interpolation.

CHAPTER 9

SUMMARY AND CONCLUSIONS

This chapter concludes the dissertation and discusses the future research directions.

9.1 Conclusions

It has been established in Kriging literature that Kriging metamodeling is applied to the interpolation, sensitivity analysis and optimization of simulation models. Kriging has the properties of precise interpolation, explicitness and differentiation. Based on these properties, Kriging is used to conduct simulation interpolation in order to reduce the expenses of running a simulation model, to assist sensitivity analysis on a simulation model, and to optimize a simulation model. Simulation models considered are computationally expensive and obtaining their sufficient outputs is difficult. Some important results of the dissertation are as follows:

1. Taylor expansion is introduced to construct a drift function of Kriging. The new drift function can capture complicated drift behaviors of data so that simulation interpolation capabilities of Kriging are enhanced. A Kriging model with Taylor expansion is defined as Taylor Kriging (TK). The empirical results from deterministic, stochastic and physical simulations demonstrate that TK has stronger interpolation capabilities. In addition, sample standard deviation is suggested to measure influence distance of covariance, which makes different problems have a consistent measurement unit for influence distance of covariance.

The simulation interpolation capabilities of Kriging, regression and artificial neural networks (ANN) are compared theoretically and empirically. A significant feature in the comparison is that multicollinearity, heteroscedasticity and specification error are considered. The conclusions of the comparison are that for interpolation accuracy, Kriging is better than regression but worse than ANN. Regression and Kriging have explicit fitted functions, but the function relationship of ANN is completely hidden.

2. The partial differentiation equation of Kriging is developed and used together with analysis of variance (ANOVA) to assist sensitivity analysis on a simulation model. Sensitivity analysis can help modelers to find critical input factors of a simulation model. Thus, controlling, improving and optimizing a simulation model become easier. The empirical results show that the sensitivity analysis assisted by Kriging is effective.

3. The dissertation integrates Kriging with evolutionary algorithms to search for optimal inputs of a simulation model with costly computational expenses. A novel algorithm is developed and named SOAKEA. The properties of SOAKEA are analyzed. Its parameter settings are empirically investigated, and some important conclusions about how to set parameters for SOAKEA are obtained. Several simulation cases are used to examine the optimization capabilities of SOAKEA. The results indicate that it is a promising optimization tool for simulation models with expensive running costs.

4. The Kriging application software is developed. The software has a user-friendly graphic interface. It is flexible and powerful and manipulating it is simple. The software is not only for simulation interpolation but for wider areas which need interpolation operations.

9.2 Future Research Directions

Kriging metamodeling has some attractive properties for simulation interpolation, sensitivity analysis and optimization. The potential research directions include the following aspects:

1. If a Kriging coefficient matrix is singular, its corresponding Kriging matrix system can not be obtained. The Kriging estimator thus does not exist. Finding a good method to avoid this problem is important.

2. While constructing a Kriging model, a modeler has to choose among some potential covariance functions. Generating an approach to efficiently identifying a covariance function with limited observations deserves further consideration.

3. Kriging usually assumes that the covariance between any two points only depends on their distance and is not influenced by their locations. However, this is not always true. When this assumption is partially or completely relaxed, how to make Kriging accommodate the relaxed assumption need be investigated.

4. Taylor expansion relies on the premise that continuous partial derivatives of a drift function have to exist. However, if this premise does not exist, it is worthy to investigate how to identify an approximate drift function to describe complicated drift of data.

5. The fundamental objective of SOAKEA is to integrate the prediction features of Kriging with optimization functions of an evolutionary algorithm in order to search for optimal inputs of a simulation model. Improving SOAKEA can be done from two aspects. One is to enhance the interpolation capabilities of Kriging; another is to find a

157

better optimization algorithm. More work is needed in these two aspects. In addition, currently SOAKEA is only applied to deterministic simulation optimization. How to extend SOAKEA to optimize stochastic simulation models should be examined.

6. The simulation models investigated in this dissertation have multiple input variables but only one output variable. However, some simulation models may have several output variables. How to use Kriging to address such simulation models deserves exploration. A Kriging model with two or more dependent variables is named cokriging. Cokriging has the potentials to fit simulation models with multiple output variables. Similarly, its partial differentiation equation can be developed and used to assist in sensitivity analysis of simulation models. Some evolutionary algorithms can handle multi-objective optimization problems. The integration of cokriging and these evolutionary algorithms can be applied to optimizing simulation models with multiple output variables.

<center>BIBLIOGRAPHY</center>

[1] Adleman, L. M., "Molecular Computation of Solutions to Combinatorial Problems," Science, Vol. 266, pp: 1021-1024, 1994.

[2] Ahmed, S. and de Marsily, G., "Comparison of Geostatistical Methods for Estimating Transmissivity Using Data on Transmissivity and Specific Capacity," Water Resources Research, Vol. 23 (9), pp: 1717-1737, 1987.

[3] Andr, B., Trochu, F. and Dansereau, J., "Approach for the Smoothing of Three-dimensional Reconstructions of the Human Spine Using Dual Kriging Interpolation," Medical and Biological Engineering and Computing, Vol. 34 (3), pp: 185-191, 1996.

[4] Armstrong, M. and Myers, D., "Robutness of Kriging: Variogram Modeling and Its Applications," Technical Report, Centre de Geostatistique, Fontainebleau, France, 1984.

[5] Aunon, J. and Gomez-Hernandez, J. J., "Dual Kriging with Local Neighborhoods: Application to the Representation of Surface," Mathematical Geology, Vol. 32 (1), pp: 69-86, 2000.

[6] Barton, R. R. "Metamodeling: A State of the Art Review," Proceedings of the 1994 Winter Simulation Conference ed. by Tew, J. D., Manivannan, S., Sadowski, D. A. and Seila, A. F., pp: 237-244, 1994.

[7] Box, G. E. P. and Jenkins, G. M., "Time Series Forecasting and Control," Holden-Day Inc., San Francisco, 1976.

[8] Branin, F. H., "Widely Convergent Methods for Finding Multiple Solutions of Simultaneous Nonlinear Equations," IBM Journal of Research Developments, Vol. 16, pp: 504-522, 1972.

[9] Bras, R. L. and Rodriguez-Iturbe, I., "Random Functions and Hydrology," Addison-Wesley Publishing Company, Reading, MA, 1985.

[10] Brass, J., Gerrard, A. M. and Peel, D. "Estimating Vessel Cost via Neural Networks," Proceedings of the 13th International Cost Engineering Congress, London, 1994.

[11] Carrier, J., Aubin, C.-E., Trochu, F. and Labelle, H., "Optimization of Rib Surgery Parameters for the Correction of Scoliotic Deformities Using Approximation Models," Journal of Biomechanical Engineering, Vol. 127 (4), pp: 680-691, 2005.

[12] Chaveesuk, R. and Smith, A. E. "Dual Kriging: An Exploratory Use in Economic Metamodeling," The Engineering Economist, Vol. 50 (3), pp: 247-271, 2005.

<center>159</center>

[13] Cheng, S. J. and Wang, R. Y., "An Approach for Evaluating the Hydrological Effects of Urbanization and its Application," Hydrological Processes, Vol. 16, pp: 1403-1418, 2002.

[14] Chiles, J. and Delfiner, P., "Geostatistics: Modeling Spatial Uncertainty," John Wiley & Sons, New York, pp: 355, 1999.

[15] Choobineh, F. and Behrens, A., "Use of Intervals and Possibility Distributions in Economic Analysis," Journal of the Operational Research Society, Vol. 43 (9), pp: 907-918, 1992.

[16] Clifton, P. M. and Neuman, S. P., "Effects of Kriging and Inverse Modeling on Conditional Simulation of the Avra Valley Aquifer in Southern Arizona," Water Resources Research, Vol. 18 (4), pp: 1215-1234, 1982.

[17] Cressie, NaA. C., "Statistics for Spatial Data," John Wiley & Sons, New York, 1993.

[18] Cressie, N., "The Origins of Kriging," Mathematical Geology, Vol. 22 (3), pp: 239-252, 1990.

[19] Curran, R., Gomis, G., Castagne, S., Butterfield, J., Edgar, T., Higgins, C. and McKeever, C., "Integrated Digital Design for Manufacture for Reduced Life Cycle Cost," International Journal of Production Economics, Vol. 109, pp: 27-40, 2007.

[20] David, M., "Geostatistical Ore Reserve Estimation," Elsevier, Amsterdam, 1977.

[21] Davis, J. C.,"Statistics and Data Analysis in Geology," John Wiley & Sons, New York, 1986.

[22] Davis, M. and David, M., "Automatic Kriging and Contouring in the Presence of Trends (Universal Kriging Made Simple)," The Journal of Canadian Petroleum Technology, Vol. 17 (1), 1978.

[23] Davis, M. W. and Culhane, P. G., "Contouring Very Large Datasets Using Kriging," Verly G. et al. (eds), Geostatistics for Natural Resources Characterization, D. Reidel Publishing Co., Part 2, pp: 599-619, 1984.

[24] Davis, M. W. and Grivet, C., "Kriging in a Global Neighborhood," Mathematical Geology, Vol. 16 (3), pp: 249-265, 1984.

[25] De Iacoa, S., Myers, D. E. and Posa, D., "Space-time Variogram and a Functional Form for Total Air Pollution Measurements," Computational Statistics and Data Analysis, Vol. 41, pp: 311-328, 2002.

[26] Delhomme, J. P., "Kriging in the Hydrosciences," Advanced Water Resources, Vol. 1 (5), pp: 251-266, 1978.

[27] Delhomme, J. P., "Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach," Water Resources Research, Vol. 15 (2), pp: 269-280, 1979.

[28] Delorme, S., Petit, Y., de Guise, J. A., Aubin, C.-E., Labelle, H., Landry, C. and Dansereau, J., "Three-dimensional Modelling and Rendering of the Human Skeletal Trunk from 2D Radiographic Images," Proceedings of the Second International Conference on 3-D Imaging and Modeling, Ottawa, Ontario, Canada, pp: 497-505, October 4-8, 1999.

[29] Diamond, P. and Armstrong, M., "Robustness of Variograms and Conditioning of Kriging Matrices," Mathematical Geology, Vol. 16, pp: 809-822, 1984.

[30] Doncker, Ph. De, Cognet, X., Dricot, J-M and Meys, R., "Electromagnetic Wave Propagation Prediction Using Spatial Statistics: Experimental Validation," Proceedings of the 9th Symposium on Communications and Vehicular Technology in the Benelux, Louvain-la-Neuve, Belgium, October 2002.

[31] Doncker, Ph. De, Dricot, J-M, Meys, R., Heliery, M. and Tabbaray W., "Kriging the Fields: A New Statistical Tool for Wave Propagation Analysis," Proceedings of the Intertional Conference on Electromagnetics in Advanced Applications, Turin, Italy, September 2003.

[32] Dorigo, M. and Gambardella, L. M., "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem," IEEE Transaction on Evolutionary Computation, Vol. 1 (1), pp: 53-66, 1997.

[33] Dowd, P. A., "A Review of Geostatistical Techniques for Contouring," Earnshaw, R. A. (eds), Fundamental Algorithms for Computer Graphics, Springer-Verlag, Berlin, NATO ASI Series, Vol. F17, 1985.

[34] Dubrule, O., "Two Methods with Different Objectives: Splines and Kriging," Mathematical Geology, Vol. 15 (2), pp: 245-257, 1983.

[35] Echaabi, J., Trochu, F. and Gauvin, R., "A General Strength Theory for Composite Materials Based on Dual Kriging Interpolation," Journal of Reinforced Plastics and Composites, Vol. 14 (3), pp: 211-232, 1995.

[36] El-Beltagy, M. A., Nair, P. B. and Keane, A. J., "Metamodelling Techniques for Evolutionary Optimization of Computationally Expensive Problems: Promise and Limitations," Proceeding of the Genetic and Evolutionary Compution Conference (GECCO99), pp: 196-203, 1999.

[37] Falkner, C. H. and Shanker, N., "Physical Simulation of Flexible Manufacturing Systems," Proceedings of the 1984 Winter Simulation Conferences ed. by Sheppard, S., Pooch, U. and Pdgden, D., pp: 397-404, 1984.

[38] Fan, S. K. S., Liang, Y. C. and Zahara E., "A Genetic Algorithm and a Particle Swarm Optimizer Hybridized with Nelder-Mead Simplex Search,", Computers & Industrial Engineering, Vol. 50, pp: 401-425, 2006.

[39] Fang, K. T., Li, R. Z. and Sudjianto, A., "Design and Modeling for Computer Experiments," Taylor & Francis Group, London, 2006.

[40] Fereshteh-Saniee, F. and Hosseini, A. H., "The Effects of Flash Allowance and Bar Size on Forming Load and Metal Flow in Closed Die Forging," Journal of Materials Processing Technology, Vol. 177, pp: 261-265, 2006.

[41] Fogel, L. J., Owens, A. J., and Walsh, M. J., "Artificial Intelligence through Simulated Evolution," John Wiley & Sons, New York, 1966.

[42] Fregeau, M., Saeed, F. and Paraschivoiu, I., "Numerical Heat Transfer Correlation for Array of Hot-air Jets Impinging on 3-Dimensional Concave Surface," Journal of Aircraft, Vol. 42 (3), pp: 665-670, 2005.

[43] Galli, A., Murillo, E. and Thomann, J., "Dual Kriging-Its Properties and Its Uses in Direct Contouring," Verly G. *et al.* (eds), Geostatistics for Natural Resources Characterization, D. Reidel Publishing Co., Part 2, pp: 621-634, 1984.

[44] Galli, A. and Sandjivy, L., "Analyse Krigeante et Analyse Spectrale," Sciences de la Terre, Serie Informatique Geologique, Vol. 21, pp: 115-124, 1985.

[45] Gerrard, A. M., Brass, J. and Peel, D. "Using Neural Nets to Cost Chemical Plants," Proceedings of the 4th European Symposium on Computer-Aided Process Engineering, pp: 475-478, 1994.

[46] Glover, F., "Tabu Search - Part I," Operations Research Society of America Journal on Computing, Vol. 1 (3), pp: 190-206, 1989 .

[47] Glover, F., "Tabu Search - Part II," Operations Research Society of America Journal on Computing, Vol. 2 (1), pp: 4-32, 1990.

[48] Goldberg, D. E., "Genetic Algorithms in Search, Optimization, and Machine Learning," Addison-Wesley Publishing Company, Reading, Massachusetts,1989.

[49] Goovaerts, P., "Factorial Kriging Analysis: A Useful Tool for Exploring the Structure of Multivariate Spatial Soil Information," Journal of Soil Science, Vol. 43, pp: 597-619, 1992.

[50] Goovaerts, P., "Using Elevation to Aid the Geostatistical Mapping of Rainfall Erosivity," Catena, Vol. 34 (3-4), pp: 227-242, 1999.

[51] Goovaerts, P., "Geostatistics in Soil Science: State-of-the-art and Perspectives," Geoderma, Vol. 89, pp: 1-45, 1999.

[52] Hemyari, P. and Nofziger, D. L., "Analytical Solution for Punctual Kriging in One Dimension," Soil Science Society of America journal, Vol. 51 (1), pp: 268-269, 1987.

[53] Hertog, D., Kleijnen, J. P. C. and Siem, A. Y. D., "The Correct Kriging Variance Estimated by Bootstrapping," Journal of the Operational Research Society , Vol. 57 (4), pp: 400-409, 2006.

[54] Hofmeyr, S. A. and Forrest, S., "Immunity by Design: An Artificial Immune System," in Proceedings of the Genetic and Evolutionary Computing Conference, Orlando, FL, pp: 1289-1296, 1999.

[55] Holland, J. H., "Adaptation in Natural and Artificial Systems," University of Michigan Press, Ann Arbor, 1975.

[56] http://en.wikipedia.org/wiki/Meta-modeling.

[57] http://en.wikipedia.org/wiki/Ramsey_RESET_test.

[58] http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/etsug_aut oreg_sect025.htm.

[59] http://papers.ssrn.com/sol3/papers.cfm?abstract_id=570142.

[60] Huang, D., Allen, T. T., Notz, W. I. and Miller, R. A., "Sequential Kriging Optimization Using Multiple-fidelity Evaluations," Structural and Multidisciplinary Optimization, Vol. 32 (5), pp: 369-382, 2006.

[61] Huang, D., Allen, T. T., Notz, W. I. and Zheng, N., "Global Optimization of Stochastic Black-box Systems via Sequential Kriging Metamodels," Journal of Global Optimization, Vol. 34 (3), pp: 441-466, 2006.

[62] Jin, R., Chen, W. and Simpson, T. W., "Comparative Studies of Metamodeling Techniques under Multiple Modeling Criteria," Structural and Multidisciplinary Optimization, Vol. 23, pp: 1-13, 2001.

[63] Jones, D., Schonlau, M. and Welch, W., "Efficient Global Optimization of Expensive Black-box Functions," Journal of Global Optimization, Vol. 13, pp: 455-492, 1998.

[64] Journel, A. G., "Nonparametric Estimation of Spatial Distributions," Mathematical Geology, Vol. 15 (3), pp: 445-468, 1983.

[65] Journel, A. G. and Huijbregis, Ch. J., "Mining Geostatistics," Academic Press, London, 1978.

[66] Jowett, G. H., "The Accuracy of Systematic Sampling from Conveyor Belts," Applied Statistics, Vol. 1, pp: 50-59, 1952.

[67] Jowett, G. H., "Multiple Regression between Variables Measured at Points of a Plane Sheet," Applied Statistics, Vol. 4, pp: 80-89, 1955.

[68] Kamanayo, G., Trochu, F. and Sanschagrin, B., "Prediction of Shrinkage by Dual Kriging for Injection Molded Polypropylene Plaques," Advances in Polymer Technology, Vol. 13 (4), pp: 305-314, 2003.

[69] Kennedy, J. and Eberhart, R., "Particle Swarm Optimization," in Proceedings of the 1995 IEEE International Conference on Neural Networks, pp: 1942-1948, 1995.

[70] Kennedy, M. C. and O'Hagan, A., "Predicting the Output of a Complex Computer Code when Fast Approximation Are Available," Biometrika, Vol. 87 (1), pp: 1-13, 2000.

[71] Khoo, L. P. and Chen, C. H., "Integration of Response Surface Methodology with Genetic Algorithms," The International Journal of Advance Manufacturing Technology, Vol. 18, pp: 483-489, 2001.

[72] Kleijnen, J. P. C., "An Overview of the Design and Analysis of Simulation Experiments for Sensitivity Analysis," European Journal of Operational Research, Vol. 164 (2), pp: 287-300, 2005.

[73] Kleijnen, J. P. C., "Experimental Design for Sensitivity Analysis, Optimization, and Validation of Simulation Models," Chapter 6 in: Handbook of Simulation, edited by Banks, John Wiley & Sons, New York, pp: 173-223, 1998.

[74] Kleijnen, J. P. C. and van Beers, W. C. M., "Application-driven Sequential Designs for Simulation Experiments: Kriging Metamodeling," Journal of the Operational Research Society, Vol. 55, pp: 876-883, 2004.

[75] Kleijnen, J. P. C. and van Beers, W. C. M., "Robustness of Kriging when Interpolating in Random Simulation with Heterogeneous Variances: Some Experiments," European Journal of Operational Research, Vol. 165 (3), pp: 826-834, 2005.

[76] Koehler, J. R. and Owen A. B., "Computer Experiments," Ghosh, S. and Rao, C. R., Handbook of Statistics, Elsevier, Amsterdam, Vol. 13, pp: 261-308, 1996.

[77] Kolmogorov, A. N., "Sur 1'interpolation et Extrapolation des Suites Stationnaires," Comptes Rendus Academie des Sciences, Paris, Vol. 208, pp: 2043-2045, 1939.

[78] Kolmogorov, A. N., "The Distribution of Energy in Local Isotropic Turbulence," Doklady Akademii Nauk SSSR, Moscow, Vol. 32, pp: 19-31, 1941.

[79] Korzybski, A. and Meyers, R., "Science and Sanity: An Introduction to Nonaristotelian Systems and General Semantics," Fourth edition. International Nonaristotelian Library Publishing Company, Lakeville, 1958.

[80] Koza, J. R., "Genetic Programming: On the Programming of Computers by Natural Selection," MIT Press, Cambridge, MA, 1992.

[81] Krige, D. G., "A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand," Journal of the Chemical, Metallurgical and Mining Society of South Africa, Vol. 52, pp: 119-139, 1951.

[82] Krige, D. G., "Lognormal-de Wijsian Geostatistics for Ore Evaluation," South African Institute of Mining and Metallurgy, Johannesburg, 1978.

[83] Kushner, H. J., "A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise," Journal of Basic Engineering, Vol. 86, pp: 97-106, 1964.

[84] Larranaga, P. and Lozano, J. A., "Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation," Kluwer Academic Publishers Group, Norwell, Massachusetts, 2002.

[85] Leary, S. J., Bhaskar, A. and Keane, A. J., "A Knowledge-based Approach To Response Surface Modelling in Multifidelity Optimization," Journal of Global Optimization, Vol. 26 (3), pp: 297-319, 2003.

[86] Leflbvre, J., Roussel, H., Walter, E., Lecointe, D. and Tabbara, W., "Prediction from Wrong Models: The Kriging Approach," IEEE Antennas and Propagation Magazine, Vol. 38 (4), pp: 35-45, 1996.

[87] Li, B. H., Rizos, C. and Lee, H. K, "Utilizing Kriging to Generate a NLOS Error Correction Map for Network Based Mobile Positioning," Journal of Global Positioning Systems, Vol. 4 (1-2), pp: 27-35, 2005.

[88] Li, B. H., Salter, J., Dempster, A. G. and Rizos, C., "Indoor Positioning Techniques Based on Wireless LAN," 1st IEEE International Conference on Wireless Broadband & Ultra Wideband Communications, Sydney, Australia, 13-16 March, paper 113, CD-ROM procs, 2006.

[89] Limaiem, A. and ElMaraghy, H. A., "Automatic Inference of Parametric Equations in Geometric Modeling Using Dual Kriging," in Proceedings of 1996 IEEE International Conference on Robotics and Automation, Vol. 2, pp: 1499-1504, 1996.

[90] McKay, M. D., Beckman, R. J. and Conover, W. J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, Vol. 21, pp: 239-245, 1979.

[91] Mamat, A., Trochu, T. F. and Sanschagrin, B., "Analysis of Shrinkage by Dual Kriging for Filled and Unfilled Polypropylene Molded Parts," Polymer Engineering and Science, Vol. 35 (19) , pp: 1511-1520, 2004.

[92] Matheron, G., "Traite de Geotatistique Appliquee, Tome I: Memoires du Bureau de Recherches Geologiques et Minieres," Editions Technip, Paris, Vol. 14, 1962.

[93] Matheron, G., "Traite de Geotatistique Appliquee, Tome II: Memoires du Bureau de Recherches Geologiques et Minieres," Editions Technip, Paris, Vol. 24, 1963.

[94] Matheron, G., "Principles of Geostatistics," Economic Geology, Vol. 58 (8), pp: 1246-1266, 1963.

[95] Matheron, G., "Les Variables Regionalisees et Leur Estimation," Une application de la theorie des fonctions aleatoires aux Sciences de la Nature, Masson et Cie, Paris, France, 1965.

[96] Matheron, G., "Le krigeage universel," Cachiers du Centre de Morphologie Mathematique, Ecole des Mines de Paris, Fontainebleau, Vol. 1, 1969.

[97] Matheron, G., "A Simple Substitute for the Conditional Expectation: The Disjunctive Kriging," In: Guarascio, David, and Huijbregts (eds), "Advanced Geostatistics in the Mining Industry," Vol. 24 of NATO Advances Study Institute Series; Series C: Mathematical and Physical Sciences, D. Reidel Publishing Company, Dordrecht (NL), pp: 221-236, 1976.

[98] Matheron, G., "Splines and Kriging: Their Formal Equivalence," in Merriam, D. F. (eds.), Down-to-earth Statistics: Solutions Looking for Geological Problems, Syracuse University Geological Contributions, pp: 77-95, 1981.

[99] Matheron, G., "Pour une Analyse Krigeante de Donnes Rgionalises," Note N- 732 du Centre de Gostatistique. Ecole des Mines de Paris, Fontainebleau, France, pp: 1-22, 1982.

[100] Matson, J. E., Barret, B. and Mellichamp, J., "Software Estimation Using Function Points," IEEE Transactions on Software Engineering, Vol. 20, pp: 275-287, 1994.

[101] McLean, P., Lger, P. and Tinawi, R., "Post-processing of Finite Element Stress Fields Using Dual Kriging Based Methods for Structural Analysis of Concrete Dams," Finite Elements in Analysis and Design, Vol. 42 (6), pp: 532-546, 2006.

[102] Michalewicz, Z., "Genetic Algorithms + Data Structures = Evolution Programs," (Third, Revised and Extended Edition), Springer-Verlag, London, UK, 1996.

[103] Mijaylova, P., Sandoval, L., Garzn, M. A., Moeller, G., Chacn, J. M. and Gonzles, A., "Pilot Scale Evaluation and Economical Feasibility on Different Wastewater Treatment Options to Obtain Reclaimed Water in Mexico," Usos Mltiples del Agua, para la Vida yel Desarrollo Sostenible, Cartagena de Indias, pp: 10-18, 2003.

[104] Mitchell, T. J. and Morris, M. D., "The Spatial Correlation Function Approach to Response Surface Estimation," Swain, J. J., Goldsman, D., Crain, R. C., and Wilson, J. R. (eds), Proceedings of the 1992 Winter Simulation Conference, pp: 565-571, 1992.

[105] Moyeed, Rana A. and Papritz, Andreas "An Empirical Comparison of Kriging Methods for Nonlinear Spatial Point Prediction," Mathematical Geology, Vol. 34 (4), 2002.

[106] Niazi, A., Dai, J. S., Balabani, S. and Seneviratne, L., "Product Cost Estimation: Technique Classification and Methodology Review," Journal of Manufacturing Science and Engineering, Vol. 128, pp: 563-575, 2006.

[107] Odeh, I., McBratney, A. and Chittleborough, D., "Spatial Prediction of Soil Properties from Landform Attributes Derived from a Digital Elevation Model," Geoderma, Vol. 63 (3-4), pp: 197-214, 1994.

[108] Odeh, I., McBratney, A. and Chittleborough, D., "Further Results on Prediction of Soil Properties from Terrain Attributes: Heterotopic Cokriging and Regression-Kriging," Geoderma, Vol. 67 (3-4), pp: 215-226, 1995.

[109] Olea, R. A., "Optimal Contour Mapping Using Universal Kriging," Journal of Geophysical Research, Vol. 79, pp: 695-702, 1974.

[110] Ong, Y. S., Nair, P. B. and Keane, A. J., "Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modeling," American Institute of Aeronautics and Astronautics Journal, Vol. 41 (4), pp: 687-696, 2003.

[111] Ord, J. K., "Kriging, Entry," Kotz, S. and Johnson, N. (eds.), Encyclopedia of Statistical Sciences, John Wiley & Sons Inc., New York, Vol. 4, pp: 411-413, 1983.

[112] Orr, E. D. and Dutton, O. R., "An Application of Geostatistics to Determine Regional Groundwater Flow in the San Andres Formation, Texas and New Mexico," Groundwater, Vol. 21 (5), pp: 619-624, 1983.

[113] Owen, A. B., "Orthogonal Arrays for Computer Experiments, Integration and Visualization," Statistica Sinica, Vol. 2, pp: 439-452, 1992.

[114] Ramsey, J. B., "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," Journal of the Royal Statistical Society, Series B (31), pp: 350-371, 1969.

[115] Ratle, A., "Accelerating Evolutionary Algorithms Using Fitness Function Models," Parallel Problem Solving from Nature, Vol. 5, pp: 87-96, 1998.

[116] Ratle, A., "Kriging as a Surrogate Fitness Landscape in Evolutionary Optimization," Artificial Intelligence for Engineering Design Analysis and Manufacturing, Vol. 15 (1), pp: 37-49, 2001.

[117] Rechenberg, I., "Evolutionsstrategie: Optimierung technischer systeme nach prinzipien der biologischen evolution," Frommann-Holzboog Verlag, Stuttgart, Germany, 1973.

[118] Rgnire, J. and Sharov, A., "Simulating Temperature- Dependent Ecological Processes at the Subcontinental Scale: Male Gypsy Moth Flight Phenology," International Journal of Biometereology, Vol. 42, pp: 146-152, 1999.

[119] Reis I. A., "Alternatives for Geosensors Networks Data Analysis," VII Simpsio Brasileiro de Geoinformtica, Campos do Jord.o, Brasil, INPE, pp: 94-104, 20-23 novembro 2005.

[120] Reisa, A. P., Sousab, A. J., Ferreira da Silvac, E., Patinhac, C. and Fonsecac, E. C., "Combining Multiple Correspondence Analysis with Factorial Kriging Analysis for Geochemical Mapping of the Gold-silver Deposit at Marrancos (Portugal)," Applied Geochemistry, Vol. 19, pp: 623-631, 2004.

[121] Reynolds, R. G., "An Introduction to Cultural Algorithms," in Proceedings of the Third Annual Conference on Evolutionary Program, San Diego, CA, pp: 131-139, 1994.

[122] Sabin, M. A., "Contouring - the State of the Art," Earnshaw, R. A. (eds), Fundamental Algorithms for Computer Graphics, Springer-Verlag, NATO ASI Series, Vol. F17, 1985.

[123] Sacks, J., Schiller, S. B. and Welch, W., "Design for Computer Experiments," Technometrics, Vol. 31, pp: 41-47, 1989.

[124] Sacks, J., Welch, W. J., Mitchell T. J. and Wynn, H. P., "Design and Analysis of Computer Experiments," Statistical Science Vol. 4 (4), pp: 409-435, 1989.

[125] Santner, T. J., Williams, B. J. and Notz, W. I., "The Design and Analysis of Computer Experiments," Springer-Verlag, New York, 2003.

[126] Sasena, M. J., "Flexibility and Efficiency Enhancements for Constrained Global Design Optimization with Kriging Approximations," Ph. D. dissertation, University of Michigan, 2002.

[127] Sasena, M. J., Papalambros, P. Y. and Goovaerts, P., " Exploration of Metamodeling Sampling Criteria for Constrained Global Optimization," Engineering Optimization, Vol. 34, pp: 263-278, 2002.

[128] Savi, M. A., Paiva, A., Baeta-Neves, A. P. and Pacheco, P. M. C. L., "Phenomenological Modeling and Numerical Simulation of Shape Memory Alloys: A Thermo-Plastic-Phase Transformation Coupled Model," Journal of Intelligent Materials Systems and Structures, Vol. 13, pp: 261-273, 2002.

[129] Scheff, H., "The Analysis of Variance," John Wiley & Sons Inc., 1959.

[130] Schwefel, H. P., "Evolution and Optimum Seeking," John Wiley & Sons, New York, 1995.

[131] Schonlau, M., "Computer Experiments and Global Optimization," Ph.D. Dissertation, University of Waterloo, 1997.

[132] Simpson, T. W., Mauery, T. M., Korte, J. J. and Mistree, F., "Comparison of Response Surface and Kriging Models for Multidisciplinary Design Optimization," 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, MO, Sept. 2-4, 1998, Collection of Technical Papers. Pt. 1 (A98-39701 10-31).

[133] Simpson, T. W., Mauery, T. M., Korte, J. J. and Mistree, F., "Kriging Models for Global Approximation in Simulation-Based Multidisciplinary Design Optimization," AIAA Journal, Vol. 39 (12), pp: 2233-2241, 2001.

[134] Smith, A. E., Mason and A. K., "Cost Estimation Predictive Modeling: Regression versus Neural Network," Engineering Economist, Vol. 42, pp: 137-161, 1996.

[135] Smith, J. L., Halvorson, J. J. and Papendick, R. I., "Using Multiple-variable Indicator Kriging for Evaluating Soil Quality," Soil Science Society of America Journal, Vol. 57 (3), pp: 743-749, 1993.

[136] Soares, A., "Geostatistical Estimation of Orebody Geometry: Morphological Kriging," Mathematical Geology, Vol. 22 (7), pp: 787-802, 2004.

[137] Song, W. B., "Shape Optimisation Of Turbine Blade Firtrees," Ph. D. Dissertation, University of Southampton, 2002.

[138] Sousa, A. J., "Factorial Kriging as a Method to Include Spatial Structure into Classification. A Case Study on a Sulphide Orebody," Chung, C. F., Fabbri, A. G. and Sinding-Larsen, R. (Eds.), Quantitative Analysis of Mineral and Energy Resources, Reidel, Dordrecht, pp: 385-392, 1988.

[139] Stein, M., "Large Sample Properties of Simulations Using Latin Hypercube Sampling," Technometrics, Vol. 29, pp: 143-151, 1987.

[140] Stein, M. L., "Asymptotically Efficient Spatial Interpolation with a Misspecified Covariance Function," Annals of Statistics, pp: 55-63, 1988.

[141] Stein, M. L. and Handcock M. S., "Some Asymptotic Properties of Kriging when the Covariance Function is Misspecified," Mathematical Geology, Vol. 21 (2), pp: 171-189, 1989.

[142] Sukhatme, S., "Kriging with Perturbed Variogram," Proceedings of the Social Statistics Section of the American Statistical Association Annual Meetings, pp: 296-299, 1985.

[143] Terriault, P., Meunier, M.-A. and Trochu F., "Application of Dual Kriging to the Construction of a General Phenomenological Material Law for Shape Memory Alloys," Journal of Intelligent Material Systems and Structures, Vol. 8 (7), pp: 605-618, 1997.

[144] Trangmar, B. B., Yost, R. S. and Uehara, G., "Application of Geostatistics to Spatial Studies of Soil Properties," Advances in Agronomy, Vol. 38, pp: 45-94, 1985.

[145] Trochu. F., "A Contouring Program Based on Dual Kriging Interpolation," Engineering with Computers, Vol. 9, pp: 160-177, 1993.

[146] Trochu, F., Sacepe, N., Volkov, O. and Turenne, S., "Characterization of NiTi Shape Memory Alloys Using Dual Kriging Interpolation," Materials Science and Engineering, Vol. 273-275, pp: 395-399, 1999.

[147] Tynan, R., OHare, G. M. P., Marsh, D. and OKane, D., "Interpolation for Wireless Sensor Network Coverage," Embedded Networked Sensors, 2005, EmNetS-II. The Second IEEE Workshop, pp: 123-131, May 30-31, 2005.

[148] Van Beers, W. C. M. and Kleijnen, J. P. C., "Kriging for Interpolation in Random Simulation," Journal of the Operational Research Society, Vol. 54 (3), pp: 255-262, 2003.

[149] Van Beers, W. C. M. and Kleijnen, J. P. C., "Kriging Interpolation in Simulation: A Survey," Ingalls, R. G., Rossetti, M. D., Smith, J. S., and Peters B. A., (eds), Proceedings of the 2004 Winter Simulation Conference, Vol. 55, pp: 113-121, 2004.

[150] Vieira, S. R., Hatfield, J. L., Nielsen, D. R. and Biggar, J. W., "Geostatistical Theory and Application to Variability of Some Agronomical Properties," Hilgardia , Vol. 51, pp: 175, 1983.

[151] Viscarra Rossel, R. A., Goovaerts , P. and McBratney, A. B., "Assessment of the Production and Economic Risks of Site-specific Liming Using Geostatistical Uncertainty Modelling," Environmetrics, Vol. 12 (8), pp: 699-711, 2001.

[152] Voltz, M. and Webster, R., "A Comparison of Kriging, Cubic Splines and Classification for Predicting Soil Properties from Sample Information," European Journal of Soil Science, Vol. 41, pp: 473-490, 1990.

[153] Wackernagel, H., "Multivariate Geostatistics: An Introduction with Applications," 2nd Edition, Springer-Verlag, 1998.

[154] Ward, T. L., "Discounted Fuzzy Cash Flow Analysis," 1985 Annual International Industrial Engineering Conference Proceedings, pp: 476-481, 1985.

[155] Warnes, J. J., "A Sensitivity Analysis for Universal Kriging," Mathematical Geology, pp: 653-676, 1986.

[156] Webster, R., "Local Disjunctive Kriging of Soil Properties with Change of Support," European Journal of Soil Science, Vol. 42, pp: 301-318, 1991.

[157] Wiener, N., "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications," Published jointly by the Technology Press of the Massachusetts Institute of Technology, Cambridge, Massachusetts, John Wiley & Sons, Inc., New York, and Chapman & Hall, Ltd., London, pp: 163, 1949.

[158] Wold, H., "A Study in the Analysis of Stationary Time Series," Almquist & Wiksell, Stockholm, Sweden, pp: 211, 1938, 2nd Edition, with Appendix by Peter Whittle, pp: 236, 1954.

[159] Yakowitz, S. J. and Szidarovszky, F., "A Comparison of Kriging with Nonparametric Regression Methods," Journal of Multivariate Analysis, Vol. 16, pp: 21-53, 1985.

[160] Yu, Y., Govindan, R., Rahimi, M. and Estrin, D., "Two Case Studies on Data Sensitivity of Wireless Sensor Network Algorithms and our Proposal on Scalable, Synthetic Data Generation," International Journal of Distributed Sensor Networks, Vol. 2, pp: 355-386, 2006.

[161] Zhou, Z. Z, Ong, Y. S. and Nair, P. B., "Hierarchical Surrogate-Assisted Evolutionary Optimization Framework," Congress on Evolutionary Computation, 2004 (CEC2004), Vol. 2, pp: 1586-1593, 2004.

[162] Zielinski, K. and Laur, R., "Stopping Criteria for a Constrained Single-Objective Particle Swarm Optimization Algorithm," Informatica, Vol. 31, pp: 51-59, 2007.

[163] Zielinski, K., Peters, D. and Laur, R., "Stopping Criteria for Single-Objective Optimization," Proceedings of the Third International Conference on Computational Intelligence, Robotics and Autonomous Systems, 2005, Singapore.

Appendix A

Proof Of Kriging Estimator In Equation (3.11)

According to Eq. (3.11), the following equations can be obtained:

$$\mathbf{F}^T\boldsymbol{\lambda} = \mathbf{f}(\mathbf{X}) \tag{A.1}$$

$$\mathbf{F}\mathbf{U} + \mathbf{S}\boldsymbol{\lambda} = \mathbf{C}(\mathbf{X}) \tag{A.2}$$

Equation (A.2) can be transformation into the form below:

$$\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F}\mathbf{U} + \mathbf{F}^T\mathbf{S}^{-1}\mathbf{S}\boldsymbol{\lambda} = \mathbf{F}^T\mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) \tag{A.3}$$

That is,

$$\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F}\mathbf{U} = \mathbf{F}^T\mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) - \mathbf{F}^T\boldsymbol{\lambda} \tag{A.4}$$

If $\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F}$ is invertible, the estimator of $\mathbf{U}$ can be calculated:

$$\hat{\mathbf{U}} = (\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) - \mathbf{F}^T\boldsymbol{\lambda}) \tag{A.5}$$

Based on Eq. (3.2), the estimator $\hat{\mathbf{U}}$ becomes

$$\hat{\mathbf{U}} = (\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) - \mathbf{f}(\mathbf{X})) \tag{A.6}$$

Substituting $\mathbf{U}$ in Eq. (A.2) with Eq. (A.6) leads to the following equation.

$$\mathbf{F}\hat{\mathbf{U}} + \mathbf{S}\boldsymbol{\lambda} = \mathbf{C}(\mathbf{X}) \tag{A.7}$$

which can be further transformed into

$$\mathbf{S}^{-1}\mathbf{F}\hat{\mathbf{U}} + \mathbf{S}^{-1}\mathbf{S}\boldsymbol{\lambda} = \mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) \tag{A.8}$$

Here $\mathbf{S}$ is supposed to be invertible. Then the estimator of $\hat{\boldsymbol{\lambda}}$ is

$$\hat{\boldsymbol{\lambda}} = \mathbf{S}^{-1}(\mathbf{C}(\mathbf{X}) - \mathbf{F}\hat{\mathbf{U}}) \tag{A.9}$$

Note that the estimator of $\hat{Z}(\mathbf{X})$ is

$$\hat{Z}(\mathbf{X}) = \hat{\boldsymbol{\lambda}}^T\mathbf{Z} = (\mathbf{S}^{-1}(\mathbf{C}(\mathbf{X}) - \mathbf{F}\hat{\mathbf{U}}))^T\mathbf{Z} = (\mathbf{C}^T(\mathbf{X}) - \hat{\mathbf{U}}^T\mathbf{F}^T)\mathbf{S}^{-1}\mathbf{Z} \tag{A.10}$$

According to Eq. (A.6), the estimator of $\hat{Z}(\mathbf{X})$ becomes:

$$\hat{Z}(\mathbf{X}) = \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}\mathbf{Z} - ((\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{S}^{-1}\mathbf{C}(\mathbf{X}) - \mathbf{f}(\mathbf{X})))^T\mathbf{F}^T\mathbf{S}^{-1}\mathbf{Z} \tag{A.11}$$

That is,

$$\hat{Z}(\mathbf{X}) = \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}\mathbf{Z} - \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}\mathbf{F}(\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{S}^{-1}\mathbf{Z} +$$
$$\mathbf{f}^T(\mathbf{X})(\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{S}^{-1}\mathbf{Z} \qquad (A.12)$$

Let

$$\hat{\mathbf{a}} = (\mathbf{F}^T\mathbf{S}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{S}^{-1}\mathbf{Z} \qquad (A.13)$$

Eq. (A.12) is simplified:

$$\hat{Z}(\mathbf{X}) = \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}\mathbf{Z} - \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}\mathbf{F}\hat{\mathbf{a}} + \mathbf{f}^T(\mathbf{X})\hat{\mathbf{a}} \qquad (A.14)$$

That is,

$$\hat{Z}(\mathbf{X}) = \hat{\mathbf{a}}^T\mathbf{f}(\mathbf{X}) + \mathbf{C}^T(\mathbf{X})\mathbf{S}^{-1}(\mathbf{Z} - \mathbf{F}\hat{\mathbf{a}}) \qquad (A.15)$$

Thus Eq. (3.11) is proved.