



Thailand Statistician
January 2017; 15(1): 17-26
<http://statassoc.or.th>
Contributed paper

Orthogonal Space Filling Design for Computer Simulated Experiments

Wiratchamas Sukdaiphueng [a], Jaratsri Rungrattanaubol [a] and Anamai Na-udom*[b]

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

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

*Corresponding author; e-mail: anamain@nu.ac.th

Received: 17 August 2015

Accepted: 16 May 2016

Abstract

At present, computer simulated experiment (CSE) has been widely used in many engineering applications, in order to explore the complex phenomenon that is hard or unable to perform physical experiment. Hence, a relationship between input variables and output response in these complex systems can be explored and constructed as a metamodel. CSE, in nature, is deterministic and time-consuming when running for a single value of output response. Therefore, an experimental design for CSE plays an essential role and is based on developing a statistical predictive model with a handful of runs. Latin hyper cube design (LHD) is one of popular design class in this context. The design with spread filling property or orthogonal property has been proposed to use in CSE. This paper proposes a combined criterion namely Φ_p and correlation coefficient criterion called multi-objective criterion along with simulated annealing search algorithm to construct an optimal LHD design. The results show that the multi-objective criterion can produce optimal LHD designs with orthogonal and space filling properties. Moreover, the predictive model conducted from these designs has good performance in terms of prediction accuracy.

Keywords: Computer experiment, orthogonal Latin hypercube, simulated annealing algorithm.

1. Introduction

Computer simulated experiments (CSEs) have been extensively used in many applicable domains to explore and investigate the complex physical phenomena, especially when physical experiments are not feasible due to time constraint or the limitation of the experimental materials (Koehler and Owen 1996) such as a reservoir simulator to quantify the ultimate recovery of oil (Cheong 2005), a chemical reaction simulation (Sacks et al. 1989). CSE is deterministic in nature. It means that giving same values of input variables always yields with the same output response.

CSEs are time consuming and computationally expensive to run. The process of CSEs can be divided into three major parts; Experimental design, Simulation Routine (“black box”) and Approximation model construction (Simpson et al. 2001) as depicted in Figure 1.

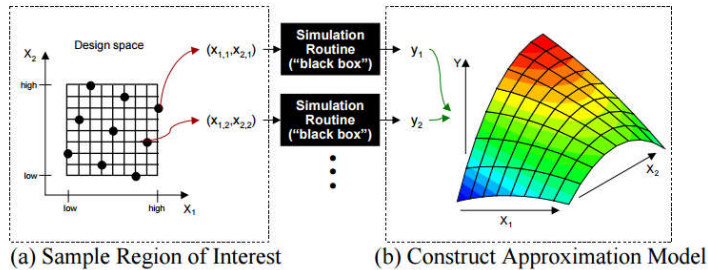


Figure 1 Design and analysis of CSEs (Simpson et al. 2001)

Experimental design focuses on efficiency by considering how to sample a design points from the design space by using fewer sample points or seeking better coverage of the design space. The good sample design points will affect the accuracy of the approximation model constructed based on fitting both linear and non-linear functions.

CSEs are usually complex and consist of many input variables and are required a large number of runs. Hence, the size of design space is relied on the dimension of problems and grows exponentially. The most popular class of space filling design in CSEs is Latin hypercube design (LHD) proposed by McKay et al. (1979). LHD is a matrix of n rows and d columns where n is the number of runs and d is the number of input variables. Each column is independent random permutation of number 1 through n as shown in Figure 2.

X_1	X_2	X_3
1	4	3
2	3	4
3	2	5
4	1	2
5	5	1

Figure 2 An example of 5×3 random LHD

With LHD class, the size of design space grows with $(n!)^d$. Therefore, search algorithm approach to construct optimal LHDs was proposed and many search algorithms have already been used and applied in CSE context such as Simulated Annealing (SA), Columnwise-pairwise (CP), Genetic algorithm (GA) and Enhanced Stochastic Evolutionary (ESE) (Morris and Mitchell 1995, Jin et al. 2005, Li and Wu 1997). Along with these algorithms, many optimality criteria have also been proposed and introduced such as entropy, maximin and ϕ_p criteria.

This paper aims to investigate a construction of a design with space filling property by using ϕ_p criterion, a design with orthogonal property by using correlation coefficient criterion and a design with space filling property and orthogonal property, namely orthogonal space filling design by combining these two criteria, called multi-objective criterion. Simulated Annealing algorithm is

selected for searching optimal designs along with these three criteria. Response surface methodology (RSM) and Kriging model (KRG) are statistical methodology for constructing approximation models. The accuracy of approximation models, for three selected functions representing ‘black box’, is measured with root mean square error (RMSE).

2. Research Methodology

2.1. Latin Hypercube Design (LHD)

LHD has been widely used in the context of CSE. It was originally proposed by McKay et al. (1979). LHD is a matrix (X) of n rows and d columns, where n is a number of runs and d is a number of input variables. LHD can be constructed based on the idea of stratified sampling to ensure that all subregions in the divided input variable space will be sampled with equally probability. A Latin hypercube sampling has

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

where π_{ij} are the elements of an $n \times d$ matrix comprising of columns π_j ($j=1,2,\dots,d$). Each column is independent random permutation of ($j=1,2,\dots,n$) and U_{ij} are values of independent $U[0,1]$ random variables independent of the π_{ij} . A new LHD can be obtained by exchanging two elements in a column as shown in Figure 3.

X_1	X_2	X_3		X_1	X_2	X_3
1	4	3	$\Rightarrow$	1	4	3
2	3	4		2	1	4
3	2	5		3	2	5
4	1	2		4	3	2
5	5	1		5	5	1

Figure 3 A new LHD by exchanging element 1 and 3 on column X_2

2.2. Simulated Annealing Search Algorithm (SA)

SA is based on the analogy between the simulation of the annealing of solids and the problem of solving large combinatorial optimization problems (Peter et al. 1987). Morris and Mitchell (1995) adapted SA to construct optimal LHD using ϕ_p optimality criterion. LHD minimizing ϕ_p value is reserved as the best optimal LHD in the search space. The entire process can be described as following steps

Step 1. Initialize $I_{\max}$, t_0 , C_t

Step 2. Set X = a random LHD, $X_{best} = X$, $t = t_0$

Step 3. Set $I = 1$, label = 0

Step 4. Construct X_{try} by performing randomly element-exchange in X

Step 5. If X is better than X_{try} then $X = X_{try}$, label = 1

else if acceptance probability $e^{-[difference\ value\ of\ criterion]/t}$ is met then

$X = X_{try}$, label = 1

Step 6. If $\mathbf{X}_{\text{try}}$ is better than $\mathbf{X}_{\text{best}}$ then $I = 1$, $X_{\text{best}} = X_{\text{try}}$
 else $I = I + 1$

Step 7. If $I < I_{\text{max}}$ then go to Step 4 else $t = t \times C_t$

Step 8. If label = 1 then go to Step 3 else STOP and report X_{best}

In this paper, I_{max} is set to 500, initial t_0 is a mean of distance matrix, C_t is 0.95 and “is better than” is substituted with different optimality criteria. In this paper, the less value means the better design.

2.3. Optimality criteria

For any dimensional LHD, the optimality criteria are used to measure the goodness of LHD, especially during SA search process. In this paper, we consider ϕ_p , correlation coefficient and multi-objective criterion.

2.3.1. ϕ_p criterion

The ϕ_p criterion, an extension of maximin distance criterion, was originally proposed by Morris and Mitchell (1995). The key concept is to find all distance between any two design points in the design. In this research, the adaptive form of ϕ_p (Jin et al. 2005) is used, since it is simpler and easier to implement as appeared in (2)

$$\phi_p = \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{1}{(d_{ij})^p} \right]^{\frac{1}{p}} \quad (2)$$

where the rectangle distance is used as shown in (3)

$$d(x_i, x_j) = \sum_{k=1}^d |x_{ik} - x_{jk}| \quad (3)$$

d_{ij} is a rectangle distance between x_i and x_j . In this paper, p is set to 5.

2.3.2. Correlation coefficient criterion

Correlation coefficient criterion in this paper is the mean of correlation between design columns, which is used to measure relationship between any pairs of input variables, can be calculated from (4).

$$\text{corr}(D) = \sum_{i=1}^{d-1} \sum_{j=i+1}^d \frac{|r_{ij}|}{\binom{d}{2}} \quad (4)$$

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

D is LHD and r_{ij} is Pearson correlation coefficient between i^{th} and j^{th} design columns. The minimized value of the mean of correlation is considered to be the better design.

2.3.3. Multi-Objective criteria

Joseph and Hung (2008) proposed the concept of multiple objectives by combining ϕ_p criterion with a concept of orthogonal design by using the mean of correlation coefficient as a criterion. Since these two criteria have different boundaries i.e. correlation coefficient value is [0, 1], while ϕ_p is varied and not in a range [0, 1], to perform multiple objectives we have to transform ϕ_p value into [0, 1] range. Roshan and Hung also proposed the upper ($\phi_p U$) and lower ($\phi_p L$) boundary of ϕ_p for each dimension design as described in (5), (6) and (7).

$$\phi_p L = \left[\binom{n}{2} \left(\frac{\lceil \bar{d} \rceil - \bar{d}}{\lfloor \bar{d} \rfloor^p} + \frac{\bar{d} - \lfloor \bar{d} \rfloor}{\lceil \bar{d} \rceil^p} \right) \right]^{\frac{1}{p}} \quad (5)$$

$$\bar{d} = \frac{(n+1)d}{3} \quad (6)$$

$$\phi_p U = \left[\sum_{i=1}^{n-1} \frac{(n-i)}{(i \times d)^p} \right]^{\frac{1}{p}} \quad (7)$$

From the equations (5)-(7), the boundary for each dimension design is shown in Table 1.

Table 1 Upper ($\phi_p U$) and lower ($\phi_p L$) boundary of ϕ_p

n	d	Upper $\phi_p U$	Lower $\phi_p L$
9	2	0.312310	0.7625394
19	3	0	0.5982345
33	4	0.077312	0.5035015
51	5	0.048227	0.4404510
73	6	0	0.3948318
99	7	0.023399	0.3599632
129	8	0.017516	0.3322554
163	9	0	0.3095899
201	10	0.010775	0.2906275

Multi-objective criterion in this paper is assigned with equal weight for each criterion as shown in (8)

$$\omega_1 \times (\phi_p(D) - \phi_p L) / (\phi_p U - \phi_p L) + \omega_2 \times \text{corr}(D) \quad (8)$$

where D is LHD and ω_1, ω_2 are positive value with a constraint that $\omega_1 + \omega_2 = 1$, in this research ω_1 is assigned to 0.5.

2.4. Statistical models

The performance of orthogonal space filling design constructed by the method described in the previous section can be validated by statistical models chosen in this paper are based on the previous work by (Timun et al. 2011, Yosboonruang et al. 2013).

2.4.1. Response surface methodology

RSM is the second-order polynomial model, which has been extensively used in many applications. All unknown parameters can be estimated by using the method of least squares (Myers

et al. 2009). RSM is based on assumption of random error, arising from a large number of insignificant input variables that are discarded from statistical modeling method. For a given output variable y , and input variables $x = (x_1, x_2, \dots, x_d)$ the relationship between y and x can be written as

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

where ε is random error, 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. (2001) and Fang and Horstemeyer (2006) rewriting the function $g(x)$ as

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

where β_i , β_{ii} ($i = 1, 2, \dots, d$) and β_{ij} ($i < j = 1, 2, \dots, d$) are unknown parameters. The function can be written in terms of the observed value as

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

which can be expressed in matrix form, $\mathbf{X}$ as

$$\mathbf{y}_0 = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (9)$$

where

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

The approximation of regression coefficients, $\boldsymbol{\beta}$, in (9) can be obtained by using the method of least squares. This method is based on the minimization of

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

which can be simplified to $\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{y}_0$. Hence, the least squares estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y}_0).$$

2.4.2. Kriging model

Kriging model (KRG), in the context of CSE, was proposed by Sacks et al. (1989). The functional form of this model can be written as

$$y_i = \sum_{j=1}^k \beta_j f_j(x) + Z(x) \quad (10)$$

where β_j ($j = 1, 2, \dots, d$) is the parameter for polynomial function of input variables, $f_j(x)$ is a polynomial function of input variables ($j = 1, 2, \dots, d$), and $Z(x)$ is a realization of stochastic process with zero mean and some forms of correlation function. In practice, the polynomial

function in (10) is considered as a constant (Sacks et al. 1989, Welch et al. 1992), the subsequent equation can be expressed by

$$y = \beta + Z(x) \quad (11)$$

The second part on the right of (10) and (11), $Z(x)$ can be considered as a Gaussian correlation function. The covariance between $Z(x_i)$ and $Z(x_j)$ can be written as

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

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

$$R(x_i, x_j) = \prod_{j=1}^d \exp(-\theta_j |x_i - x_j|^p) \quad (12)$$

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

KRG is normally fitted using an 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 maximum likelihood estimators (MLE) can be obtained by maximizing the log likelihood function (Welch et al. 1992)

$$l(\beta, \sigma^2, R | y_0) = -\frac{1}{2} [n \ln \sigma^2 + \ln |R| + (y_0 - l\beta)^T R^{-1}] \quad (13)$$

where l is a column vector of length d which all elements are one. Given the correlation parameters θ and p in (12), the generalized least squares estimate of β is

$$\hat{\beta} = (l^T R^{-1} l)^{-1} l^T R^{-1} y_0$$

and the MLE of $\hat{\sigma}^2$ is

$$\hat{\sigma}^2 = \frac{1}{n} (y_0 - l\hat{\beta})^T R^{-1} (y_0 - l\hat{\beta}).$$

Substituting $\hat{\beta}$ and $\hat{\sigma}^2$ into (13), the problem is to numerically maximize

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

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

$$\hat{y}(x) = \hat{\beta} + r^T(x) R^{-1} (y_0 - l\hat{\beta}).$$

where $r^T(x)$ is the $1 \times n$ vector of correlation function between error $Z(x)$ and prediction point x .

KRG has received wide attention in many applications of CSE due to its interpolation property (Sacks et al. 1989).

2.5. Test problems

In this paper, three different test problems were selected to compare the predictive accuracy of two statistical models on three different optimal LHDs obtained from different optimality criteria. The test problems consist of two 2-dimension problems and 7-dimension problems, which are Welch function, Mystery function and Cyclone function respectively. The details of each test problems are shown in Table 2.

Table 2 The selected test problems

Problem	<i>d</i>	<i>n</i>	
Welch Function	2	9	$y = [30 + x_1 \sin(x_1)](4 + e^{-x_2}); 0 \leq x_1, x_2 \leq 5$
Mystery Function	2	9	$y = 2 + 0.01(x_2 - x_1^2)^2 + (1 - x_1)^2 + 2(s - x_2)^2 + 7 \sin(0.5x_1)(0.7x_1x_2);$ $0 \leq x_1, x_2 \geq 5$
Cyclone function	7	99	$f(x_1, x_2, ..., x_7) = 174.42 \left(\frac{x_1}{x_5} \right) \left(\frac{x_3}{x_2 - x_1} \right)^{0.85}$ $\times \sqrt[3]{ \frac{1 - 2.62 \left[1 - 0.36 \left(\frac{x_4}{x_2} \right)^{-0.56} \right]^{\frac{3}{2}} \left(\frac{x_4}{x_2} \right)^{1.16}}{x_6 x_7} };$ $0.09 \leq x_1, x_3, x_4 \leq 0.11, 0.27 \leq x_2 \leq 0.33, 1.35 \leq x_5 \leq 1.65,$ $14.4 \leq x_6 \leq 17.6, 0.675 \leq x_7 \leq 0.825$

3. Results and Discussion

The implementation in this study can be divided into two parts, constructing LHDs and creating approximation models. The first part is to construct the design based on simulated annealing (SA) search algorithm with three different criteria. According to second order polynomial, we have set up the dimension of design to be 2×9, 5×51 and 7×99. By implementing a program in R, we construct 10 optimal LHDs for each dimension design and then mean, max and min value of each criteria are reported in Tables 3-5. The last column labeled ϕ_p in each table presents the average ϕ_p of all 10 optimal LHDs obtained from each criterion.

Table 3 LHDs obtained from correlation criterion

Correlation Coefficient criterion			
	Mean	Max	Min
2 × 9	0	0	0
5 × 51	0.000061	0.000842	0.000371
7 × 99	0.000537	0.000757	0.000418

Table 4 LHDs obtained from ϕ_p criterion

ϕ_p criterion			
	Mean	Max	Min
2 × 9	0.407314	0.407314	0.407314
5 × 51	0.060143	0.060261	0.060012
7 × 99	0.290539	0.029199	0.028898

Table 5 LHDs obtained from multi-objective criterion

Multi-objective criterion				
	Mean	Max	Min	ϕ_p
2×9	0.106937	0.119820	0.105506	0.408603
5×51	0.019789	0.021179	0.019172	0.063478
7×99	0.009943	0.010464	0.000418	0.029802

LHDs in Table 3 will have orthogonal property according to correlation coefficient criterion, while LHDs in Table 5 will have space-filling property and in Table 6 will have a combination between orthogonal and space-filling property.

The results in Tables 3-5 indicates that LHDs obtained from correlation criterion have a little high value of ϕ_p , when comparing to LHDs from multi-objective criterion.

The second part of implementation is applying these designs (10 designs for each dimension) to create approximation models, which are RSM and Kriging. The three test problems are selected as described before. The 121 sample points are fixed as a test data for two-dimension problems; Welch function and Mystery and 500 sample points for seven-dimension problem; Cyclone model. The accuracy of the model is measured with Root Mean Square Error (RMSE). The result in Table 6 shows that RSM can perform better for small dimensional problems and Kriging performs better in a large dimensional problem. Kriging with LHDs from multi-objective criterion performs best with the lowest RMSE.

Table 6 RMSE values for RSM and Kriging for LHDs from ϕ_p , correlation and multi-objective criteria

		RSM		Kriging	
		Mean	S.D.	Mean	S.D.
Welch function	ϕ_p	1.12349	0.03852	1.02657	0.01443
	Corr	0.98865	0.31544	1.44351	0.24000
	Multi	1.10422	0.10152	1.00733	0.04897
Mystery	ϕ_p	5.52599	0.12122	5.41439	0.01718
	Corr	4.72598	0.69283	5.48841	1.05028
	Multi	5.33700	0.46630	5.29862	0.33136
Cyclone model	ϕ_p	0.00247	0.00016	0.00182	0.00013
	Corr	0.00336	0.00028	0.00193	0.00026
	Multi	0.00279	0.00012	0.00174	0.00018

4. Conclusions

This paper presented multi-objective criterion to construct orthogonal space filling designs and then developed approximation models from the designs. The results review that orthogonal space filling design can perform well, especially for Kriging model. In 2-dimension problem models, RSM with correlation coefficient criterion performs best, while multi-objective criterion performs second best. From the result, it can be concluded that orthogonal space filling design is potentially suitable for constructing accurate approximation models. Further, larger dimension problems and

the methods to adjust the weight between the two criteria in multi-objective should be studied and investigated.

References

- Cheong YP. Experimental design and analysis methods in reservoir uncertainty studies, PhD [dissertation], Perth: Curtin University of Technology. 2005.
- Fang H, Horstemeyer MF. Global response approximation with radial basis functions, *Eng. Optimiz.* 2006; 38: 407-424.
- Jin R, Chen W, Sudjianto A. An efficient algorithm for constructing optimal design of computer experiments. *J. Stat. Plan. Infer.* 2005; 134: 268-287.
- Joseph VR, Hung Y. Orthogonal-Maximin Latin hypercube designs. *Stat. Sinica.* 2008; 18: 171-186.
- Koehler J, Owen RJ. *Handbook of Statistics.* 1996; 13: 261-308.
- Li W, Wu CFJ. Columnwise-pairwise algorithms with applications to the construction of supersaturated designs. *Technometrics.* 1997; 39:171-179.
- McKay MD, Beckman RJ, Conover WJ. 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.
- Morris MD, Mitchell TJ. Exploratory design for computational experiments. *J. Stat. Plan. Infer.* 1995; 43: 381-402.
- Myers RH, Montgomery DC, Anderson-Cook CM. *Response surface methodology: Process and product optimization using designed experiments.* 3rd edition, United States of America, Wiley, 2009.
- Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments, *Stat. Sci.* 1989; 4(4): 409-435.
- Simpson TW, Lin DKJ, Chen W. Sampling strategies for computer experiments: Design and analysis, *Int. J. Reliab. Appl.* 2001; 2: 209-240.
- Timun R, Na-udom A, Rungrattanaubol J. Exploratory optimal Latin hypercube designs for computer simulated experiments. *Thail Stat.* 2011; 9(2): 171-93.
- van Laarhoven PJM, Aarts EHL. Simulated annealing, *Theory Appl. Math.* 1987; 37: 7-15.
- Welch WJ, Buck RJ, Sacks J, Wynn HP, Mitchell TJ, Morris MD. Screening, predicting, and computer experiments, *Technometrics.* 1992; 34: 15-25.
- Yosboonruang N, Na-udom A, Rungrattanaubol J. A comparative study on prediction accuracy of statistical models for modeling deterministic output responses. *Thail Stat.* 2013; 11(1): 1-15.