SciencePG
Science Publishing Group

Application of Latin Hypercube Sampling Based Kriging Surrogate Models in Reliability Assessment

Liu Chu[*]**, Eduardo Souza De Cursi, Abdelkhalak El Hami, Mohamed Eid**

Laboratory of Optimization and Reliability in Mechanical Structure, Department of Mechanics, National Institute of Applied Science of Rouen, Rouen, France

Email address:

Liu.chu@insa-rouen.fr (Liu Chu)

Abstract: Reliability assessment is one of the necessary and critical parts in structural design under uncertainties. The methods for structural reliability assessment aim at evaluating the probability of limit state by considering the fluctuation of acting loads, variation of structural component or system, and complexity of operating environment. Latin Hypercube sampling (LHS) method as advanced Monte Carlo simulation (MCS) has higher efficiency in sampling. It will be chosen and applied in this paper in order to obtain an effective database for building Kriging surrogate models. In this paper, we propose an effective method to have reliability assessment by Latin Hypercube sampling based Kriging surrogate models. This method keeps the certain level of accuracy in prediction of the response of a structural finite element model or other explicit mathematical functions.

Keywords: Latin Hypercube Sampling, Kriging Models, Reliability Assessment

1. Introduction

Uncertainty is an inevitable issue in the process of manufacture, infrastructure, and engineering design. Quantifying and propagating the uncertainty in the simulation or design process as a key component of risk analysis, robustness evaluation or reliability based optimization attracts attention of researchers and designer [1]. The traditional deterministic model is not effective for structural analysis because of avoiding the effects of uncertainties in the parameters. To consider parameter fluctuation in the real operation environment, the Monte Carlo simulation (MCS) is chosen to perform the stochastic simulation. It is one of most popular discrete algorithmic for uncertainty analyses and used with increasing frequency. MCS (sampling based approach) is useful for several reasons. First, the sampling based approach covers the full range of each uncertain variable in a complicated system. Second, modification of the model is not required, and direct estimates of distribution functions are provided. In addition, in the process of sampling, a variety of sensitivity analysis procedures are available. Last but not the least, analysis procedures can be developed and allow the propagation of results through systems of linked models [2].

However, MCS can reach a certain level of accuracy only if a very large number of iterations are performed. It is obvious that MCS methods become computational prohibitive when simulation model is complicated. To be more efficient than the random sampling method, several improved MCS methods with different sampling techniques have been developed. Importance sampling (weighted sampling), is expected to reduce error to zero if the probability density function is properly selected [3]. L. M. Berliner [4] applied importance sampling Monte Carlo method in sequential problems of Bayesian updating. P. Beaurepaire [5] attempted performing importance sampling technique in reliability based optimization of structure. In the literatures, the first-order sensitivity method, as a variance reduction technique, is also utilized to accelerate MCS estimation convergence [6]. The variance reduction techniques are especially important when MCS is applied to estimate small failure probability [7].

Latin Hypercube Sampling (LHS) method operates by subdividing the sample space into smaller regions and sampling within these regions. The produced samples more effectively fill the sample space and therefore reduce the

variance of computed statistical estimators [8]. Stein M [9] had research in the large sample properties of simulations using Latin hypercube sampling. Owen [10] and Huntington [11] tested the limitation of Latin Hypercube sampling. Improved LHS have been developed, Stocki R [12] projected samples onto a known subspace to minimize integrated mean square error and maximize entropy. Iman [13] had efforts to reduce spurious correlations, Florian [14] rearranged the matrix of samples based on a transformation of the rank number matrix. Further, methods for constructing orthogonal LHS are proposed to possess enhanced space filling properties [15], utilized iterative optimization methods are applied in LHS in order to reduce spurious correlations [16].

A surrogate model can be thought of as a regression to a set of data, where the data is a set of input-output pairing obtained by evaluating a black-box model of the complex system [17]. Surrogate models may be classified into two general categories based on their proposed: local and global models. An example of optimization using local is Response Surface Method expressed (RSM) as polynomial surrogate models. The kriging models are interpolation models based on the assumption that there is a spatial correction between the values of the function to be approximated [18]. With these models a known or sampled value of the limit state function to be approximated is exactly predicted. The Kriging models do not assume an underlying global functional form as assumed in the polynomial regression models and can approximate arbitrary functions with high accuracy in global as well as local approximations [19].

Kriging model was originally developed in spatial statistics by Krige. Matheron [20] had contribution in mathematical formulation of Kriging model. In the pioneering work of Sacks et al [21], Kriging model was chosen to predict deterministic functions in physical processes. A univariate kriging model combining a regression term with a zero-mean stationary Gaussian disturbance process was developed by Handcock [22]. Van Beer [23] started with the application of Kriging model in random simulation. The track record of application of Kriging models in random simulation holds great promise.

This paper presents an effective method to have reliability assessment by Latin Hypercube sampling based Kriging surrogate models. For that purpose, Latin Hypercube Sampling is utilized to build a reliable database for approximating the response of a structural finite element model or other explicit mathematical functions. Once the database is defined, we compare the relative performance of approximation methods to fit the probabilistic response of original models: namely, response surface method (first order and second order polynomial regressions) and Kriging models. Then, reliability assessment is predicted by the surrogate models, which heavily reduced computational cost and also kept the certain level of accuracy.

2. Latin Hypercube Sampling Method

A compromise method of advanced MCS is Latin hypercube sampling (LHS) approach. This approach divides the range of each variable into disjoint intervals of equal probability, and one value is randomly selected from each interval [24]. It improves MCS stability and also maintains the tractability of random sampling. Consider a statistic system described by the relation:

$$Y = F(X) \tag{1}$$

$$X = \{X_1, X_2, \cdots, X_n\} \tag{2}$$

Where the random vector X possesses independent components defined over the sample space S describing n input random variables with marginal cumulative distribution functions P. F is the operator commonly represents computer simulation such as finite element model with propagation of uncertainties in the system.

Traditional Monte Carlo methods rely on the so-called Simple Random Sampling (SRS) in which realization of X, denoted x_k, $k = 1, \cdots, N$ (samples), are generated as independent distribution realization on sample space by

$$x_{ki} = P_{X_i}^{-1}(U_i) ; \; i = 1, \cdots, n \tag{3}$$

Where U_i are uniformly distributed samples on [0, 1]. The realizations of x are then applied to the system $y = F(x)$ and y is statistically evaluated.

Correlated random variables are not considered because of Principal Component Analysis and Nataf or Rosenblatt transformations can be applied to produce a set of uncorrelated random variables from correlated relationships between variables.

LHS divides the range of each vector components $X_1, X_2, \cdots, X_n$ into disjoint subsets of equal probability. Samples of each vector components are drawn from the respective subset according to,

$$x_{ki}^j = P_{X_i}^{-1}(U_{ij}) \tag{4}$$

Where $i = 1, \cdots, n; j = 1, \cdots, m; \; n$ refers to the total number of vector components or dimensions of vector. m is the number of subset in a design. Subscript k denotes a specific sample,

Where U_{ij} are uniformly distributed samples on $\left[\xi_j, \xi_j' \right]$,

$$\xi_j = \frac{j-1}{m} \quad, \quad \xi_j' = \frac{j}{m} \tag{5}$$

The samples $x_k = \left[x_{k1}, x_{k2}, \cdots x_{kn} \right]$, $k = 1, \cdots, N$ are assembled by uniformly randomly grouping the terms of the generated vector components.

In Latin Hypercube Sampling approach, the range of all random input variables is divided into n intervals with equal probability, which are restricted within the respective interval avoid the disadvantage of clustering together, as demonstrated in Fig. 1.
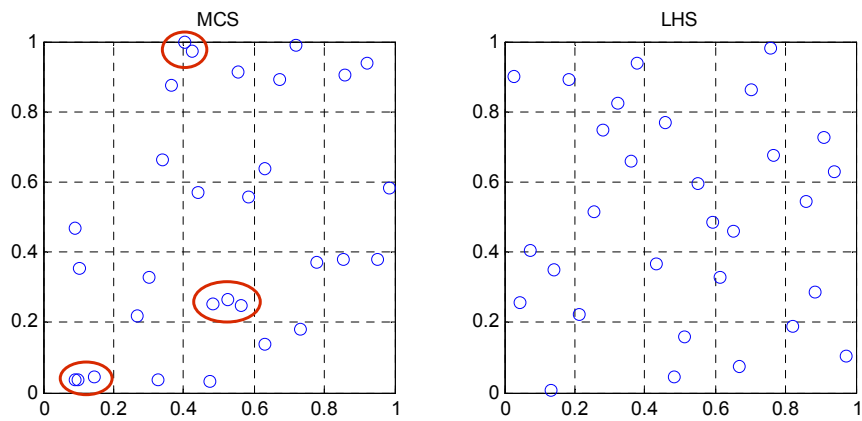
Fig. 1. *Comparison between MCS and LHS in samplingExample 1: Consider the cubic polynomial function given by as in reference [25].*
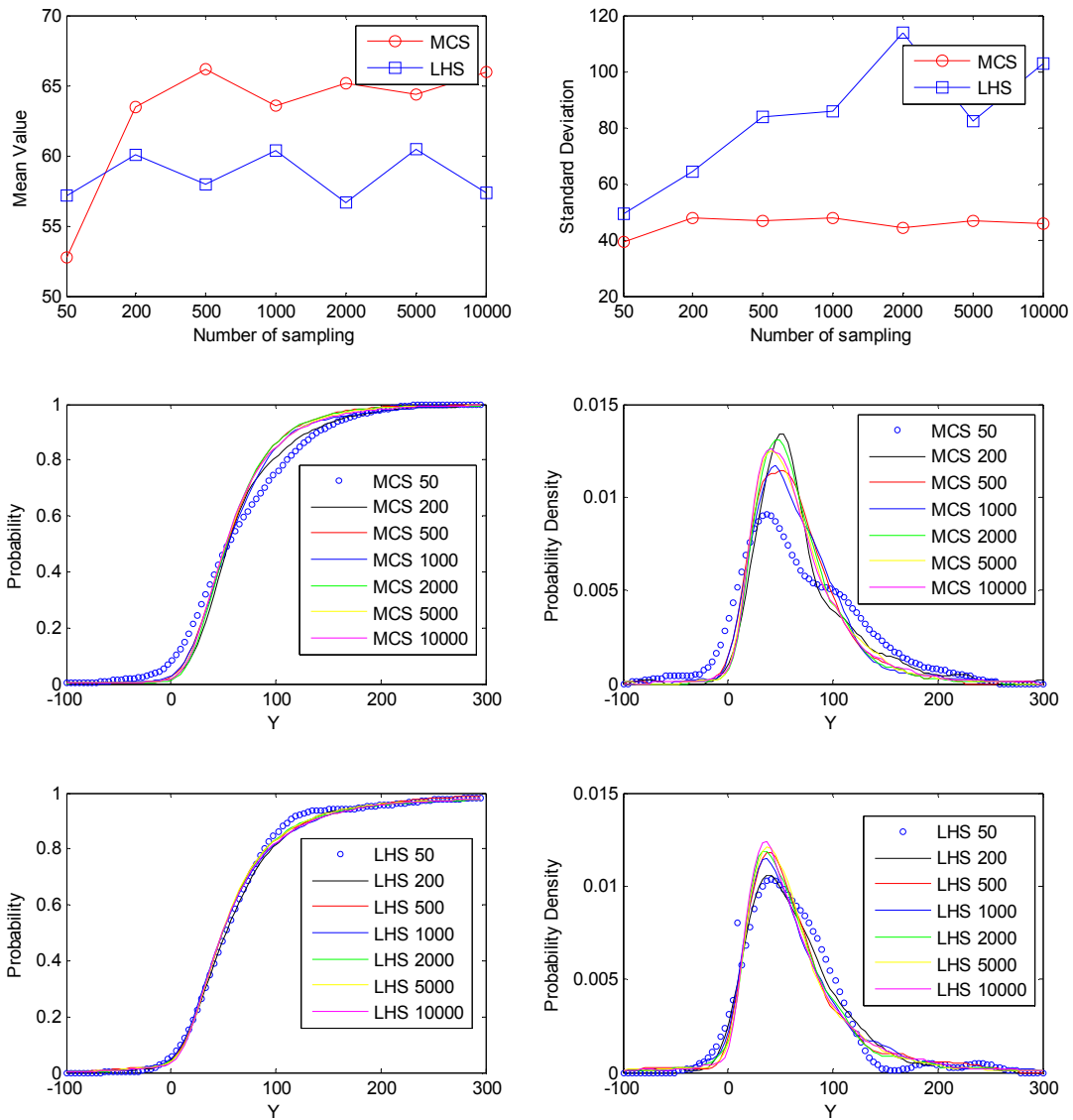

Fig. 2. *Results of Example 1 (Fig 2.a and Fig 2.b are the mean value and standard deviation records of MCS and LHS; Fig 2.c and Fig 2.d are the probability and probability density of MCS with different number of sampling; Fig 2.e and Fig 2.f are the probability and probability density of LHS with different number of sampling.)*
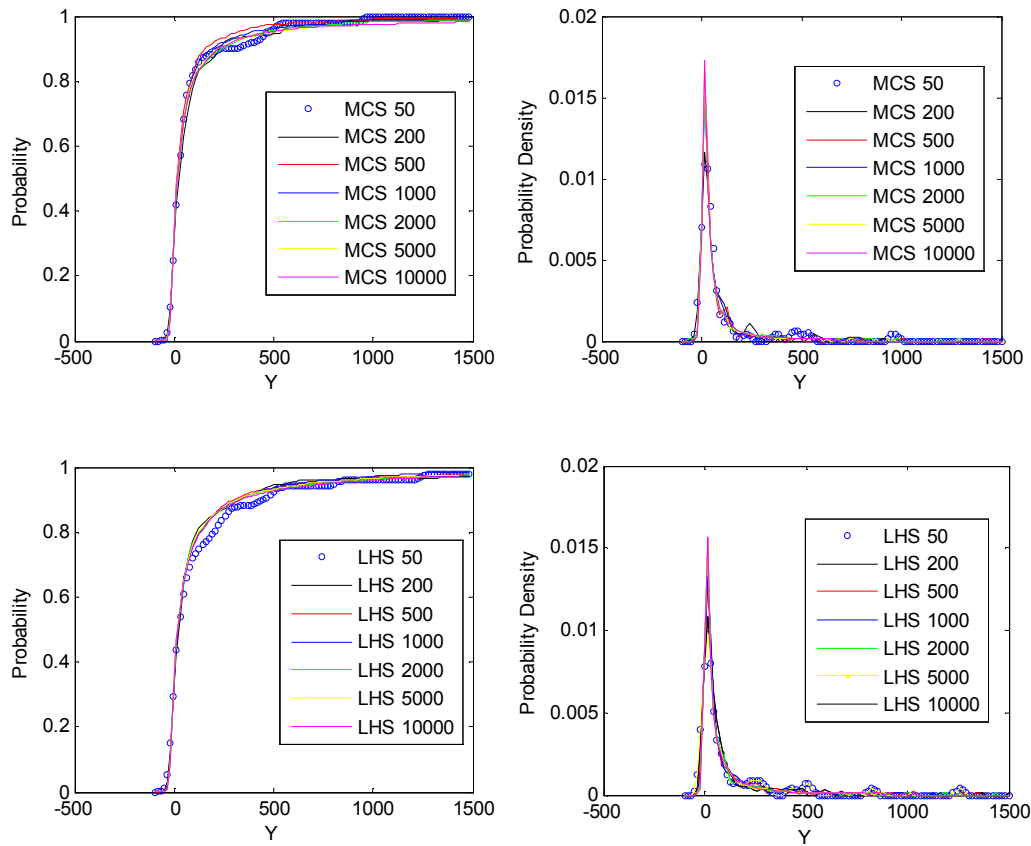
Fig. 3. *Results of Example 2 (Fig 3.a and Fig 3.b are the probability and probability density of MCS with different number of sampling; Fig 3.c and Fig 3.d are the probability and probability density of LHS with different number of sampling.)*

$$Y = F(\mathrm{X}) = X_1^2 X_2 - \alpha X_1 X_2^2 + X_1 X_2 \qquad (6)$$

The function possesses three random variables: X_1, X_2 and α, the probability density distribution of X_1, X_2 and α are Gaussian distribution (5, 1), Log-normal distribution (1, 0.5), and uniform distribution (0, 1), respectively.

Example 2: Suppose the limit state function expressed as,

$$Y = ML^2 / E \qquad (7)$$

Where M, L, E are three random variables. Here we defined the probability density distribution of M, L, E are Gaussian distribution (2, 1), Log-normal distribution (1, 0.5), and uniform distribution (0.01, 1), respectively.

The results of Example 1 and Example 2 of MCS and LHS in different number of sampling are presented in Fig. 2 and Fig. 3. To begin generating the LHS, an interval of each feature is chosen at random. The intersection of these intervals in the multi-dimensional feature space is a small hypercube, from which a sample is taken at random. Next, type of interval is selected at random for each feature. A sample is produced at random from that small hypercube. This continues until N samples have been generated. Each interval of each parameter is sampled exactly once in the process. In contrast to random sampling, the entire range of each feature is always represented in a LHS. Unbiased estimates of the sample means of the outcomes are obtainable. By comparing Fig 2.c with Fig 2.e, it is obvious that LHS method is easier to have convergent and accurate result in statistic with smaller number of sampling in Example 1 than MCS method. Fig 2.d and Fig 2.e support the same conclusion. However, in Example 2, when the probability density of output result is very concentrated in a small range, the advantage of LHS is not evident then MCS as in Fig. 3.

To compare MCS with LHS more precisely, here we defined relative error can be calculated as in Eq. (8)

$$e = \frac{\sum_{i=1}^{n} (p_i - \overline{p}_i)^2}{n}, \quad e_d = \frac{\sum_{i=1}^{n}(f_i - \overline{f}_i)^2}{n} \qquad (8)$$

Therefore, the relative errors in probability and probability density can be calculated as e and e_d. We divided the range of Y into $n-1$ intervals, accordingly n points in the probability and probability density can be obtained and expressed as p_i and f_i, where $\overline{p}_i$ and $\overline{f}_i$ are the probability and probability density of 10000 sampling in MCS or LHS.

Fig. 4 presents relative errors of MCS and LHS in Example 1 and Example 2. In Fig 4.a and Fig 4.b, the relative errors in probability density and probability of LHS are smaller than that of MCS when they have same number of sampling. However, in Fig 4.c and Fig 4.d, the advantage of LHS is not absolutely evident.
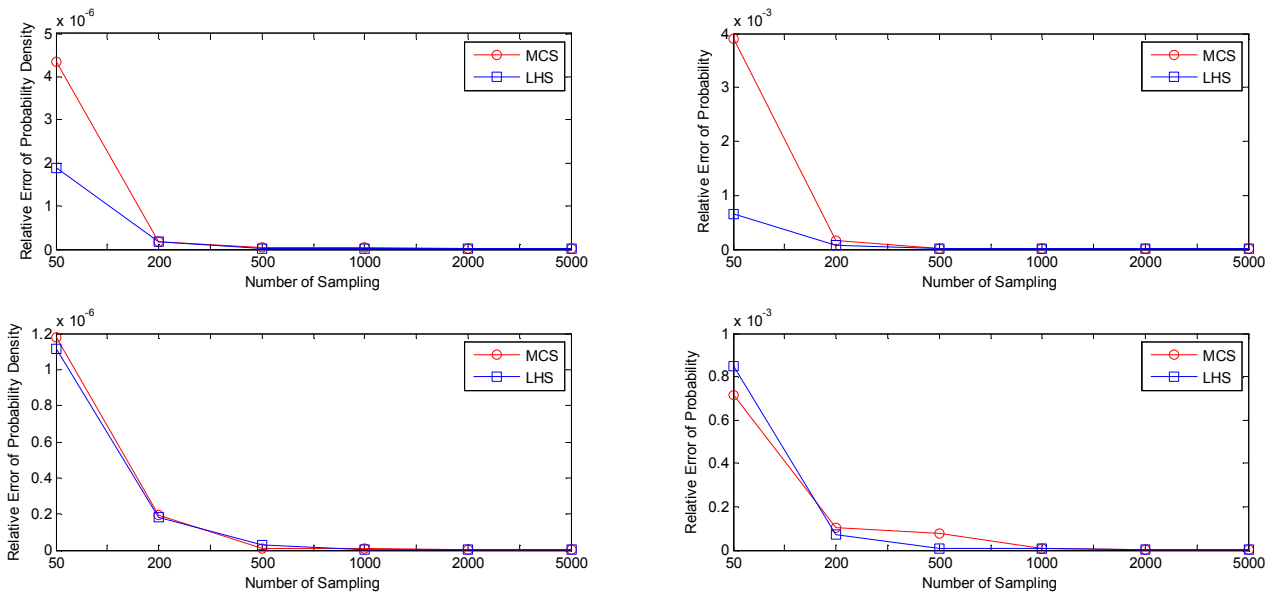
Fig. 4. *Relative errors of MCS and LHS in Example 1 and Example 2 (Fig 4.a and Fig 4.b are the relative errors in probability density and probability of MCS and LHS in Example 1; Fig 4.c and Fig 4.d are the relative errors in probability density and probability of MCS and LHS in Example 2.)*

3. Surrogate Models

A surrogate model can be thought of as a regression to a set of data, where the data is a set of input-output pairing obtained by evaluating a black-box model of the complex system [26]. Surrogate models may be classified into two general categories based on their proposed: local and global models. An example of optimization using local is Response Surface Method expressed (RSM) as polynomial surrogate models. RSM sequentially fits local first and second order regression models to a small region of the overall search space, as in Eq. (9) and Eq. (10).

$$F(\beta:x) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \tag{9}$$

$$F(\beta:x) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{ij} x_i x_j \tag{10}$$

In the other hand, a global surrogate model is a function that approximates the system across the design space. Kriging models fit a spatial correlation function to a data set consisting of input-output pairs obtained by evaluating the underlying function.

$$G(x) = F(\beta:x) + z(x) \tag{11}$$

Where $F(\beta:x)$ is a deterministic component defined by a regression model that gives an approximation to $G(x)$ in mean value and $z(x)$ is a stationary Gaussian process with zero mean and covariance,

$$Cov[z(x), z(x')] = \sigma^2 R(\theta:x,x') \tag{12}$$

That interpolates the errors between the regression model predictions $F(\beta:x)$ and the true limit state function values $G(x)$ at the m realizations of the vector of basic random variables x, with σ^2 the constant process variance and R is a prescribed correlation function.

Several correlation functions are available, such as the exponential, linear and Gaussian correlation functions, the most widely used correlation function for structural reliability problems is the anisotropic Gaussian correlation function

$$R(\theta:x,x') = \prod_{i=1}^{n} \exp(-\theta_i d_i^2) \tag{13}$$

With $d_i = x_i - x_i'$ the distance between the evaluation point x and the reference point x' in the ith direction of the basic random variables space and $\theta = [\theta_1, \ldots, \theta_n]^T$ a vector of parameters that define the inverse of the correlation length in each direction.

A kriging interpolation model is completely defined by a vector of regression coefficients β, a vector of correlation parameter θ and the variance σ^2 of the stationary Gaussian process. These parameters are estimated by fitting the Kriging model to a sample of support points.

Where F is the regression matrix and y is the vector of true limit state function values. The matrix R defines the correlation between each pair of support points according to the prescribed correlation function.

θ has to be first estimated using the method of maximum likelihood:

$$\hat{\theta} = \arg\min_{\theta} \{L(\theta)\} \tag{14}$$

$$L(\theta) = |R(\theta)|^{1/m} \sigma^2(\theta) \qquad (15)$$

Its prediction at a given point of the space of basic random variables can be obtained,

$$\hat{G}(x) = f(x)^T \hat{\beta} + r(x)^T \hat{\gamma} \qquad (16)$$

$$\hat{\gamma} = R^{-1}(y - F\hat{\beta}) \qquad (17)$$

$$r(x)^T = [R(\theta:x,x^{(1)}), \dots, R(\theta:x,x^{(m)})] \qquad (18)$$

A vector with the correlations between the prediction point and the m realizations $x^{(k)}$ $(k = 1, \dots, m)$ of the vector of basic random variables used in the Kriging model fitting corresponds to the expected or mean value of the Kriging model prediction, an estimate for the variance or uncertainty associated with the model predictions can be given by:

$$\sigma_G^2 = \sigma^2 \left[1 + u(x)^T (F^T R^{-1} F)^{-1} u(x) - r(x)^T R^{-1} r(x) \right] \qquad (19)$$

$$u(x) = F^T R^{-1} r(x) - f(x) \qquad (20)$$

$$F(\beta:x) = \beta_0 \prod_{i=1}^{n} x_i^{\beta_i} \qquad (21)$$

To gain some insight into the behavior of the Kriging surrogate models, we created Kriging models by database of MCS and LHS, which has different amount of sampling. Zero-order, first-order and second-order Kriging models are applied to predict the results when the input variables are changed in Example 1 and Example 2. In Example 1, the ranges of input variables (X_1, X_2 and α) are (4, 7), (0.1, 1), and (0.2, 0.8) respectively. They are totally included in the range of MCS and LHS in Section 2. In Example 2, we set up the ranges of input variables (M, L, E) as (0.5, 3.5), (0.7, 1.3), and (0.2, 0.8) respectively. It insures that the ranges of input variables in surrogate Kriging models do not exceed the scope of MCS and LHS.

Here, the relative errors are calculated by comparing the difference between the results predicting by Kriging models with the exact results from the analytical function in Example 1 and Example 2.

$$e_r = \sum_{i=1}^{i=n} \left(\frac{y_i - \overline{y}_i}{\overline{y}_i} \right)^2 \qquad (22)$$

Where y_i is the result predicted by Kriging models, $\overline{y}_i$ is the exact result from the analytical function, they have same input (X_{1i}, X_{2i} and α_i) in Example 1, and (M_i, L_i, E_i) in Example 2.

Table 1 and Table 3 present time cost of Kriging models fitting and predicting in Example 1 and Example 2. Table 2 and Table 4 provide the values of relative errors, which demonstrate the difference between the results predicted by Kriging models and the exact results when the input variables

are certain. We find that the number of sampling in MCS or LHS which is the original database for Kriging models is the most important factor to time cost. The sampling in MCS or LHS will be transferred to Kriging models as the original database to define parameters in Kriging models. According to the increase of numbers of sampling, the computational cost in Kriging models will sharply grow. This conclusion is proved in both Table 1 and Table 3.

In Table 2, the relative errors from zero-order, first-order and second-order Kriging model are very small. In Example 1, Kriging models have satisfied accuracy as surrogate models to predict. In addition, when the number of sampling in LHS and MCS is same, original database provided by LHS is always better than that of MCS in Example 1. Along to the increase in size of original database from LHS and MCS, results predicted by Kriging models are more precise. However, the computational cost also sharply increases as in Table 1. Therefore, for creating Kriging surrogate model, there is a trade-off between accuracy and computational cost.

Table 1. *Time cost of Kriging models fitting and predicting in Example 1.*

	KM0	KM1	KM2
LHS(200)	0.0615	0.1438	0.1323
LHS(500)	0.3296	0.3353	0.3499
LHS(1000)	1.5545	1.5945	1.6401
LHS(2000)	6.6056	6.7608	9.2099
LHS(5000)	62.829	58.204	65.002
MCS(200)	0.0756	0.0747	0.0825
MCS(500)	0.3356	0.3488	0.3381
MCS(1000)	1.5767	1.5835	1.6291
MCS(2000)	6.6484	6.8306	6.8552
MCS(5000)	52.465	58.248	59.021

Units in Table 1 for time cost are second. The results are tested in the same computer.

Table 2. *Relative error of Kriging models prediction in Example 1.*

	KM0	KM1	KM2
LHS(200)	4.4316 e-4	5.5266 e-4	9.2171 e-4
LHS(500)	2.0669 e-7	1.1147 e-7	1.0959 e-5
LHS(1000)	1.0467 e-7	1.1143 e-6	4.5837 e-6
LHS(2000)	3.9162 e-8	2.1744 e-8	3.4246 e-8
LHS(5000)	7.0082 e-8	3.2972 e-8	3.4849 e-8
MCS(200)	9.1157 e-5	2.2110 e-4	0.0529
MCS(500)	3.2057 e-6	4.9294 e-4	1.2637 e-4
MCS(1000)	4.7832 e-6	1.6464 e-5	7.9070 e-6
MCS(2000)	1.4250 e-7	8.8909 e-8	7.5655 e-6
MCS(5000)	5.1851 e-7	7.2244 e-7	1.1270 e-5

The results predicted by Kriging model in Example 2 do not have the same level of accuracy as in Example1. In Table 4, we can find the relative errors are unable to be ignored. To make Kriging surrogate model to be more appropriate in complicated situation, we still have a lot of work to do. In order to track the property of Kriging models, tests of zero-order, first –order and second-order Kriging models which are based on the database of LHS and MCS are performed and recorded in Fig. 5. Firstly, the results predicted by Kriging models which are based on database of LHS are

more approximated to the exact results of analytical function in Example 2 than that based on database of MCS. Therefore, for creating Kriging surrogate models, LHS method is a more effective and competitive method than MCS method. Secondly, at the beginning of the prediction, Kriging models have large fluctuation and the predicted results are not precise when compared with the exact results, this part should be taken into consideration and chosen to be removed or filtered. Lastly, compared the results predicted by zero-order, first –order and second-order Kriging models which are based on the database of LHS with different number of sampling, choice of best Kriging model depends on the certain situation. It is difficult to define which Kriging model is the best.
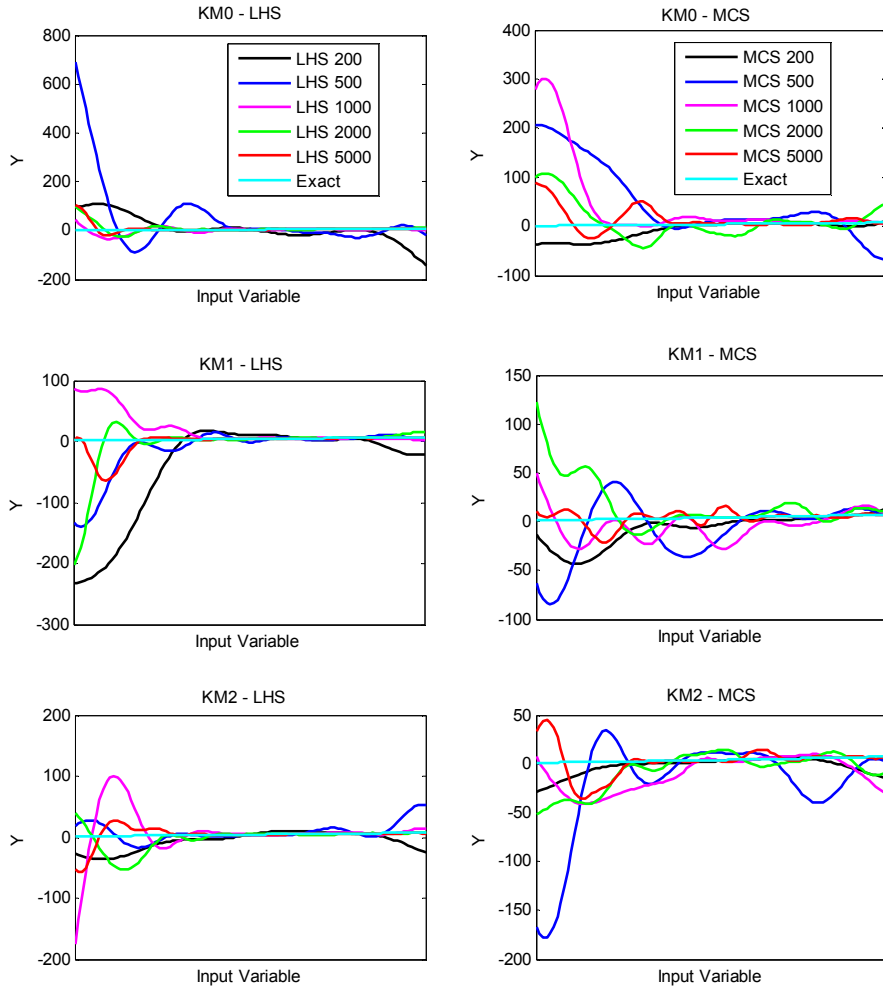
Fig. 5. *Kriging models prediction in Example 2 (Fig 5.a, Fig 5.c and Fig 5.e are the results predicted by zero-order, first-order and second-order Kriging models respectively which are based on database of LHS; Fig 5.b, Fig 5.d and Fig 5.f are the results predicted by zero-order, first-order and second-order Kriging models respectively which are based on database of MCS).*

Table 3. *Time cost of Kriging models fitting and predicting in Example 2.*

	KM0	KM1	KM2
LHS(200)	0.0724	0.0880	0.0858
LHS(500)	0.3934	0.3660	0.3690
LHS(1000)	1.7373	1.4166	1.8091
LHS(2000)	4.6267	4.6012	6.1435
LHS(5000)	27.170	27.082	27.405
MCS(200)	0.0699	0.0861	0.0674
MCS(500)	0.3038	0.3661	0.3729
MCS(1000)	1.3988	1.5859	1.6189
MCS(2000)	5.9070	3.8678	4.6432
MCS(5000)	26.815	32.289	43.2507

(Units in Table 3 for time cost are second. The results are tested in the same computer.)

Table 4. *Relative error of Kriging models prediction in Example 2.*

	KM0	KM1	KM2
LHS(200)	2.0766 e6	1.1141 e4	1.6238 e5
LHS(500)	1.8833 e3	4.5290 e5	2.0888 e5
LHS(1000)	7.3861 e3	5.1375 e4	1.9754 e4
LHS(2000)	219.7084	3.9539 e5	5.6788 e3
LHS(5000)	1.7566 e4	473.6334	2.6469 e3
MCS(200)	1.8039 e5	1.0552 e4	4.9243 e5
MCS(500)	1.9474 e5	2.2941 e4	1.9898 e3
MCS(1000)	1.1578 e5	9.4862 e4	1.3671 e5
MCS(2000)	9.2426 e3	1.6142 e5	3.8570 e3
MCS(5000)	2.4799 e5	6.4827 e3	1.7360 e4

4. Example in Finite Element Model

When the limit state function is implicit and each deterministic sampling is computational expensive, we propose creating Kriging surrogate models to have reliability assessment, LHS is chosen as an effective method to provide original database for Kriging models. Modal frequency is an important area of structural dynamics which has deservedly received much attention. Usually, researchers and designers identify the basic modal frequencies of a specific structural system and avoid the periodic loading coincide with them in order to prevent the damage or failure of resonance. The dynamic equation can be written as,

$$[M]\{\ddot{x}\}+[C]\{\dot{x}\}+[K]\{x\}=\{F\} \qquad (23)$$

Where $[M]$ is the mass matrix describing the distribution of mass, it is about the structural degree of freedom, $\{\dot{x}\}$ and $\{\ddot{x}\}$ are the first and second derivatives of the displacement with respect to time. Note that the force applied to the system is now a function of time. While mass and stiffness of a structure are measured and relatively easily derived, the mechanism whereby energy is lost through damping is less easily modeled. The viscous damping model represented by matrix $[C]$ is commonly but by no means exclusively used, being proportional to velocity. If there is no damping, the equation of motion is

$$[M]\{\ddot{x}\}+[K]\{x\}=\{F\} \qquad (24)$$

For free (unforced) vibrations the following relationship is obeyed

$$[M]\{\ddot{x}\}+[K]\{x\}=0 \qquad (25)$$

The solution to which can be written in the form

$$\{x\}=\{\psi\}_j\, e^{iw_j t} \qquad (26)$$

Where ω_j is the resonant frequency. Substituting back into the vibration equation leads to the well-known eigenvalue problem

$$[K]\{\psi\}_j = \lambda_j [M]\{\psi\}_j \qquad (27)$$

Where $\lambda_j = \omega_j^2$, and $\{\psi\}_j$ can be thought of the mode shapes corresponding to the system natural frequencies $\{\omega\}_j$.

While the eigenvalues have an exact relationship with the resonant frequencies, the eigenvectors are arbitrarily scaled; it is common practice to define a uniquely scaled set of eigenvectors such that

$$[\phi]^T[M][\phi]=[I] \qquad (28)$$

The result is

$$[\phi]^T[K][\phi]=diag(\lambda) \qquad (29)$$

Where $[\phi]$ is the matrix of mass normalized eigen-vector.

In this paper, our finite element model of wing structure, as presented in Fig. 1, is constructed by ANSYS Parameter Design Language. The parameters in the original deterministic model are corresponding with geometrical properties and material properties. Where S is the parameter representing the ratio of area between the two airfoil sections, it is 0.25 as in initial. L and D as presented in the Fig. 6, are 6.25 m and 1.42 m respectively. For material property, Young's module is 7e10 Pa, Poisson ratio is 0.33, and physical density is 2700 kg/m^3.

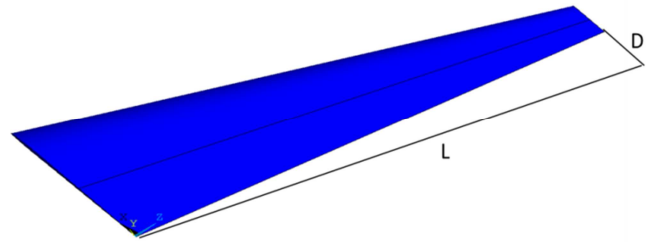**Fig. 6.** *Finite element model of wing structure.*

Table 5. *Results of deterministic finite element model.*

	1	2	3	4	5
Natural frequency / Hz	61468	197798	291869	447981	578028
Minimum stress / N/M2	0.243e10	0.705e10	0.188e11	0.294e11	0.482e11
Maximum stress / N/M2	0.249e12	0.189e13	0.143e13	0.631e13	0.391e13

The results of modal frequencies of wing structure in the deterministic finite element model are as presented in Table 5. Latin Hypercube sampling method is performed in the deterministic finite element model to calculate the modal frequencies. The parameters corresponding with geometry (S, D, L) and material property (E, P, R) are as input variables in the process of Latin Hypercube sampling, while the modal frequencies of specific wing structure are output variable in each sampling iteration. We chose uniform distribution as probability distribution for input variables. The ranges of S, D, L are (0, 0.5), (0, 3) and (4, 10) respectively. The ranges of E, P, R are (le10, 1e11), (0, 0.5) and (1000, 8000) respectively. The units for input variables keep unchanging as in the deterministic finite element model. We had 10000 samplings for each input variables by LHS method. Fig.7 provides the records of five modal frequencies in the process of stochastic simulation. To be more obvious, the accumulative probabilities of five modal frequencies of wing structure are presented in Fig.8 as numerical statistics. The evaluation of the stochastic simulation in finite element model of wing by Latin Hypercube sampling method is presented in Table 2. The mean value, standard deviation, skewness, and also the minimum and maximum of five modal frequencies for wing structure by 10000 LHS are obtained and concluded in Table 2.
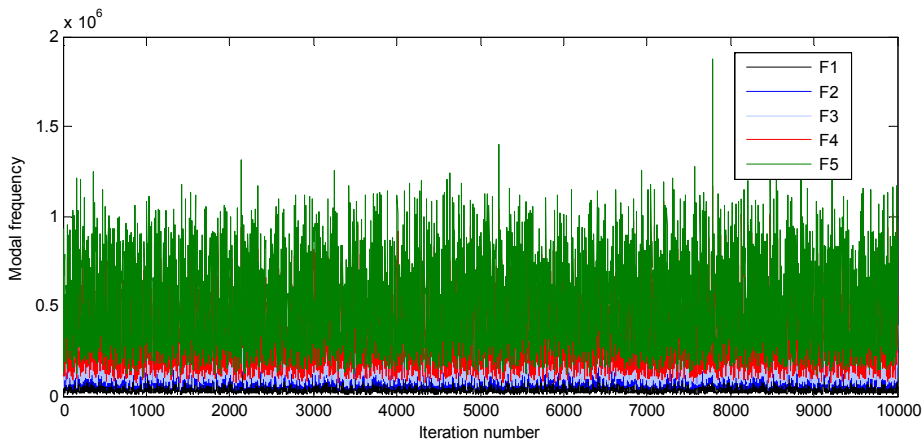
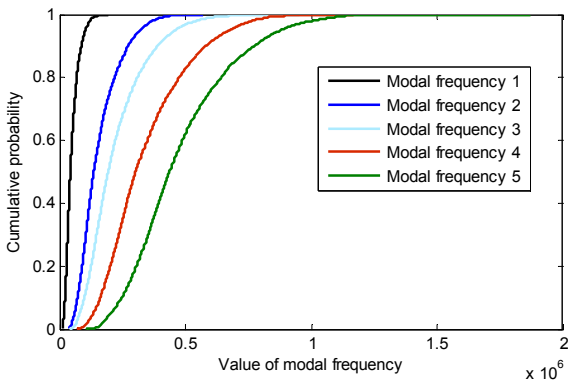**Fig. 7.** *Records of modal frequencies in the process of Latin Hypercube sampling.*

Fig. 8. *Cumulative probability of five modal frequencies in LHS.*

Table 6. *Results of Latin Hypercube Sampling method.*

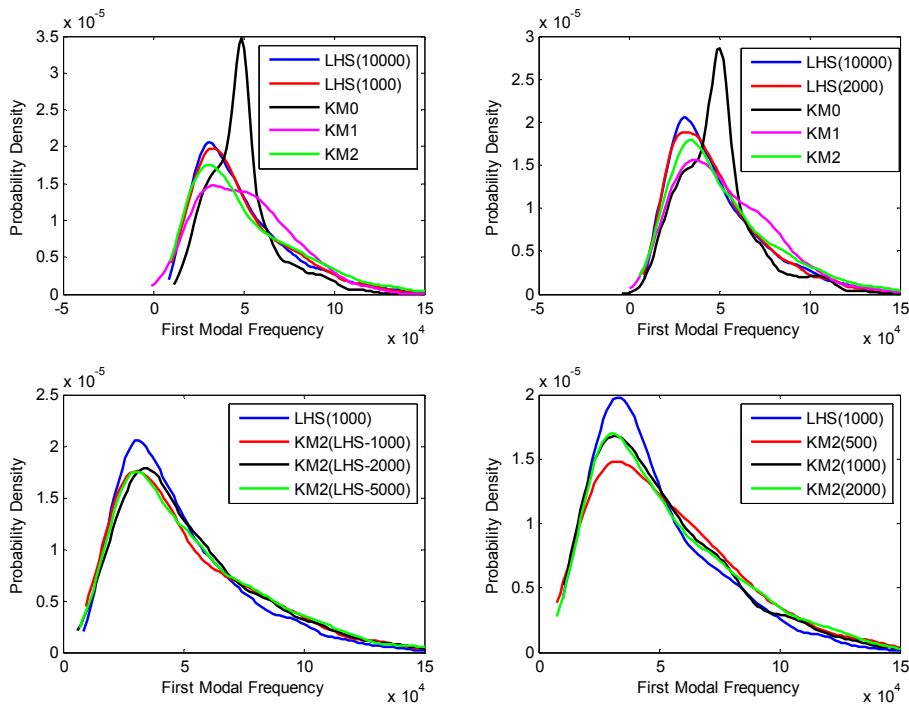|	**F1**	**F2**	**F3**	**F4**	**F5**
Mean value / *e5 Hz	0.48595	1.5632	2.2246	3.4208	4.7813
Standard deviation / *e5 Hz	0.26679	0.81028	1.1919	1.6633	2.0616
Skewness / *e5 Hz	1.2572	1.1292	1.2509	0.96708	0.91258
Minimum / *e5 Hz	0.08809	0.32320	0.41714	0.74041	1.0691
Maximum / *e5 Hz	1.9210	5.7092	9.4982	11.670	18.737

Fig. 9. *Tests of Kriging models as surrogate models for finite element model (Fig 9.a and Fig 9.b are the results of probabiltiy density of fist modal frequency predicted by Kriging models which are based on 1000 LHS and 2000 LHS, respectively; Fig 9.c presents the results predicted by second-order Kriging model based on different number of LHS; Fig 9.d presents the results predicted by second-order Kriging model with different amount of points).*

In order to identify one Kriging model as a surrogate model to replace time-cost computation of original model, we should find answers for three questions: firstly, which Kriging model is the most appropriate to this specific case; secondly, how many LHS should be provided as original database for Kriging model; lastly, we set up how many points in Kriging model for prediction. These three questions can be solved in Fig. 9. Fig 9.a presents the results of probabiltiy density of fist modal frequency predicted by zero-order, first-order and second-order Kriging models based on 1000 LHS and compared with the results of 1000 LHS and 10000 LHS. It is obvious that second-order Kriging model is more approximated to the original finite element model. Fig 9.b which presents the results predicted by Kriging models based on 2000 LHS supports the same point. Then, second-order Kriging model is chosen as surrogate model to have prediction.

In Fig 9.c, we find that the results predicted by second-order Kriging model based on different amount of LHS, namely 1000 LHS, 2000 LHS and 5000 LHS, are close to each other and convergent very well to the precise result. In the other hand, having prediction to 1000 points and 2000 points by defined second-order Kriging model has the same level of accuracy and stability as demonstrated in Fig 9.d. Kriging model as a surrogate model, in it, the second order regression provides convergent and accurate prediction results. In addition, the advantages of Kriging model are not only at their accuracy, but also reflect at time-saving process. The 5000 LHS and performing calculation of modal frequencies of wing structure in the finite element model costs 1955.491 s, and if 1000 sampling, it also cost 371.237 s; while in the surrogate model, fitting 1000 LHS in second-order Kriging model only costs 9.632 s, and predict the corresponding result of 1000 input random sampling, it only costs 10.713 s. The advantage of time-saving is very competitive as a surrogate model.

The convenience of Kriging surrogate model is significant in reliability assessment. Mechanical resonance may cause violent swaying motions and even catastrophic failure in improperly constructed structures including bridges, buildings and airplanes, a phenomenon known as resonance disaster. It is the tendency of a mechanical system to respond at greater amplitude when the frequency of its oscillations matches the system's modal frequency of vibration. Designers struggle to avoid this physical phenomenon happening in the operation situation. However, the traditional deterministic model ignores the effects of uncertainties in the real complicated operation environment. To propagate the uncertainty in the deterministic model of complicated system, the calculation expense is a heavy burden. It is necessary to create an appropriate surrogate model to have prediction and provide reliability assessment.
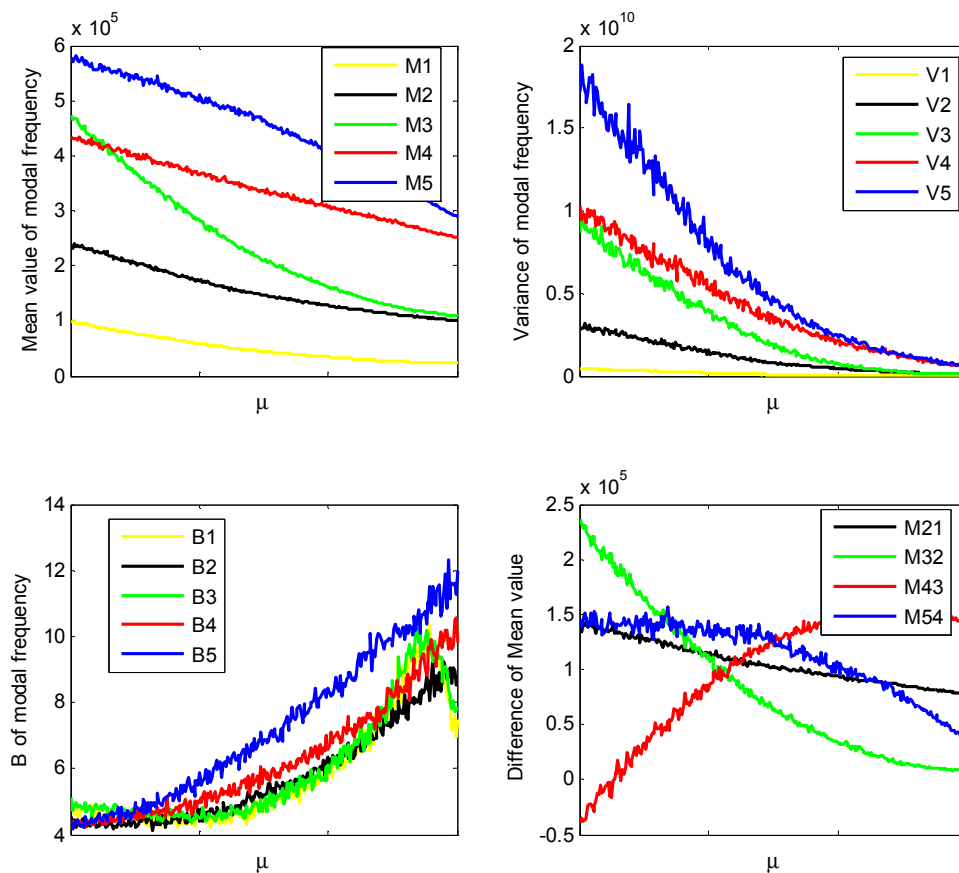

Fig. 10. *Results predicted by Kriging model based on LHS for reliability assessment (Fig 10.a and Fig 10.b present median value and variance for five modal frequency respectively; Fig 10. c provides results of B for five modal frequency; Fig 10.d presents the difference between two neighbor modal frequencies).*

In the example of finite element model of wing structure, the parameters corresponding with geometry (S, D, L) and material property (E, P, R) are supposed to be uncertain and fluctuate in a specific range in order to simulate the uncertainties in the real situation. To be general, the type of probability distribution is chosen to be Gaussian distribution, as (μ_S, σ_S), (μ_D, σ_D), (μ_L, σ_L), (μ_E, σ_E), (μ_P, σ_P) and (μ_R, σ_R) for the parameters respectively. The standard deviation is settled by 10% of the mean value for each parameter to simulate the fluctuation. Second–order Kriging model based on 1000 LHS is applied as surrogate model. Then, the relation between the probability property (mean value, variance, etc.) of five modal frequencies and mean value of input variables (the mean value of the input variables in this paper Synchronous increase, they can be set as any other tendency for reliability assessment).

In Fig 10.a and Fig 10.b, we can find that both mean value and variance for all five modal frequency will reduce when the mean value of input variables increase. B is the result of mean value divided by the standard deviation for every modal frequency. In contrary with mean value and variance of modal frequency, B has positive gradient as in Fig 10.c. The difference between two neighbor modal frequencies is an important parameter in structural design. Fig 10.d presents difference of mean value between the second and first modal frequency, between the third and the second modal frequency, and between the fifth and the forth modal frequency has the same tendency, they all have negative gradient. In contrary, the difference of mean value between the fourth and the third modal frequency has positive gradient. Therefore, Fig.10 provides essential information for structural designers.

5. Conclusion

1. LHS method is more effective than MCS method for providing comprehensive original database to Kriging models.
2. When the relation between the input variables and output results can be expressed as polynomial, Kriging models have satisfied accuracy as surrogate models to predict. Along with the increase in the number of sampling in LHS, the computational cost of creating Kriging models sharply grows, in the same time, the results predicted by Kriging models can be more precise. Therefore, there is a trade-off between computational expense and accuracy.
3. Comparison between zero-order, first-order and second-order Kriging model is not evident, which one is the most appropriate model depends on the specific relation between the input variables and output results.
4. Kriging models as surrogate models are very promising in reliability assessment. Its advantages of time-saving and high level of accuracy are competitive in uncertainty analysis.

References

[1] Donald L. Phillips, Danny G. Mark. Spatial uncertainty analysis: propagation of interpolation errors in spatially distributed models. Ecological Modelling, Volume 91, Issues 1–3, 15 November 1996, Pages 213-229.

[2] L. Goel, R. Billinton. Monte Carlo simulation applied to distribution feeder reliability evaluation. Electric Power Systems Research, Volume 29, Issue 3, May 1994, Pages 193-202.

[3] R.E. Melchers. Importance sampling in structural systems. Structural Safety, Volume 6, Issue 1, July 1989, Pages 3-10.

[4] L. Mark Berliner, Christopher K. Wikle. Approximate importance sampling Monte Carlo for data assimilation. Physica D: Nonlinear Phenomena, Volume 230, Issues 1–2, June 2007, Pp 37-49.

[5] P. Beaurepaire, H. A. Jensen, G. I. Schuëller, M. A. Valdebenito. Reliability-based optimization using bridge importance sampling. Probabilistic Engineering Mechanics, Volume 34, October 2013, Pp 48-57.

[6] Munoz Zuniga Miguel, Garnier Josselin, Remy Emmanuel. An original sensitivity statistic within a new adaptive accelerated Monte-Carlo method. Procedia - Social and Behavioral Sciences, Volume 2, Issue 6, 2010, Pages 7712-7713.

[7] Yaning Liu, M. Yousuff Hussaini, Giray Ökten. Optimization of a Monte Carlo variance reduction method based on sensitivity derivatives. Applied Numerical Mathematics, Volume 72, October 2013, Pages 160-171.

[8] Michael D. Shields, Kirubel Teferra, Adam Hapij, Raymond P. Daddazio. Refined Stratified Sampling for efficient Monte Carlo based uncertainty quantification. Reliability Engineering & System Safety, Volume 142, October 2015, Pp 310-325.

[9] Stein M. Large sample properties of simulations using Latin hypercube sampling. Technometrics 1987.

[10] Owen A. A central limit theorem for Latin hypercube sampling. J. R. Stat Soc B 1992, 54(2): 541-51.

[11] Huntington D. Improvements to and limitations of Latin hypercube sampling. Probabilistic Engineering Mechanics, 1998, 13(4): 245-253.

[12] Stocki R. A study on algorithms for optimization of Latin hybercubes. J Stat Plan Inference 2006, 136(9): 3231-3247.

[13] Iman R. A distribution-free approach to inducing rank correlation among input variables. Common Stat: Simul Comput 1982: 11(3):3 11-334.

[14] Florian A. An efficient sampling scheme: updated Latin hypercube sampling. Probabilistic Engineering Mechanics, 1992, 7: 123-130.

[15] Cioppa T, Lucas T. Efficient nearly orthogonal and space-filling Latin hypercubes. Technometrics 2007, 49(1), 45-55.

[16] Novak D. Correlation control in small-sample Monte Carlo type simulations i: a simulated annealing approach. Probab Eng Mech 2009, 24, 452-462.

[17] B. Gaspar, A. P. Teixeira, C. Guedes Soares. Assessment of the efficiency of Kriging surrogate models for structural reliability analysis. Probabilistic Engineering Mechanics, Volume 37, July 2014, Pages 24-34.

[18] M. Zakerifar, W. E. Biles and G. W. Evans. Kriging metamodeling in multi-objective simulation optimization. 2115-2122, 2009.

[19] Van Beers, W. C. M., J. P. C. Kleijnen. 2003. Kriging for Interpolation in Random Simulation. Journal of the Operational Research So-ciety, No. 54: 255-262.

[20] G. Matheron. Principles of geo-statistics. Economic Geology, 58(8)(1963), pp 1243-1266.

[21] J. Sacks, S. B. Schiller, W. Welch. Designs for computer experiment. Technometrics, 31(1) (1989), pp 41-47.

[22] M. Handcock, M. Stein. A Bayesian analysis of kriging. Technometrics, 35(3) (1993), pp 403-410.

[23] Van Beers, J. P. C. Kleijnen. Customized sequential designs for random simulation experiments: Kriging metamodeling and bootstrapping. European Journal of Operational Research, 186(3) 2008, pp. 1099-1113.

[24] Aleš Florian. An efficient sampling scheme: Updated Latin Hypercube Sampling. Probabilistic Engineering Mechanics, Volume 7, Issue 2, 1992, Pages 123-130.

[25] Michael D. Shields, Kirubel Teferra, Adam Hapij, Raymond P. Daddazio. Refined Stratified Sampling for efficient Monte Carlo based uncertainty quantification. Reliability Engineering and System Safety 142 (2015) 310-325.

[26] B. Gaspar, A. P. Teixeira, C. Guedes Soares. Assessment of the efficiency of Kriging surrogate models for structural reliability analysis. Probabilistic Engineering Mechanics, Volume 37, July 2014, Pages 24-34.