

A Hybrid Algorithm Combining Kriging Model and Importance Sampling for Structural Reliability Analysis

Juan Wang* and Yizhong Ma

School of Economics and Management, Nanjing University of Science and Technology, Nanjing, China, 210094

*Corresponding author

Abstract—One primary goal of structural reliability analysis aims at computing the probability of failure of a component or system with respect to some prescribed performance functions. In modern engineering, most performance functions usually resort to running an expensive-to-evaluate computational model, which leads to the infeasibility of simulation involving 10³–6 runs. Surrogate models, such as response surfaces approach or Kriging are introduced as a substitute for the original model to cope with the computational cost. However, the analysis based only on the surrogate model fails to quantify the error made by this substitution. In addition, the existing variance technique in simulation is still less efficient in terms of complex computational model. Therefore, a hybrid algorithm combining Kriging surrogate model and importance sampling for structural reliability analysis is proposed to calculate the failure probability. The applications in several reliability problems proves the approach efficient, robust and accurate.

Keywords—probability of failure; Kriging model; Importance Sampling; augmented failure probability; correction term

I. INTRODUCTION

In structural reliability analysis, failure probability is an important measure. Once it is calculated, the optimal scheme to increase reliability or to decrease failure probability is made accordingly. Normally the failure probability is expressed as the multiple integration. However, it is very difficult to calculate such failure probability. Various approaches concerning failure probability are proposed, including numerical-optimization-based FORM/SORM (First/Second Order Reliability Method)[1] and a series of simulation methods, such as directional simulation[2], and importance sampling[3].

Monte Carlo simulation is easy to implement, but it is not practical when the performance function involves complex computational model. Response surface methods are the mostly used surrogate model, but it fails to provide globally accurate interpolation. Within the latest 20 years, Kriging model[4] appear in the reliability analysis. However, most of the methods can not be categorized as the active learning ones since the updating of Kriging model did not use the predicted variance. An active learning method EGRA(Efficient Global Reliability Analysis)[5] is suggested to update the Kriging model by sequentially adding points into the design of experiment. Meanwhile, B. Echard[6] proposed a similar method AK-MCS(Active Reliability Method Containing

Kriging and Monte Carlo) with an altered updating criterion. Comparing AK-MCS with EGRA, the former need a global optimization algorithm; while, the minimal points of criteria function can be easily obtained through a surrogate model. When the accurate surrogate model is achieved, Monte Carlo simulation can be used to calculate the failure probability on the surrogate model instead of the original one. In addition, V.Dubourg[7] proposed an algorithm combining surrogate model and importance sampling to analyze structural reliability. It used MCMC and cross validation several times to calculate two important terms.

In this paper, a new algorithm based on Kriging model and importance sampling is proposed to calculate the failure probability in structural reliability analysis. The method separates the model updating and importance sampling. Compared with V.Dubourg[7], it is much more simple and efficient; meanwhile, the method guarantees a more accurate result. Firstly, the Kriging model is updated to be accurate, and the potential points are chosen as the least m points of a criteria function. Next, importance sampling is used to calculate the ‘augmented’ failure probability, and eventually the failure probability is expressed as the product of the ‘augmented’ failure probability and ‘correction term’. It is important to notice that the proposed method is implemented in the standard normal space, and if there exist non-normal variable or correlation among variables, the generalized Rosenblatt transformation is needed.

II. METHODOLOGY

A. Brief Introduction to Kriging Model

In the 80th of last century, Krige and Matheron proposed the famous Kriging model, which was soon applied in the computer experiment. It demonstrates its unique character: not only the predicted mean value can be provided, but the predicted variance. In the context of structural reliability, kriging assumes that the performance function is a realization of Gaussian process, and comprised of a regression model and a random error:

$$G(x) = f(x)^T \beta + Z(x) \quad (1)$$

where $\{f_i, i = 1, 2, \dots, p\} \in L_2$ are the basis function in regression model and $Z(x)$ is a stationary Gaussian process with mean value of zero. $Z(x)$ can be completely defined by an auto covariance function:

$$C_{GG}(x, x') = \sigma_G^2 R(x - x', l) \quad (2)$$

in which l is the parameter vector for defining the correlation coefficient. The extensively used auto covariance function is the non-homogeneous square exponential model:

$$R(x - x', l) = \exp\left(-\sum_{k=1}^n \frac{(x_k - x'_k)^2}{l_k}\right) \quad (3)$$

In summary, the optimal linear unbiased estimate of G at the point x can be specified as:

$$\hat{G}(x) \sim N(\mu_G(x), \sigma_G^2(X)) \quad (4)$$

The mean value and variance can be formulated as:

$$\mu_G(x) = f(x)^T \hat{\beta} = r(x)^T R^{-1}(y - F \hat{\beta}) \quad (5)$$

$$\sigma_G^2(x) = \sigma_G^2(1 - r(x)^T R^{-1} r(x) + \mathbf{u}(x)^T (F^T R^{-1} F)^{-1} \mathbf{u}(x)) \quad (6)$$

where $y = (g(x_1), \dots, g(x_m))^T$ and R is the correlation matrix from $R_{ij} = R(x_i - x_j, l), i, j = 1, \dots, m$, $r(x)$ is the correlation coefficient vector between experimental points and predicted points ($r_i(x) = R(x - x_i, l), i = 1, \dots, m$), F is the regression matrix ($F_{ij} = f_j(x_i), i = 1, \dots, m; j = 1, \dots, p$).

B. Updating of Kriging Model for Performance Function

EGO (Efficient Global Optimization) algorithm depends on the EIC (Expectation Improvement Criterion). For problems in structural reliability, the Kriging predicted mean value instead of true value of performance function is used in Monte Carlo simulation. Hence a different criteria, irrelevant with optimization, is needed to update the surrogate model: EGI (Efficient Global Inversion).

The so-called U function in B. Echard[6] is defined as the ratio of absolute value of predicted mean and the predicted variance, so the minimal points of U function is preferable:

$$U(x) = \frac{|\mu_G(x)|}{\sigma_G(x)} \quad (7)$$

Based on the U function, the proposed model updating is implemented as following procedures.

Step1: Generate $N(50000)$ points as the training set P . In this step, no performance function need to be evaluated; P only provides the potential points for model updating;

Step 2: Build an initial DoE: the points in DoE are obtained through kmeans or LHS(Latin Hypercube Sampling);

Step 3: Evaluate the performance function at the experimental points, and an initial Kriging model is established;

Step 4: Obtain the predicted values for the points in P according to the Kriging model;

Step 5: Calculate the U function values for P , and find the least m points in P ;

Step 6: Add the m points into the DoE to update the Kriging model;

Step 7: The termination condition is: $\min U(x) \geq 2$. If the condition is not achieved, loop the procedure 1-6.

C. Importance Sampling Based on Kriging Model

A fundamental method to calculate the failure probability is Monte Carlo simulation, and a multiple integration can be used to express this quantity:

$$P_f = \int_D f_X(X) dx = E_X[1_{g \leq 0}(X)] \quad (8)$$

where $f_X(X)$ is the joint probability density function and D is the failure domain. Hence the Monte Carlo estimate of failure probability can be expressed as:

$$\hat{P}_{fMC} = \frac{1}{N} \sum_{k=1}^N 1_{g \leq 0}(X^{(k)}) \quad (9)$$

Importance sampling attempts to sample points around the limit state function to improve the efficiency of simulation. Given the importance sampling density function, failure probability can be rewritten as:

$$P_f = \int_{R^n} 1_{g \leq 0}(x) \frac{f_X(x)}{h(x)} h(x) dx \equiv E_h[1_{g \leq 0}(x) \frac{f_X(x)}{h(x)}] \quad (10)$$

consequently, the corresponding estimate of (10) is:

$$\hat{P}_{fIS} = \frac{1}{N} \sum_{k=1}^N 1_{g \leq 0}(x^{(k)}) \frac{f_X(x^{(k)})}{h(x^{(k)})} \quad (11)$$

where $\{x^{(1)}, \dots, x^{(N)}\}$ are samples from $h(x)$. Besides, its effectiveness can be characterized by:

$$Var(\hat{P}_{f_{IS}}) = \frac{1}{N-1} \left(\frac{1}{N} \sum_{k=1}^N 1_{g \leq 0}(x^{(k)}) \frac{f_X(x^{(k)})^2}{h(x^{(k)})^2} - \hat{P}_{f_{IS}}^2 \right) \quad (12)$$

According to Rubinstein and Kroese[8], when $h(x)$ takes:

$$\hat{h}^*(x) = \frac{1_{g \leq 0}(x) f_X(x)}{\int 1_{g \leq 0}(x) f_X(x) dx} = \frac{1_{g \leq 0}(x) f_X(x)}{P_f} \quad (13)$$

the variance of $\hat{P}_{f_{IS}}$ is zero. Because it is impossible to achieve, a sub-optimal density function $\hat{h}^*(x)$ is a better option. Therefore a probability classification function[9] is introduced:

$$\pi(x) \equiv P[G(x) \leq 0] \quad (14)$$

Combined with Kriging model, Eq (14) can be rewritten as:

$$\pi(x) = \Phi\left(\frac{0 - \mu_G(x)}{\delta_G(x)}\right), \text{ if } x \notin DoE \quad (15)$$

For points in DoE:

$$\pi(x) = \begin{cases} 1, & g(x) \leq 0 \\ 0, & g(x) > 0 \end{cases} \quad (16)$$

Substituting $\pi(x)$ for $1_{g \leq 0}(x)$, we will get:

$$P_{f_{f\epsilon}} \equiv \int_{R^n} \pi(x) f_X(x) dx = E_X(\pi(x)) \quad (17)$$

Here, $P_{f_{f\epsilon}}$ is regarded as 'augmented' failure probability.

Meanwhile, the substitution also lead to a sub-optimal importance sampling density function:

$$\hat{h}^*(x) = \frac{\pi(x) f_X(x)}{\int \pi(x) f_X(x) dx} = \frac{\pi(x) f_X(x)}{P_{f_{f\epsilon}}} \quad (18)$$

In importance sampling, the adoption of $\hat{h}^*(x) = \frac{\pi(x) f_X(x)}{P_{f_{f\epsilon}}}$ as a density function results in the following:

$$P_f = \int 1_{g \leq 0}(x) \frac{f_X(x)}{\hat{h}^*(x)} \hat{h}^*(x) dx = P_{f_{f\epsilon}} \cdot \int \frac{1_{g \leq 0}(x)}{\pi(x)} \hat{h}^*(x) dx \equiv P_{f_{f\epsilon}} \alpha_{corr} \quad (19)$$

where $\alpha_{corr} = \int \frac{1_{g \leq 0}(x)}{\pi(x)} \hat{h}^*(x) dx = E_h [1_{g \leq 0}(x) / \pi(x)]$, and we define it as a 'correction term'. The respective Monte Carlo estimates of $P_{f_{f\epsilon}}$ and α_{corr} can be written as:

$$\hat{P}_{f_{f\epsilon}} = \frac{1}{N_{f\epsilon}} \sum_{i=1}^{N_{f\epsilon}} \pi(x^{(i)}) \quad (20)$$

$$\hat{\alpha}_{corr} = \frac{1}{N_{corr}} \sum_{j=1}^{N_{corr}} \frac{1_{g \leq 0}(h^{(j)})}{\pi(h^{(j)})} \quad (21)$$

To generate samples from \hat{h}^* , modified Metropolis-Hasting Sampler algorithm is used, in which the thinning procedure guarantees the independence between samples.

III. NUMERICAL EXAMPLE

An analytical performance function is defined as follows: $f(x) = b - x_2 - k \times (x_1 - 1)^2$, with $b = 5, k = 0.5, e = 0.1$. This is the famous Kiureghian function. With respect to FORM, this performance function has more than two design points. To improve the efficiency of updating surrogate model, we add 6 points to DoE in each iteration, and after 3 iteration, the Kriging reaches its precision, as described in the algorithm. To achieve of same accuracy of failure probability, B.Echard[6] calculated 120 values of performance function, as well as 120 model updating. Using the result $P_f = 0.000359$ from crude Monte Carlo with 500000 simulation as a reference value, the error from the proposed method is very small. Not surprisingly, the result from FORM deviates the reference a lot since the performance function is highly nonlinear. Figure 1, 2, 3 demonstrate the evolution of Kriging surrogate model. From these figures, we can observe that after 3 iteration, which corresponds to 18 most potential points into the DoE, a very accurate surrogate model of the original model can be obtained, which guarantee the precision of the subsequent calculation. Moreover, the devised method also compensate the 'error' with a correction term. Table 1 presents the different results resulted from distinct methods. It clearly shows the efficiency and robustness of the proposed method.

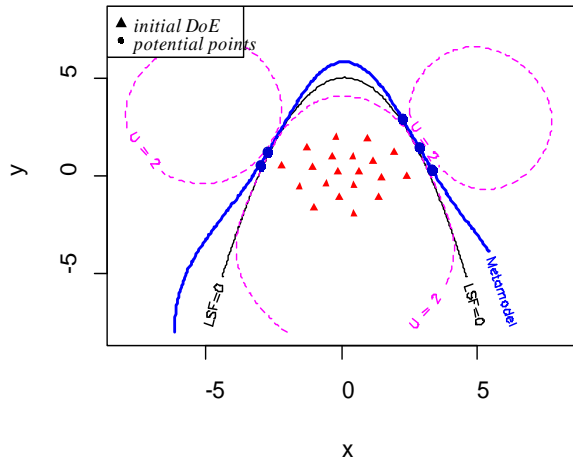


FIGURE I. INITIAL DOE AND KRIGING MODEL

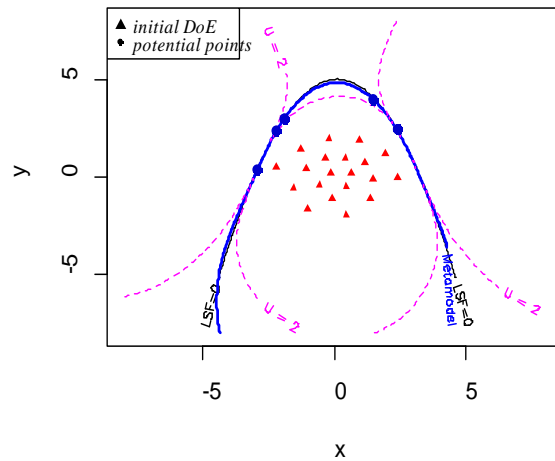


FIGURE II. KRIGING MODEL AFTER 1 ITERATION

TABLE I RESULTS FROM DIFFERENT METHODS

	Monte Carlo Simulation	the proposed method	Importance Sampling	AKMCS	FORM
P_f	0.000359	0.000313	0.000282	0.000379	0.000109
error	-----	0.35%	2.71985%	0.268%	-----
Number of g-evaluation	500000	32	4750	120	1000000

reliability analysis is proposed. It is proved to be efficient and accurate through a numerical example.

The proposed method provide a useful and efficient way to access the failure probability of a component or of a system. It will shed light on the more complicated problem: reliability-based design optimization, which need to be decoupled in the modeling.

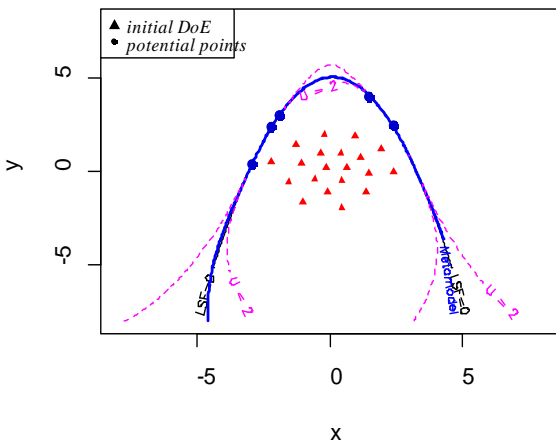


FIGURE III. AN ACCURATE KRIGING MODEL

IV. CONCLUSION

Due to the fact that the early surrogate-based reliability analysis can't quantify the error resulted from the substitution, and that the existing variance reduction techniques remain time-consuming when the performance function involves the complex computational model, a hybrid algorithm combining Kriging model and importance sampling for structural

REFERENCES

- [1] Zhao Y G, Ono T. A general procedure for first/second-order reliability method (FORM/SORM), *J. Structural safety*, 21 ,1999, pp.95-112.
- [2] Bjerager P. Probability integration by directional simulation, *J. J of Engi. Mechanics*, 114,1988, pp. 1285-1302.
- [3] Engelund S, Rackwitz R. A benchmark study on importance sampling techniques in structural reliability, *J. Structural safety*, 12,1993, pp. 255-276.
- [4] Sacks J, Welch W J, Mitchell T J, et al. Design and analysis of computer experiments, *J. Statistical Sci.*, 1989, pp. 409-423.
- [5] Bichon B J, Eldred M S, Swiler L P, et al. Efficient global reliability analysis for nonlinear implicit performance functions, *J. AIAA journal*, 46,2008, pp. 2459-2468.
- [6] Echard B, Gayton N, Lemaire M. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation, *J. Structural Safety*, 33,2011, pp. 145-154.
- [7] Dubourg V, Sudret B, Deheeger F. Metamodel-based importance sampling for structural reliability analysis, *J. Probabilistic Engineering Mechanics*, 33, 2013, pp. 47-57.
- [8] Rubinstein R Y, Kroese D P. *Simulation and the Monte Carlo method*, M. John Wiley & Sons, 2016
- [9] Picheny V. *Improving accuracy and compensating for uncertainty in surrogate modeling*, D. University of Florida, 2009