



Thailand Statistician

July 2011; 9(2) : 171-193

<http://statassoc.or.th>

Contributed paper

Exploratory Optimal Latin Hypercube Designs for Computer Simulated Experiments

Rachadaporn Timun [a,b] Anamai Na-udom* [a,b] and Jaratsri Rungrattanaubol [c]

[a] Department of Mathematics, Faculty of Science, Naresuan University,
Phitsanulok 65000, Thailand.

[b] Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd., Bangkok 10400,
Thailand.

[c] Department of Computer Science and Information Technology, Faculty of Science,
Naresuan University, Phitsanulok 65000, Thailand.

* Author for correspondence; e-mail: anamain@nu.ac.th

Received: 21 April 2011

Accepted: 9 June 2011

Abstract

The aim of this paper is to present the construction of the optimal design for computer simulated experiments (CSE) based on three different classes of Latin hypercube design (LHD), random Latin hypercube design (RLHD), symmetric Latin hypercube design (SLHD), and orthogonal array-based Latin hypercube design (OALHD), respectively. We first consider the property of design through various optimality criteria such as ϕ_p criterion, maximin distance criterion, and the mean of correlation coefficient between design columns. After the design properties of each class of design are validated, we compare the prediction accuracy of the surrogate models namely Response surface methodology (RSM) and Kriging model (KRG), conducted by using the optimal design from those three classes of LHD. The results indicate that OALHD has the best design property over all dimensions of problem under consideration. Moreover, OALHD is superior to SLHD and RLHD in terms of prediction accuracy when both of RSM and KRG models are performed. Hence OALHD is recommended as the best design choice for CSE.

Keywords: computer simulated experiments, optimal Latin hypercube design, statistical modeling method.

1. Introduction

Currently computer simulated experiments (CSE) have replaced the physical experiments to investigate a physical complex phenomena, especially when physical experiments are not feasible. CSE is deterministic in nature; hence an identical setting of input variables always produces an identical of output response. The CSEs are usually time consuming and computationally expensive to run. Moreover, the process of CSE is mainly relied on the dimension of problems. Therefore the optimal Latin hypercube designs (OLHD) that aim to spread the design points over the region of interest is more desirable.

Normally, Kriging models (KRG) along with the OLHD are practiced in the context of CSE. The Latin hypercube design (LHD) was originally proposed by McKay et al. [1] and has received wide attention in various applications of CSE. LHD is a matrix $\mathbf{X}$ which contains n rows d columns, where n is the number of runs, and d is the number of input variables. The total numbers of possible LHD are $(n!)^d$ and hence when the dimension of problem is increased, the searching time for constructing OLHD is time consuming.

The construction of optimal design for CSE consists of two approaches called non-search based and search based, respectively. The first approach (non-search based) aims to construct the optimal design by using mathematical or statistical theory. These methods are complicated in the construction of designs and the run size of design is not flexible. For example, Tang [2] presented the construction of orthogonal array-based Latin hypercube design (OALHD) using the generation of the random orthogonal array (random OA). The OALHD has a good property but the run size is not flexible as its construction requires the validity of a random orthogonal array. Ye [3] used the algebraic method called Kronecker product to construct an orthogonal LHD. This method performs well in the applications of CSE but there is one disadvantage as the run size of design is not flexible e.g. $n = 2^d$ or $2^d + 1$ with $2d + 2$ columns. Butler [4] proposed the construction of optimal and orthogonal LHD generated by William transformation. Though this design has a good projection property, the number of run must be an odd number or prime number only. Therefore the construction of this design is quite

complicated. The second approach (search based) of constructing an optimal LHD is based on exchange of elements by using search algorithm. This method first creates a design with different class of LHD e.g. random Latin hypercube design (RLHD), symmetric Latin hypercube design (SLHD), and OALHD, respectively. Then try to search for the optimal LHD using search algorithms under pre-specified optimality criteria. For instance, Ye et al. [5] presented the construction of symmetric Latin hypercube design (SLHD), and compared the search algorithm which proposed by Park [6], Morris and Mitchell [7] (SA), and Li and Wu [8] (Columnwise-pairwise algorithm (CP)) by using the six different dimension problem. The authors recommended that SLHD is superior to random LHD in terms of design property such as the maximin distance criterion. Further, the construction of optimal SLHD requires less searching time than random LHD with respect to entropy and maximin distance criteria. It has also been reported that SA performed better than CP for the large dimension of problem whereas CP performs much more efficient than SA for small dimension of problem. Rungrattanaubol and Na-udom [9] compared the efficiency of two popular evolutionary search algorithms for finding optimal LHD called SA and Genetic algorithm (GA). The result indicated that SA performed better than GA for all cases under consideration.

There are various research papers that have been published in the area of construction of the optimal design for CSE. Leary et al. [10] studied the construction of LHD by using the idea of orthogonal array and compared the design property obtained from SA and CP. The results indicated that OALHD is superior to RLHD with respect to ϕ_p criterion while SA performed better than CP in terms of rate of convergence and the optimum values of optimality criteria. Na-udom [11] compared the efficiency of various types of LHD classes by considering prediction accuracy. The results revealed that SLHD performed better than other classes of design for small dimension problem. For larger dimension problem, the orthogonal Latin hypercube designs generated by William transformation performed the best.

According to the results that have been published in the previous studies, we observe that RLHD, SLHD and OALHD provide both of advantage and disadvantage. Hence, it is crucial to explore the performance of RLHD, SLHD and OALHD when various types of problems are considered. In this paper, we investigate the design property of these three classes of LHD and compare the prediction accuracy of statistical models when using the optimal designs obtained from these three classes. The design property is validated through the popular optimality criteria such as ϕ_p criterion (Morris

and Mitchell [7]), maximin distance criterion [12], the mean of correlation coefficient between design columns. The prediction accuracy is implemented by using Response Surface Methodology (RSM) [13] and KRG [14] with respect to the root mean square error (RMSE).

2. Methodology

In this section, we present the details of three classes of designs, followed by the steps of SA for finding the optimal design in the class. After the optimal design from each dimension is obtained, the statistical modeling methods will be used to implement the prediction accuracy of statistical models when three different classes of design are used. All simulation studies discussed here were implemented in R program version 2.10.1.

2.1 Design Used

2.1.1 Random Latin hypercube design (RLHD)

The RLHD was originally proposed by McKay et al. [1] and received wide attention from the practitioners in the context of CSE. RLHD is a matrix $\mathbf{X}$ which contains n rows d columns, where n is the number of runs, and d is the number of input variables, and be denoted by $RLHD(n, d)$. The total numbers of possible RLHD are $(n!)^d$. Generally the RLHD can be constructed based on the idea of stratified sampling to ensure that all sub-regions in the divided input variable space will be sampled with equal probability. A LHD can be generated from

$$X_{ij} = \frac{\pi_{ij} - U_{ij}}{n} \quad (1)$$

where $\pi_{i1}, \pi_{i2}, \dots, \pi_{id}$ are independent random permutation of $\{1, 2, \dots, n\}$ and U_{ij} are $n \times d$ values of *i.i.d.* uniform $U(0, 1)$ random variables independent of the π_{ij} ($i = 1, 2, \dots, n; j = 1, 2, \dots, d$).

In practice, RLHD can be easily generated by random permutation of each column which contains $\{1, 2, \dots, n\}$. Thus the d columns are combined together to form the design matrix $\mathbf{X}$. For instance, in the case of $RLHD(9, 2)$ design is

considered, this means that $n = 9$ and $d = 2$, the total number of possible RLHD is $(9!)^2 \approx 1.3168 \times 10^{11}$. Thus searching for the optimal design in the class is time consuming. In practice the range of each input variable is scaled into unit interval. Ye et al. [5] recommended the transformation as

$$\{0, 1/(n-1), 2/(n-1), \dots, 1\} \quad (2)$$

The example of $RLHD(9, 2)$ are visualized in Figure 1.

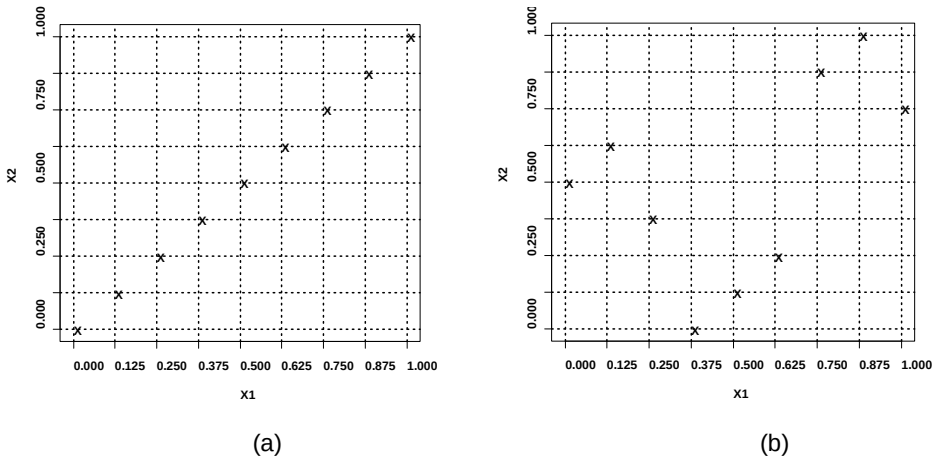


Figure 1. The example of $RLHD(9, 2)$.

From the Figure 1, it should be noted that RLHD represents a very bad design space on two dimensions. Hence the optimal RLHD could be generated by using search algorithm under pre-specified optimality criteria.

2.1.2 Symmetric Latin hypercube design (SLHD)

In this section we adapt the approaches of Ye et al. [5] and Na-udom [15] to study the construction of SLHD which is a special case of LHD and can be denoted by $SLHD(n, d)$. Any LHD is called a SLHD if it has a reflection property. The SLHD comprises of $n \times d$ LHD with the levels $\{1, 2, \dots, n\}$. Since SLHD has reflection of runs in the design, thus if one of runs have been selected as a design point, then another design point that reflects this design point would be selected into the design matrix as well. For instance, if $(a_1, a_2, \dots, a_d)$ is one of the rows in design matrix, then the vector

$(n+1-a_1, n+1-a_2, \dots, n+1-a_d)$ must be another row in design matrix. The construction of SLHD can be divided into two cases, where n is even number and n is odd number, respectively. The steps of generating each case of design runs are quite similar, except for the case of n is odd number, the center point of the design $(\frac{n+1}{2})$ does not play any role in the exchange of design points in design matrix, e.g. the construction of $SLHD(9, 2)$, when the first pair of element in the design is randomly generated, if the design point $(a_1, a_2) = (1, 2)$ is the first element of any row, then $(n+1-a_1, n+1-a_2) = (9+1-1, 9+1-2) = (9, 8)$ will be another row in design matrix which it is the ninth rows in this case. The construction of other rows can be made by using the same approach as mentioned previously, in this case the fifth rows with the element $(5, 5)$ does not have the reflection point. The example of $SLHD(9, 2)$ is shown in Figure 2(a) and the consequence design which the range of input variables are scaled into $(0, 1)^2$ is also visualized in Figure 2(b). The spread of design points for $SLHD(9, 2)$ is presented in Figure 3. It can be clearly seen from Figure 3 that the design points spread well over the design space. Hence SLHD seems to have a good space filling property.

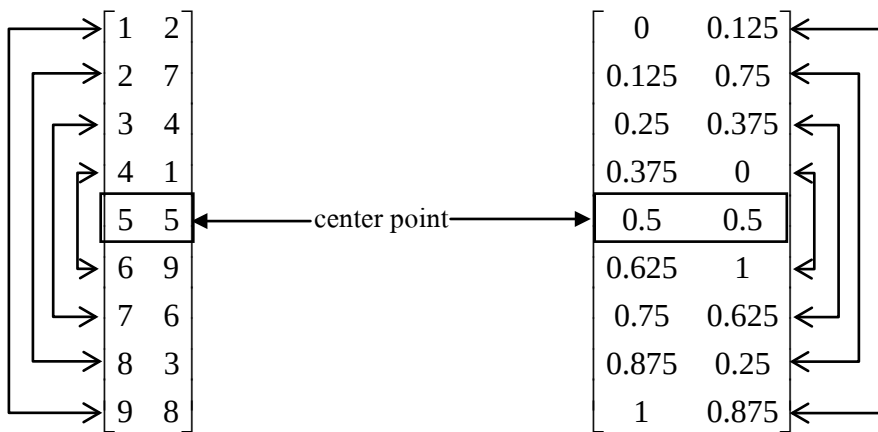


Figure 2. The design matrix of $SLHD(9, 2)$; (a) real range (b) unit interval.

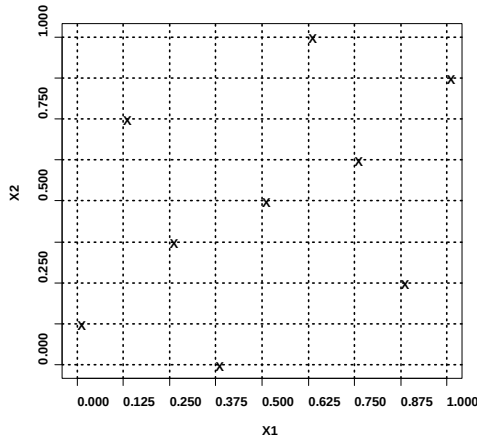


Figure 3. The scatter diagram for $SLHD(9, 2)$.

2.1.3 Orthogonal Array-based Latin hypercube design (OALHD)

In this section, we present the construction of OALHD which adapted from the approach of Tang [2], Leary et al. [10] and Fang et al. [16]. An $n \times d$ matrix $\mathbf{A}$ is an orthogonal array of strength r , where $2 \leq r \leq d$, comprises of n runs, d input variables, and q levels. If $n \times r$ is a sub-matrix of $\mathbf{A}$, which contains all possible $1 \times r$ rows with the same frequency λ , where λ represents the index of an array which $\lambda = n / q^r$. The OALHD is denoted by $OA(n, d, q, r)$.

The construction of an OALHD can be performed by generating the matrix $\mathbf{A}$. For each column of $\mathbf{A}$, the λq^{r-1} positions with entry k will be replaced by the permutation of

$$\{(k-1)\lambda q^{r-1} + 1, (k-1)\lambda q^{r-1} + 2, \dots, (k-1)\lambda q^{r-1} + \lambda q^{r-1} = k\lambda q^{r-1}\} \quad (3)$$

for all $k = 1, 2, \dots, q$.

After every columns of $\mathbf{A}$ is fully replaced by using equation (3), a new design matrix $\mathbf{A}$ will be the LHD class. The OALHD can be constructed if and only if the random orthogonal arrays are available; hence in this study we limit the dimension of problem by considering the availability of a random orthogonal array proposed by Tang [2]. For the sake of completeness, we present the steps of constructing an OALHD as follows.

Step 1: Specify dimension of problem, e.g. $OA(8, 2, 2, 2)$.

Step 2: Construct the random orthogonal arrays (matrix **A**), and find the λ value ($\lambda = 8 / 2^2 = 2$). The design matrix of random orthogonal arrays is shown in Figure 4(a).

Step 3: Create the OALHD. Since $\lambda = 2$, then the $\lambda q^{r-1} = (2)(2)^{2-1} = 4$ position in each value of k ($k=1, 2$) must be randomly replaced by the elements computed from equation (3). If $k=1$, the position with the element equal to 1 must be replaced by the permutation of $\{1, 2, 3, 4\}$. If $k=2$, the position with element equal to 2 must be replaced by the permutation of $\{5, 6, 7, 8\}$. Thus, the new design matrix is an $OA(8, 2, 2, 2)$ and presented in Figure 4(b). The design matrix in the scaled format is also shown in Figure 4(c). The spread of design points is visualized in Figure 5.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 2 & 1 \\ 2 & 1 \\ 2 & 2 \\ 2 & 2 \end{bmatrix}$$

(a)

$$\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 6 \\ 4 & 5 \\ 8 & 4 \\ 6 & 3 \\ 5 & 8 \\ 7 & 7 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 0 & 0.143 \\ 0.286 & 0 \\ 0.143 & 0.714 \\ 0.429 & 0.571 \\ 1 & 0.429 \\ 0.714 & 0.286 \\ 0.571 & 1 \\ 0.857 & 0.857 \end{bmatrix}$$

(c)

Figure 4. Design matrix: (a) random orthogonal arrays (b) $OA(8, 2, 2, 2)$ with real range (c) $OA(8, 2, 2, 2)$ with unit range.

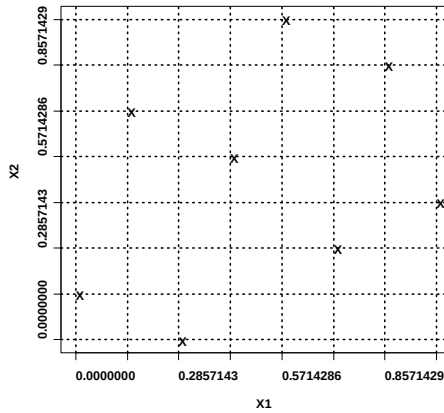


Figure 5. The spread of design points for $OA(8, 2, 2, 2)$.

Clearly, the spread of design points presented in Figure 3 and Figure 5 show better space filling property than the design presented in Figure 1. This indicates that SLHD and OALHD classes are superior over RLHD class of design in terms of space filling property.

2.2 Search Algorithms

As we already mentioned that the optimal LHD can be constructed by using the search algorithm under a pre-specified optimality criterion. In this paper, we adopt a version of SA proposed by Morris and Mitchell [7] to construct the optimal designs with respect to ϕ_p criteria. For each design class, SA was repeated for 10 times to vary the starting point. The prediction accuracy values are implemented based on these 10 optimal designs. The steps and parameter setting of SA can be found in Morris and Mitchell [7], Leary et al. [10], and Rungrattanaubol and Na-udom [9].

2.3 Optimality Criteria

For a given dimension of problem, the optimality criteria are used to consider the goodness of designs conducted from three different classes of LHD. In this study we consider ϕ_p criterion, maximin distance criterion, and the mean of correlation coefficient between design columns, respectively. The details of these criteria are presented as follows.

2.3.1 Maximin Distance Criterion

Maximin distance criterion was developed by Johnson et al. [12]. Any design $\mathbf{X}$ is called a maximin design if it maximizes the minimum intersite distance:

$$\text{maximin} = \min_{1 \leq i, j \leq n} d(X_i, X_j); i \neq j \quad (4)$$

where $d(X_i, X_j)$ is the Euclidean distance between i^{th} and j^{th} design points:

$$d(X_i, X_j) = \left[\sum_{k=1}^d (X_{ik} - X_{jk})^2 \right]^{1/2} \quad (5)$$

This criterion guarantees that the design points are not close to each other. In this study, each class of design is constructed for 10 times to vary the starting designs point. Once the optimal design from each class is constructed, the Euclidean distance matrix is calculated using equation (5) and the maximin value as stated in equation (4) is obtained. After this process is repeated for 10 times, the average of maximin value is calculated. Any class of design which maximizes the average of maximum value is considered as the best class of design for a specific dimension of problem.

2.3.2 ϕ_p Criterion

The ϕ_p criterion, an extension of maximin distance criterion, was proposed by Morris and Mitchell [7]. This criterion can be calculated as

$$\phi_p = \left[\sum_{j=1}^m J_j d_j^{-p} \right]^{1/p} \quad (6)$$

where p is a positive integer, J_j is index list $(J_1, J_2, \dots, J_m)$ which J_j is the number of pairs of runs in the design separated by distance d_j , d_j is distance list

$(d_1, d_2, \dots, d_m)$ with $d_1 < d_2 < \dots < d_m$, and m is the value between 1 and $\binom{n}{2}$. In

this study, we use the adaptive form of ϕ_p [17], which is simpler than equation (4). This equation can be expressed as

$$\phi_p = \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{1}{d_{ij}^p} \right]^{1/p} \quad (7)$$

where d_{ij} is Euclidean distance, $d(x_i, x_j) = d_{ij} = \left[\sum_{l=1}^d (x_i^{(l)} - x_j^{(l)})^2 \right]^{1/2}$, of the element in i^{th} and j^{th} runs, and l is the number of input variables ($l = 1, 2, \dots, d$).

2.3.3 Mean of correlation coefficient between design columns

The mean of correlation between design columns is the mean of Pearson correlation coefficient [18] which used to measure relationship between any pairs of input variables, can be calculated from

$$r_{ij} = \frac{\sum_{u=1}^n (x_{ui} - \bar{x}_i)(x_{uj} - \bar{x}_j)}{\sqrt{\sum_{u=1}^n (x_{ui} - \bar{x}_i)^2} \sqrt{\sum_{u=1}^n (x_{uj} - \bar{x}_j)^2}} \quad (8)$$

where r_{ij} is Pearson correlation coefficient between i^{th} and j^{th} design columns.

Once we obtain all $\binom{d}{2}$ values of the correlation coefficient between the

design columns, we then calculate the mean of these values. The class of design which maximizes the mean of correlation coefficient is considered as the best class of design.

2.4 Statistical Models

2.4.1 Response Surface Methodology (RSM)

RSM has received wide attention in the statistical modeling method. The second-order polynomial model has been extensively used. All unknown parameters can be estimated by using the method of least squares [19]. RSM is based on assumption of random error that arising from a large number of insignificant input variables that are discarded from statistical modeling method. For a given output variables y , and input variables $x = (x_1, x_2, \dots, x_d)$, the relationship between y and x can be written as

$$y = f(x) + \varepsilon \quad (9)$$

where ε is random error that is assumed to be normal distribution with mean zero and constant variance σ^2 . Since the true response surface function ($f(x)$) is unknown, the response surface $g(x)$ is constructed to approximate $f(x)$. Thus, the prediction values are obtained by using $\hat{y} = g(x) + \varepsilon$, which is the polynomial function of $(X_1, X_2, \dots, X_d)$. Simpson et al. [13] and Fang and Horstemeyer [20] rewrite the function $g(x)$ as

$$\hat{y}(x) = \beta_0 + \sum_{i=1}^d \beta_i x_i + \sum_{i=1}^d \beta_{ii} x_i^2 + \sum_{i=1}^{d-1} \sum_{i < j}^d \beta_{ij} x_i x_j \quad (10)$$

where $\beta_i, \beta_{ii} (i = 1, 2, \dots, d)$ and $\beta_{ij} (i < j = 1, 2, \dots, d)$ are unknown parameters.

The function in equation (10) can be written in terms of the observed value as

$$y_i = \beta_0 + \sum_{j=1}^d \beta_{ij} x_{ij} + \sum_{j=1}^d \beta_{jj} x_{ij}^2 + \sum_{j=1}^{d-1} \sum_{j < k}^d \beta_{jk} x_{ij} x_{ik} + \varepsilon_i \quad (11)$$

Equation (11) can be expressed in matrix form, $\mathbf{X}$ as

$$\mathbf{y}_0 = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (12)$$

$$\text{where } \mathbf{y}_0 = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_d \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

The approximation of regression coefficients, $\boldsymbol{\beta}$, in equation (12) can be obtained by using the method of least squares. This method is based on the minimization of

$$L = \sum_{i=1}^n \varepsilon_i^2 = (\mathbf{y}_0 - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y}_0 - \mathbf{X}\boldsymbol{\beta}) \quad (13)$$

The equation (13) can be simplified to

$$\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{y}_0 \quad (14)$$

Hence, the least squares estimator of $\boldsymbol{\beta}$ is

$$\hat{\beta} = \left(X^T X \right)^{-1} X^T y_0 \quad (15)$$

2.4.2 Kriging Model (KRG)

Kriging model (KRG), in the context of CSE, was proposed by Sacks et al. [14]. The functional form of this model can be written as

$$y = \sum_{j=1}^k \beta_j f_j(\mathbf{x}) + Z(\mathbf{x}) \quad (16)$$

where β_j ($j = 1, 2, \dots, d$) is the parameter for polynomial function of input variables, $f_j(\mathbf{x})$ is a polynomial function of input variables ($j = 1, 2, \dots, d$), and $Z(\mathbf{x})$ is a realization of stochastic process with zero mean and some forms of correlation function. In practice, the polynomial function in equation (16) is considered as a constant [14,21], the subsequent equation can be expressed by

$$y = \beta + Z(x) \quad (17)$$

The second part on the right of equation (16) and (17), $Z(x)$ can be considered as a Gaussian correlation function. The covariance between $Z(x_i)$ and $Z(x_j)$ can be written as

$$\text{Cov}[Z(x_i), Z(x_j)] = \sigma^2 R(x_i, x_j) \quad (18)$$

where σ^2 is the process variance and $R(x_i, x_j)$ is the correlation function. The most frequently used form can be written as

$$R(X_i, X_j) = \prod_{j=1}^d \exp\left(-\theta_j |X_i, X_j|^p\right) \quad (19)$$

where $0 \leq p \leq 2$ and $\theta_j > 0$.

KRG is normally fitted using the idea of generalized least squares method. The problem of estimating all unknown parameters is reduced to the estimation of the parameter of the correlation function. The method of maximum likelihood estimation (MLE) proposed by Welch et al. [21] are widely used. The maximum likelihood estimators can be obtained by maximizing the log likelihood function

$$l(\boldsymbol{\beta}, \sigma^2, \mathbf{R} | \mathbf{y}_0) = -\frac{1}{2} \left[n \ln \sigma^2 + \ln |\mathbf{R}| + (\mathbf{y}_0 - \mathbf{1}\boldsymbol{\beta})^T \mathbf{R}^{-1} \right] \quad (20)$$

where $\mathbf{1}$ is a column vector of length d which all elements are one.

Given the correlation parameters θ and p in equation (19), the generalized least squares estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1})^{-1} \mathbf{1}^T \mathbf{R}^{-1} \mathbf{y}_0 \quad (21)$$

and the MLE of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n} (\mathbf{y}_0 - \mathbf{1}\hat{\boldsymbol{\beta}})^T \mathbf{R}^{-1} (\mathbf{y}_0 - \mathbf{1}\hat{\boldsymbol{\beta}}) \quad (22)$$

Substituting $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ into equation (20), the problem is to numerically maximize

$$-\frac{1}{2} (n \ln \hat{\sigma}^2 + \ln |\mathbf{R}|) \quad (23)$$

After all unknown parameter are estimated, the next step is to construct a predictor, $\hat{y}(x)$, at prediction point x . This predictor can be written as

$$\hat{y}(x) = \hat{\boldsymbol{\beta}} + \mathbf{r}^T(x) \mathbf{R}^{-1} (\mathbf{y}_0 - \mathbf{1}\hat{\boldsymbol{\beta}}) \quad (24)$$

where $\mathbf{r}^T(x)$ is the $1 \times n$ vector of correlation function between error $Z(x)$ and prediction point x . The correlation vector $\mathbf{r}^T(x)$ can be computed from

$$\hat{y}(x) = \hat{\boldsymbol{\beta}} + \mathbf{r}^T(x) \mathbf{R}^{-1} (\mathbf{y}_0 - \mathbf{1}\hat{\boldsymbol{\beta}}) \quad (25)$$

KRG has received wide attention in many applications of CSE due to its interpolation property [14].

2.5 Test Problems

In order to investigate the relation of design property and prediction accuracy, we fit RSM and KRG by using the obtained optimal design for a specified dimension of problem as described in Section 2.1. Five test problems are employed and the prediction accuracy is validated through root mean square error (RMSE). A range of test problems are included in this study, the complex test problem are Welch function [21], Cyclone

model [5], and Borehole function [22], while the 3D function [23] and 10D function [24] are non-complex test problems. In order to validate the prediction accuracy of RSM and KRG models through RMSE values, the additional test points are used in this study. For two-dimensional test problem, the 81 test points are constructed using the grid points over the design space. For larger dimension of problems the 500 random test points are selected to validate the prediction accuracy. All test problems along with their dimensions included in this study are summarized in Table 1.

Table 1. The details of test problems.

d	n	Designs	Functions
2	9	RLHD SLHD OALHD	$f(x_1, x_2) = [30 + x_1 \sin(x_1)](4 + e^{-x_2}); 0 \leq x_1, x_2 \leq 5$
3	16	RLHD SLHD OALHD	$f(x_1, x_2, x_3) = (x_1 + x_2)^2 + (x_2 + x_3)^2; -10 \leq x_1, x_2, x_3 \leq 10$
7	49	RLHD SLHD OALHD	$f(x_1, x_2, \dots, x_7) = 174.42 \left(\frac{x_1}{x_5} \right) \left(\frac{x_3}{x_2 - x_1} \right)^{0.85} \times$ $\sqrt{\frac{1 - 2.62 \left\{ 1 - 0.36 (x_4 / x_2)^{-0.56} \right\}^{3/2} (x_4 / x_2)^{1.16}}{x_6 x_7}}$; $0.09 \leq x_1 \leq 0.11, 0.27 \leq x_2 \leq 0.33, 0.09 \leq x_3 \leq 0.11,$ $0.09 \leq x_4 \leq 0.11, 1.35 \leq x_5 \leq 1.65, 14.4 \leq x_6 \leq 17.6,$ $0.675 \leq x_7 \leq 0.825$

d	n	Designs	Functions
8	49	RLHD SLHD OALHD	$f(x_1, x_2, \dots, x_8) = \frac{2\pi x_3 (x_4 - x_6)}{\ln\left(\frac{x_2}{x_1}\right) \left\{ 1 + \left[\frac{2x_7 x_3}{\ln\left(\frac{x_2}{x_1}\right) x_1^2 x_8} \right] + \frac{x_3}{x_5} \right\}}$ $; 0.05 \leq x_1 \leq 0.15, 100 \leq x_2 \leq 50000, 63070 \leq x_3 \leq 115600,$ $990 \leq x_4 \leq 1110, 63.1 \leq x_5 \leq 116, 700 \leq x_6 \leq 820,$ $1120 \leq x_7 \leq 1680, 9855 \leq x_8 \leq 12045$
10	81	RLHD SLHD OALHD	$f(x_1, x_2, \dots, x_{10}) = \sum_{j=1}^{10} \left[\frac{3}{10} + \sin\left(\frac{16}{15} x_j - 1\right) + \sin^2\left(\frac{16}{15} x_j - 1\right) \right]$ $; -1 \leq x_1, x_2, \dots, x_{10} \leq 1$

The RMSE values is calculated from

$$RMSE = \sqrt{\frac{\sum_{i=1}^k (y_i - \hat{y}_i)^2}{k}} \quad (26)$$

where k is the number of test points (specified in Table 1), y_i is the actual response from the i^{th} test point and $\hat{y}_i$ is the predicted value by using RSM or KRG for the i^{th} test point.

3. Result

In this section we present the design property of the optimal designs obtained from three classes of design. The design property considered here are ϕ_p criterion, maximin distance criterion, and the mean of correlation between columns in the design matrix. All results are presented in Table 2.

Table 2. Design property of three classes of LHD.

d	n	Design	ϕ_p	Maximin	Correlation
2	9	RLHD	4.3011 (0.8284)	0.3786 (5.6921)	0.0167 (210.8185)
		SLHD	4.7771 (6.0697)	0.3133 (14.4452)	0.1733 (253.3147)
		OALHD	4.3011 (0.8284)	0.3786 (5.6921)	0 (-)
3	16	RLHD	4.3875 (0.6897)	0.4219 (3.5648)	0.0079 (416.7345)
		SLHD	4.4948 (1.2406)	0.4051 (4.7570)	0.0216 (374.4601)
		OALHD	4.3749 (0.5118)	0.4229 (4.2482)	-0.0019 (-1405.7984)
7	49	RLHD	4.1250 (0.1871)	0.7381 (2.0447)	0.0017 (257.8105)
		SLHD	4.1622 (0.1559)	0.7412 (2.1486)	-0.0019 (-403.8402)
		OALHD	4.1236 (0.1276)	0.7413 (1.4894)	0.0013 (339.7717)
8	49	RLHD	3.7721 (0.1037)	0.8264 (1.7346)	0.0010 (339.0918)
		SLHD	3.8243 (0.0600)	0.8188 (0.7805)	-0.0038 (-200.9162)
		OALHD	3.7714 (0.1375)	0.8302 (2.3622)	-0.0002 (-1702.5178)
10	81	RLHD	4.1368 (0.0601)	0.9092 (2.0332)	0.0002 (883.9840)
		SLHD	4.1755 (0.0616)	0.9113 (0.3999)	0.0025 (146.1437)
		OALHD	4.1363 (0.0530)	0.9116 (1.9795)	0.0001 (1579.4033)

The results in Table 2 indicate that OALHD shows a better design property than RLHD and SLHD for all dimensions of problem since OALHD has a minimum mean of ϕ_p criterion, while the mean value of maximin distance criterion for OALHD is larger than those obtained from RLHD and SLHD. OALHD also has a minimum average value of the mean of correlation between design columns, which indicates the orthogonal between columns in the design matrix. Hence it could be concluded that OALHD is the best design class when comparing with RLHD and SLHD for 3-10 dimensional problems. In the case of two dimensional problems, it should be noted that RLHD and OALHD have the same design property when ϕ_p criterion and maximin distance criterion are considered. Further, OALHD is the most consistence class of design since the coefficient of variation (C.V.) obtained from ten different starting points is very small.

After considering the design properties of three different design classes, we used them to construct the statistical models. For each test problem, 10 LHDs from each

class of design are used to get RMSE values and the descriptive statistics are presented in Table 3. The box plot of RMSE generated from each statistical modeling method is presented in Figure 6 to Figure 10.

Table 3. RMSE values from RSM and KRG for all test problem.

Test problems	Designs	RSM			KRG		
		Mean	S.D.	C.V. (%)	Mean	S.D.	C.V. (%)
Welch Function	RLHD	3.3092	0.1114	3.3661	3.8744	0.4972	12.8320
	SLHD	3.6703	0.3124	8.5123	4.5583	0.6699	14.6973
	OALHD	3.3081	0.0898	2.7140	3.7331	0.5079	13.6061
3D Function	RLHD	115.655	5.8368	5.0467	28.6686	1.9842	6.9213
	SLHD	109.3415	6.6224	6.0566	28.8497	8.7014	30.1612
	OALHD	97.1094	4.6363	4.7743	24.9280	1.4909	5.9807
Cyclone model	RLHD	0.0494	0.0170	34.4645	0.0534	0.0150	28.1537
	SLHD	0.0527	0.0165	31.2989	0.0569	0.0157	27.6163
	OALHD	0.0487	0.0112	23.0076	0.0537	0.0097	18.1185
Borehole function	RLHD	201.6870	224.8513	111.4853	15.9201	0.5328	3.3465
	SLHD	46.4112	52.5110	113.1428	16.0270	0.4245	2.6486
	OALHD	34.4127	58.8035	170.8773	14.9726	0.6337	4.2327
10D function	RLHD	0.3207	0.1442	44.9729	0.2645	0.0058	2.2078
	SLHD	0.3057	0.0082	2.6869	0.2830	0.0101	3.5619
	OALHD	0.3030	0.1507	49.7409	0.2496	0.0065	2.5923

It can be clearly seen from the Table 3 that the statistical models constructed from OALHD performs best in terms of prediction accuracy, as it provides lower RMSE values on average, when both of RSM and KRG are used. It is, however, when Cyclone model ($d = 7$) is considered, the KRG constructed from RLHD performs slightly better than OALHD, but the RMSE values obtained from these two design classes are still close to each other. It should be noted that RSM and KRG created from SLHD perform worst among three design classes when Welch function and 10D test problems are considered. SLHD, however, is superior to RLHD when 3D and 10D with RSM test problems are considered. In the case of prediction capability between RSM and Kriging is compared, we observe that Kriging model performs much better than RSM for Borehole and 10D functions. This indicates that Kriging model along with optimal LHD is the best choice to use for modeling response from CSE. RSM model, however, performs not far from Kriging model when less complex test functions are employed (Cyclone

function and 10D). According to the C.V. of RMSE values, it could be concluded that the prediction accuracy of RSM and KRG models conducted from OALHD are quite stable as the C.V. values obtained from this design are small.

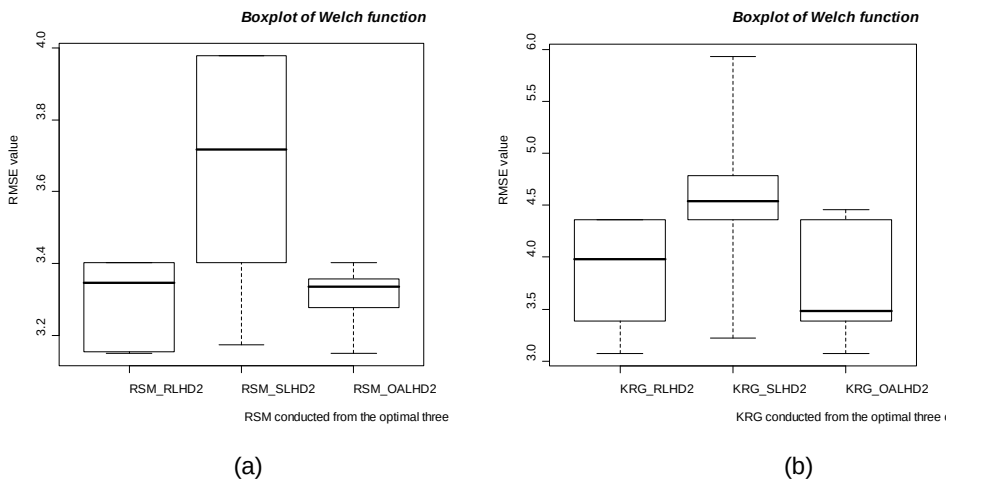


Figure 6. Box plot of RMSE values for Welch function; (a) RSM (b) KRG.

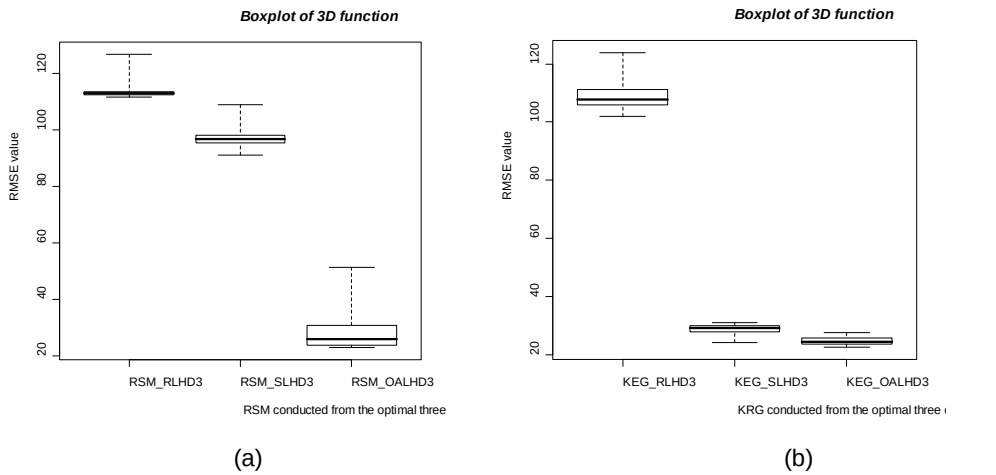
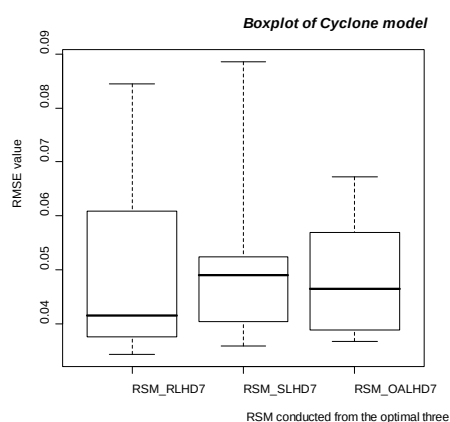
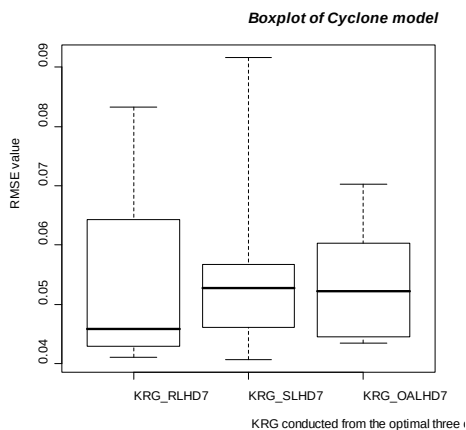


Figure 7. Box plot of RMSE values for 3D function; (a) RSM (b) KRG.

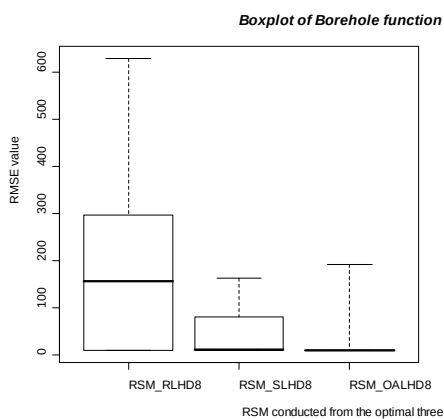


(a)

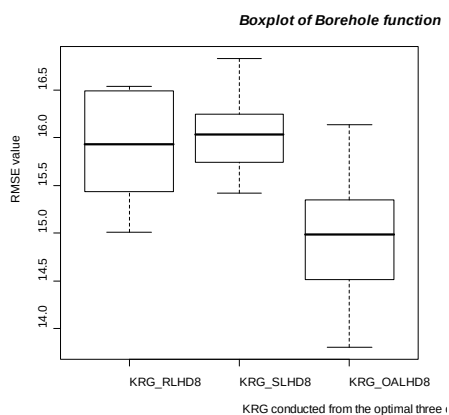


(b)

Figure 8. Box plot of RMSE values for Cyclone model; (a) RSM (b) KRG.



(a)



(b)

Figure 9. Box plot of RMSE values for Borehole function; (a) RSM (b) KRG.

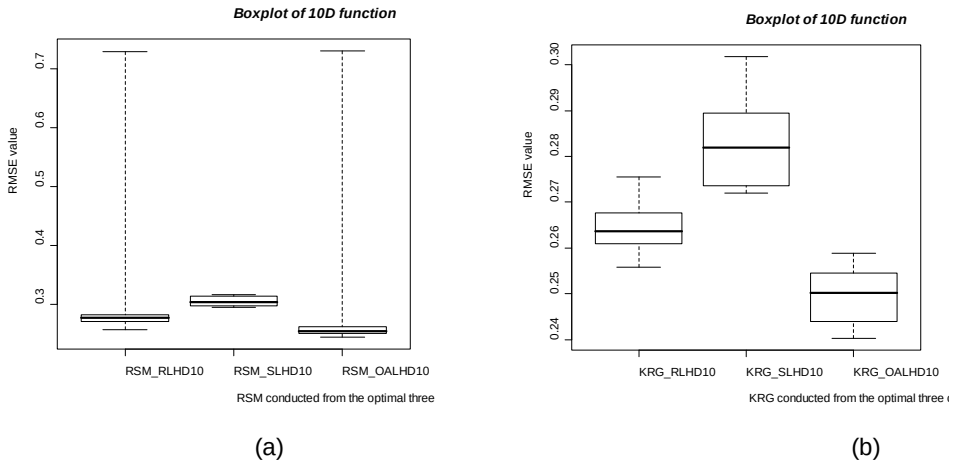


Figure 10. Box plot of RMSE values for 10D function; (a) RSM (b) KRG.

4. Conclusions

The construction of optimal design based on three different classes of Latin hypercube design has been studied. According to the design property results presented in the previous section, it could be concluded that OALHD is the best design for most of the dimensional problems with respect to the optimality criteria under consideration. It should be noted that for two dimensional problems, both of RLHD and OALHD perform best with respect to ϕ_p and maximin distance criterion. When the prediction accuracy based on RMSE values is considered, OALHD with RSM model performs best for all dimension problems, while OALHD with KRG model is the most accurate approach for Welch function, 3D function, Borehole function, and 10D function. It is, however, RLHD with Kriging models can perform well when Cyclone model is validated. Further, SLHD with RSM can perform as well as OALHD for small dimensional problem.

As already presented in the previous section, OALHD should be recommended as the best design choice for designing and modeling the response in the context of CSE. Since the dimension of problem is assigned with respect to the validity of the random orthogonal arrays, hence SLHD can be used as an alternative of OALHD whenever the random orthogonal array does not exist. As we restrict our approach to three classes of Latin hypercube design only, hence other methods should be further investigated.

Acknowledgements

This research is supported by the Centre of Excellence in Mathematics, the Commission on Higher Education, Thailand.

References

- [1] 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*, 1979; 21: 239-245.
- [2] Tang, B., Orthogonal Array-Based Latin Hypercubes, *Journal of American Statistical Association*, 1993; 88: 1392-1397.
- [3] Ye, K.Q., Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments, *Journal of American Statistical Association*, 1998; 93: 1430-1439.
- [4] Butler, N.A., Optimal and Orthogonal Latin Hypercube Designs for Computer Experiments, *Biometrika*, 2001; 88: 847-857.
- [5] Ye, K.Q., Li, W., and Sudjianto, A., Algorithmic construction of optimal symmetric Latin hypercube designs, *Journal of Statistical Planning and Inference*, 2000; 90: 145-159.
- [6] Park, J.S., Optimal latin hypercube designs for computer experiments, *Journal of Statistical Planning and Inference*, 1994; 39: 95-111.
- [7] Morris, M.D., and Mitchell, T.J., Exploratory designs for computational experiments, *Journal of Statistical Planning and Inference*, 1995; 43: 381-402.
- [8] Li, W., and Wu, C.F.J., Columnwise-pairwise algorithms with applications to the construction of supersaturated designs, *Technometrics*, 1997; 39: 171-179.
- [9] Rungrattanaubol, J., and Na-udom, A., Comparison of Evolutionary Search Algorithms in Computer Simulated Experiments, *Proceedings of the National Computer Science and Engineering Conference*, 2007; 102-106.
- [10] Leary, S., Bhaskar, A., and Keane, A., Optimal Orthogonal Array-Based Latin Hypercubes, *Journal of Applied Statistics*, 2003; 30: 585-598.
- [11] Na-udom, A., Comparison of Design for Computer Simulated Experiments, *Proceedings of the Annual meeting on Statistics*, 2007; 148-152.

- [12] Johnson, M.E., Moore, I.M., and Ylvisaker, D., Minimax and Maximin distance designs, *Journal of Statistical Planning and Inference*, 1990; 26: 131-148.
- [13] Simpson, T.W., Lin, D.K.J., and Chen, W., Sampling Strategies for Computer Experiments: Design and Analysis, *International Journal of Reliability and Applications*, 2001; 2: 209-240.
- [14] Sacks, J., Welch, W.J., Mitchell, T.J., and Wynn, H.P., Design and analysis of computer experiments, *Statistical Science*, 1989; 4: 409-435.
- [15] Na-udom, A., Review of Design and Analysis of Computer Simulated Experiments, *KKU Science Journal*, 2008; 36: 106-120.
- [16] Fang, K.T., Li, R., and Sudjianto, A., Design and Modeling for computer experiment, America, Chapman & Hall/CRC.
- [17] Jin, R., Chen, W., and Sudjianto, A., An efficient algorithm for constructing optimal design of computer experiments, *Journal of Statistical Planning and Inference*, 2005; 134: 268-287.
- [18] Montgomery, D.C., Peck, E.A., and Vinning, G.G., Introduction to Linear Regression Analysis, 4th ed, New Jersey, Wiley, 2006.
- [19] Myers, R.H., Montgomery, D.C., and Anderson-Cook, C.M., Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 3rd ed, United States of America, Wiley, 2009.
- [20] Fang, H., and Horstemeyer, M.F., Global Response Approximation with Radial Basis Functions, *Engineering Optimization*, 2006; 38: 407-424.
- [21] Welch, W.J., Buck, R.J., Sacks, J., Wynn, H.P., Mitchell, T.J., and Morris, M.D., Screening, Predicting, and Computer Experiments, *Technometrics*, 1992; 34: 15-25.
- [22] Palmer, K., and Tsui, K.L., A minimum bias latin hypercube design, *IIE Transactions*, 2001; 33: 793-808.
- [23] Hock, W.F., and Schittkowski, K., Test examples for nonlinear programming codes, New York, Springer, 1981.
- [24] Muller, A.A., and Messac, A., Metamodeling using Extended Radial Basis Functions: A Comparative Approach, *Engineering with Computers*, 2006; 21: 203-217.