

## An Efficient Variable Screening Method for Effective Surrogate Models for Reliability-Based Design Optimization

**Hyunkyoo Cho<sup>1</sup>, Sangjune Bae<sup>1</sup>, K.K. Choi<sup>1\*</sup>, David Lamb<sup>2</sup>, Ren-Jye Yang<sup>3</sup>**

<sup>1</sup>The University of Iowa, Iowa City, Iowa, USA, \*kyung-choi@uiowa.edu

<sup>2</sup>US Army RDECOM/TARDEC, Warren, Michigan, USA, david.a.lamb40.civ@mail.mil

<sup>3</sup>Ford Motor Company, Dearborn, Michigan, USA, ryang@ford.com

### 1. Abstract

In the reliability-based design optimization (RBDO) process, surrogate models are frequently used to reduce the number of simulations since analysis of a simulation model takes a great deal of computational time. On the other hand, to obtain accurate surrogate models, we have to limit the dimension of the RBDO problem and thus mitigate the curse of dimensionality. Therefore, it is desirable to develop an efficient and effective variable screening method for reduction of the dimension of the RBDO problem. In this paper, requirements of the variable screening method for deterministic design optimization (DDO) and RBDO are compared, and it is found that output variance is critical for identifying important variables in the RBDO process. An efficient approximation method based on the univariate dimension reduction method (DRM) is proposed to calculate output variance efficiently. For variable screening, the variables that induce larger output variances are selected as important variables. To determine important variables, hypothesis testing is used in this paper so that possible errors are contained in a user-specified error level. Also, an appropriate number of samples is proposed for calculating the output variance. Moreover, a quadratic interpolation method is studied in detail to calculate output variance efficiently. Using a 44-dimensional example, performance of the proposed method is verified. It is shown that the proposed method finds important variables efficiently and effectively.

**2. Keywords:** Variable Screening, RBDO, Output Variance, 1-D Surrogate Model, Univariate Dimension Reduction Method

### 3. Introduction

The variable screening method is a useful method in the design optimization process since it can select essential design variables for accurate surrogate models and effective design optimization. In the formulation of a design optimization problem, a set of design variables that describe the system need to be identified [1]. Design variables are selected to be independent of each other as much as possible in the design space. The number of independent design variables is known as the degrees of freedom, and this is the dimensionality of the optimization problem. To obtain an appropriate optimum design, a minimum number of design variables are required. For this reason, it is better to identify as many design variables as possible, and then fix some of the variables at certain values according to the variable screening result. The variable screening method can play a key role, especially in reliability-based design optimization (RBDO), because the RBDO process requires a larger number of analyses than the deterministic design optimization (DDO) process due to reliability analyses and the design sensitivities of probabilities of failure. To this end, surrogate models are usually used to reduce the number of analyses required in RBDO. Various surrogate model methods such as the radial basis function (RBF), polynomial response surface (PRS), support vector regression (SVR), Kriging, and dynamic Kriging have been developed [2-10]. However, even for the surrogate model, the number of design variables becomes critical factor since surrogate model generation is difficult for high-dimensional problems, due to the curse of dimensionality.

Variable screening methods have been developed in various disciplines. In statistics, important variables were found to create an accurate surrogate model of computer simulation using the maximum likelihood estimator (MLE) of correlation parameters of the Gaussian process for a deterministic problem [11]. Using a regression model, essential variables among candidate variables were efficiently identified based on data [12, 13]. In physics, a variable screening model was developed for the quasi-molecular treatment of ion-atom collision [14]. In engineering, a confidence interval of the coefficient of a linear surrogate model was proposed to detect key variables for a car crash DDO problem [15]. A sampling-based sensitivity measure using a small amount of data was introduced to rank the importance of variables and was applied to long-term performance of a geologic repository for high-level radioactive waste [16]. Moreover, the design sensitivity method can be extended to the variable screening method because vital variables have larger design sensitivity. In the deterministic problem, the design sensitivity, which shows the rate of change in the performance measure at the design point, can be

obtained using various methods [17, 18] and is called local sensitivity analysis (LSA) [19, 20]. For a reliability problem, the variability of the input random variable should be incorporated to assess the design sensitivity of a probabilistic constraint. The design sensitivity of the probabilistic constraint using the first-order reliability method (FORM) [21-23], dimension reduction method (DRM) [24, 25], and sampling-based stochastic sensitivity [26, 27], could be used to identify important design variables. In addition, global sensitivity analysis (GSA), such as correlation ratio [28], global sensitivity indices [29], and analytical GSA methods [20], can be used for variable screening as well [30].

However, previous works may have limitations to being directly applied to RBDO with surrogate models. If a method depends entirely on existing data [12, 13], it may not be possible to carry out RBDO since design variables change during the optimization process. A method developed for a specific problem [14] will be inadequate for broad. Variable screening and design sensitivity methods for a deterministic problem [11, 15, 17, 18] may not be applicable for RBDO since input randomness is not considered. Methods that require a very large number of analyses [28, 29] could be ineffective for RBDO of computationally demanding problems and become unstable when sufficient numbers of analyses are not provided [16]. The design sensitivity of the probabilistic constraint using FORM or DRM [21-25] requires searching for the most probable point (MPP), which may be very difficult to obtain for a large-dimensional problem. If a method was developed based on assumption that accurate full-dimensional surrogate models are available *a priori* [20], RBDO could be carried out using the surrogate models as it can provide accurate response and sensitivity of the problem, unless the optimization algorithm has limitation of the number of design variables, which is not a common case. From the previous works, key desirable properties of a variable screening method for RBDO with a surrogate model were found: it should (1) be efficient, (2) consider input randomness, (3) not require a full-dimensional surrogate model, and (4) be applicable to broader problems.

Therefore, the objective of this paper is to develop a variable screening method that can satisfy the above desirable properties. In this paper, a partial output variance, which is the output variance when one design variable has variability while others are fixed at their mean, is used to find important design variables [31]. The output variance is a combination of the input variability and local deterministic design sensitivity of the performance function. The partial output variance is simple to calculate and requires a 1-D surrogate model for each design variable. The reliability analysis in RBDO captures the output variability induced by the input variability and the sensitivity of the performance function. The variable that induces larger output variability is important in the RBDO process. The method introduced in this paper has strengths and weaknesses. Its main strength is its efficiency and practical applicability. Its weakness is accuracy; the interactions between the design variables are not fully captured. However, practical applicability is the focus in this paper because it is very important for large-scale problems. In the following sections, the proposed method will be explained in detail, and their strengths and weaknesses will be fully discussed. To demonstrate effectiveness of the proposed method, a large-scale industrial problem is used.

#### 4. Variable Screening

As explained in the introduction, screening out variables means finding important variables among all random variables. Here, the word “important” could have different meanings depending on the problem we are dealing with. In the following two sections, the difference between variable screening for DDO and RBDO will be explained. Based on the difference, the required properties of variable screening for RBDO will be introduced.

##### 4.1. Variable Screening for DDO

A DDO problem can be formulated as,

$$\begin{aligned}
 & \text{minimize} && \text{cost}(\mathbf{d}) \\
 & \text{subject to} && G_j(\mathbf{d}) \leq 0, \quad j = 1, \dots, NC \\
 & && \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U, \quad \mathbf{d} \in \mathbb{R}^{NDV}
 \end{aligned} \tag{1}$$

where  $\mathbf{d}$ ,  $G_j$ ,  $NC$ , and  $NDV$  are the design variable vector,  $j^{\text{th}}$  constraint function, number of constraints, and number of design variables, respectively.

As stated before, in the DDO problem, the input design variables do not have uncertainty, and thus the design sensitivity can be used as a barometer to determine the importance ranking of design variables with respect to the performance measure. The question is: “Where should the importance ranking of design variables be determined?” or “Where should the design sensitivity be calculated?”

The LSA calculates the design sensitivity at a given design point [19, 20]. Usually, LSA is used to provide the direction of design movement in the optimization process. For variable screening, LSA can provide the importance ranking of design variables at the current design point. However, the importance ranking at the given

design point could be different from the ranking at other design points if the performance measure is a nonlinear function of design variables. On the other hand, GSA is used to calculate overall design sensitivity on the entire design domain. The GSA is like averaged design sensitivity in the design domain. As it is an average, the importance ranking using GSA could mislead at specific points or even regions. Hence, LSA and GSA have advantages and disadvantages for variable screening [19].

#### 4.2. Variable Screening for RBDO

A general RBDO problem can be formulated as,

$$\begin{aligned}
& \text{minimize} && \text{cost}(\mathbf{d}) \\
& \text{subject to} && P_{F_j} = P[G_j(\mathbf{X}) > 0] \leq P_{F_j}^{Tar} \quad j = 1, \dots, NC \\
& && \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U, \mathbf{d} \in \mathbb{R}^{NDV}, \text{ and } \mathbf{X} \in \mathbb{R}^{NRV}
\end{aligned} \tag{2}$$

where  $\mathbf{d}$ ,  $G_j$ ,  $P_{F_j}^{Tar}$ ,  $NC$ ,  $NDV$ , and  $NRV$  are the design variable vector,  $j^{\text{th}}$  constraint function,  $j^{\text{th}}$  target probability of failure, number of constraints, number of design variables, and number of random variables, respectively.

In the RBDO process, design variable vector  $\mathbf{d}$  is the mean vector of the corresponding random variable  $\mathbf{X}$ . Though the design variable  $\mathbf{d}$  is deterministic, the design sensitivity for RBDO should consider the randomness of  $\mathbf{X}$  because the constraints are based on the probabilistic performance measure  $P[G_j(\mathbf{X}) > 0]$  as shown in Eq.(2). Therefore, the design sensitivity of the performance measure alone cannot be used as a barometer. The design sensitivity of the probabilistic performance measure can be obtained using by several methods, such as FORM [21-23], DRM [24, 25], and sampling-based stochastic sensitivity [26, 27]. The design sensitivity of the probabilistic performance measure can be used for variable screening. The design sensitivities by those methods are LSA because they provide different sensitivities at different designs. The GSA method is also applicable for variable screening in RBDO problems as it is in DDO problems. Again, both LSA and GSA methods have advantages and disadvantages.

The random parameters will not increase the dimensionality of the optimization problem since they are not random design variables. However, the surrogate model that includes random parameters is still required because they affect the output distribution. The main objective of this paper is to select important design variables so that accurate surrogate models can be generated and, at the same time, an appropriate optimum design (i.e., not suboptimum) can be obtained in the RBDO process. Hence, once variable screening is done, the screened-out random design variables need to be fixed, not to be a random parameter. However, fixing a random variable as a deterministic variable will reduce the total output variability.

Consider a simple example:

$$\begin{aligned}
X_i &\sim N(5, 3^2), \quad i = 1, 2, \dots, 10 \\
Y &= \sum_{i=1}^{10} X_i \sim N\left(50, (3\sqrt{10})^2\right)
\end{aligned} \tag{3}$$

If the probabilistic performance measure is  $P[Y > 60]$ , then the reliability analysis result is

$$P[Y > 60] = 1 - \Phi\left(\frac{60 - 50}{3\sqrt{10}}\right) = 0.1459 \tag{4}$$

However, if one dimension is reduced by screening out  $X_{10} = \mu_{10} = 5$  while the other variables remain random, then the probabilistic performance measure changes to

$$\tilde{Y} = \sum_{i=1}^9 X_i + 5 \sim N(50, 9^2) \tag{5}$$

As a consequence, the reliability analysis result yields

$$P[\tilde{Y} > 60] = 1 - \Phi\left(\frac{60 - 50}{9}\right) = 0.1333 \tag{6}$$

From Eqs.(4) and (6), 0.0126 (1.26%) of the output variability is decreased by screening out one variable. A more fundamental problem is that the lost amount 1.26% cannot be estimated without the full-dimensional

reliability analysis result of Eq.(4). On the other hand, let's assume that  $X_{10}$  has a smaller variance of one. Then, the full-dimensional reliability analysis yields

$$P(Y > 60) = 1 - \Phi\left(\frac{60-50}{\sqrt{82}}\right) = 0.1347 \quad (7)$$

From Eqs.(6) and (7), the difference is 0.0014 (0.14%), which could be acceptable. Therefore, in this case,  $X_{10}$  could be fixed at the mean value. As shown in the example, the output variability decreases if any random variable is fixed at a deterministic value. However, there are some variables that affect the output variability a small amount. The variable screening method for effective surrogate models for RBDO is to find those variables that have small effects on the output variability. It is noted that the random parameters are considered as much as the random design variables in this paper. Even though the random parameters are not changing during the RBDO process, they will influence the output variability. Hence they should be considered in the variable screening process, so that reliability of the performance measure can be accurately approximated using reduced dimension.

## 5. Variable Screening with 1-D Surrogate Model

The output variability can be quantified by the output variance. The exact output variance of a nonlinear implicit performance measure is very difficult to obtain. Hence, an approximated output variance is used in this paper. In the following sections, the output variance is decomposed into partial output variances, which are the output variances when each input variable is random and the others are fixed at their mean values. Then, a method to find the design variables that have a large impact on output variance is developed using a hypothesis testing.

### 5.1. Approximated Output Variance

A univariate DRM is a well-known approximation method for statistical moments using multiple 1-D integrations [32]. Consider a performance measure  $Y$  and its realization  $y$  subject to random input vector  $\mathbf{X}=\{X_1, \dots, X_N\}^T$ :

$$Y(\mathbf{X}) = Y(X_1, \dots, X_N), \quad y(\mathbf{x}) = y(x_1, \dots, x_N) \quad (8)$$

Define a function  $Y_i$ , which is the performance measure when  $X_i$  is random and other variables are fixed at their mean values, as

$$Y_i = Y(\mu_1, \dots, \mu_{i-1}, X_i, \mu_{i+1}, \dots, \mu_N) \quad (9)$$

The realization of the performance measure at the input mean point  $\boldsymbol{\mu}_x$  is defined as

$$y_0 = y(\boldsymbol{\mu}_x) \quad (10)$$

The  $l^{\text{th}}$  statistical moment of  $Y$ , which is approximated using the univariate DRM, is defined as [32]

$$m_l \cong E \left[ \left\{ \sum_i Y_i - (N-1)y_0 \right\}^l \right] \quad (11)$$

Then, the output variance  $\sigma_Y^2$  can be approximated as

$$\sigma_Y^2 \cong m_2 - m_1^2 = \sum_i \sigma_{Y_i}^2 + \sum_{i>j} \rho_{Y_i Y_j} \sigma_{Y_i} \sigma_{Y_j} \quad (12)$$

where  $\sigma_{Y_i}^2$  is the variance of Eq.(9), which is the partial output variance when only  $X_i$  is random, and  $\rho_{Y_i Y_j}$  is the correlation coefficient between  $Y_i$  and  $Y_j$ .

As shown in Eq.(12), the partial output variances  $\sigma_{Y_i}^2$  are the main variables for approximating the output variance  $\sigma_Y^2$ . When  $\sigma_{Y_i}^2$  is larger than other partial output variances, it takes the largest portion in the output variance  $\sigma_Y^2$ . Therefore, if some  $X_i$  produces larger partial output variance than others, then  $X_i$  should be

selected as an important variable. Even though there are correlation terms  $\rho_{Y_j} \sigma_{Y_j} \sigma_{Y_i}$  in Eq.(12), their effect is less than the partial output variances. Moreover, when we are looking for important variables, not the approximated output variance  $\sigma_Y^2$  itself, we can focus on partial output variances. It is noted that calculation of  $\sigma_{Y_i}^2$  requires only 1-D integration, and thus, only 1-D surrogate models are required.

The partial output variance of  $\sigma_{Y_i}^2$  is like LSA because it can have different values at different input mean point  $\boldsymbol{\mu}_X$ , which is the current design point in the RBDO process. Hence, the variable screening result could be changed as the design point changes. There are several recommended points at which to perform variable screening using LSA. The first one is the DDO optimum. As the DDO optimum is usually close to the RBDO optimum, the variable screening result at the DDO optimum is likely to be similar to the result at the RBDO optimum. Also, the design point where most of the deterministic constraints are active can be a good candidate point. It is noted that DDO or the design point where the constraints are active could be obtained using the finite difference method in a practical engineering problem. Also, DDO could be achieved using the sensitivity obtained from a 1-D surrogate model because DDO requires only the deterministic LSA, which is 1-D.

## 5.2. Variable Screening Using Hypothesis Testing

Using the 1-D surrogate model, the partial output variance  $\sigma_{Y_i}^2$  can be calculated approximately as

$$s_{Y_i}^2 = \frac{1}{ns-1} \sum_{j=1}^{ns} \{y_i(\mu_1, \dots, x_i^{(j)}, \dots, \mu_N) - \bar{y}_i\}^2 \quad (13)$$

where  $x_i^{(j)}$  is the  $j^{\text{th}}$  realization of the input random variable  $X_i$ ,  $ns$  is the number of samples, and  $\bar{y}_i$  is the mean of  $y_i$  as,

$$\bar{y}_i = \frac{1}{ns} \sum_{j=1}^{ns} y_i(\mu_1, \dots, x_i^{(j)}, \dots, \mu_N) \quad (14)$$

As explained in previous sections, the partial output variance  $s_{Y_i}^2$  can be used to determine important design variables. To make the variable screening procedure systematic, hypothesis testing is applied in this paper. Hypothesis testing can prevent undesirable choices that could occur during the decision-making procedure. Calculated partial output variance  $s_{Y_i}^2$  depends on the number of samples  $ns$ . When  $ns$  is large enough, the variable screening result will be accurate. However, it would require a large computational time. Also, it is hard to determine when  $ns$  is large enough. When  $ns$  is small, it will include statistical error. If calculated  $s_{Y_i}^2$  are distinctive from each other or with respect to the screening threshold value, then the effect of  $ns$  may not be significant. However,  $ns$  could cause an error when some  $s_{Y_i}^2$  are similar to each other or are near the screening threshold value. Hypothesis testing can prevent this problem in a statistical manner by letting users control the error level.

Various hypothesis testing methods have been developed for the decision-making problem [33]. Among those methods, we need the one that is not sensitive to distribution type because the distribution type of  $Y_i$  or  $s_{Y_i}^2$  is not known in general. The one-sample  $t$ -test is developed based on the central limit theorem, which states that the sample mean of non-normal distribution follows normal distribution approximately for a large number of samples. The one-sample  $t$ -test is not sensitive to underlying distribution types, so it is used in this paper. As the  $t$ -test is a method for sample mean,  $s_{Y_i}^2$  is calculated  $nr$  times for its statistical moments as

$$\bar{v}_i = \frac{1}{nr} \sum_{k=1}^{nr} s_{Y_i}^{2(k)} \quad (15)$$

$$s_{v_i}^2 = \frac{1}{nr-1} \sum_{k=1}^{nr} (s_{Y_i}^{2(k)} - \bar{v}_i)^2 \quad (16)$$

where  $s_{Y_i}^{2(k)}$  is the  $k^{\text{th}}$  repetition of  $s_{Y_i}^2$  and  $nr$  is the number of repetitions. Now, the hypothesis is constructed:

$$H_0 : \bar{v}_i \leq \mu_0 \text{ versus } H_1 : \bar{v}_i > \mu_0 \quad (17)$$

where  $\mu_0$  is the criterion of hypothesis testing. According to Eq.(17), the design variable that corresponds to  $\bar{v}_i$ , which is greater than  $\mu_0$  ( $H_1$  is true), will be selected as an important variable. Using the one-sample  $t$ -test, the hypothesis can be tested by checking the following statement:

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } q \geq t_{nr-1,1-\alpha} \quad (18)$$

where  $\alpha$  is the significance level,  $t_{nr-1,\bullet}$  is  $t_{nr-1}^{-1}(\bullet)$ , and the test statistics  $q$  is defined as

$$q \equiv (\bar{v}_i - \mu_0) \Big/ \frac{s_{v_i}}{\sqrt{nr}} \quad (19)$$

In Eqs.(15) and (16), the uncertainty induced by  $ns$  is transferred to  $nr$ . Hence,  $ns$  can be a fixed number, whereas  $nr$  should be decided appropriately. Also,  $\mu_0$  needs to be identified in Eq.(17).  $\mu_0$  is the key criterion that decides important variables, and it should be a value relative to  $\bar{v}_i$  because the relative difference of partial output variances should be checked for variable screening. At the same time,  $\mu_0$  needs to be statistically independent from  $\bar{v}_i$  for reasonable hypothesis testing. In this paper, preliminary testing is proposed to obtain reasonable  $nr$  and  $\mu_0$  as follows. First, choose  $nr_0$ , which is large enough so that the central limit theorem holds. Then, calculate the initial statistical moments of  $s_{v_i}^2$  as

$$\bar{v}_i^{(0)} = \frac{1}{nr_0} \sum_{k=1}^{nr_0} s_{v_i}^{2(k)} \quad (20)$$

$$s_{v_i}^{2(0)} = \frac{1}{nr_0 - 1} \sum_{k=1}^{nr_0} (s_{v_i}^{2(k)} - \bar{v}_i^{(0)})^2 \quad (21)$$

Using the value from Eq.(20), the testing criterion  $\mu_0$  relative to  $\bar{v}_i$  can be calculated as

$$\mu_0 = \frac{\gamma}{N} \sum_{i=1}^N \bar{v}_i^{(0)} \quad (22)$$

where  $\gamma$  is a constant that the user selects.  $nr$  is calculated by limiting type II error ( $H_0$  is accepted when  $H_1$  is true) at the level of false negative rate  $\beta$  as [33]

$$nr = \max \left( \frac{s_{v_i}^2 (t_{nr_0-1,1-\alpha} + t_{nr_0-1,1-\beta})}{\bar{v}_i^{(0)} - \mu_0}, nr_0 \right) \quad (23)$$

In Eq.(23),  $t_{nr-1,\bullet}$  should be used instead of  $t_{nr_0-1,\bullet}$  for accurate calculation of  $nr$ . However, Eq.(23) requires the value of  $nr$  on the right side to calculate  $nr$ . To avoid this problem,  $t_{nr_0-1,\bullet}$  is used instead, and  $t_{nr_0-1,\bullet}$  produces a conservative result as it is larger than  $t_{nr-1,\bullet}$  because  $nr$  is larger than  $nr_0$  in Eq.(23) and  $\alpha$  is small. Finally,  $nr$  and  $\mu_0$  are determined so that the proposed hypothesis testing can be utilized.

### 5.3. 1-D Surrogate Model

In previous sections, the 1-D surrogate model is treated as the given one because it is not difficult to generate. However, efficiently creating a 1-D surrogate model could be an issue. For efficiency, quadratic interpolation is proposed as a basic 1-D surrogate model in this paper. Quadratic interpolation may not be an adequate method for creating a surrogate model for a highly nonlinear performance measure. However, a nonlinear performance measure can be effectively approximated by a quadratic function on a small region. If  $X$  follows normal distribution, the domain of  $X$  is  $(-\infty, \infty)$ , whereas 99.73% of  $X$  is in  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$ , which is much smaller than the infinite domain. Even if  $X$  does not follow normal distribution, the region  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$  can cover almost all (approximately 98%) of  $X$ . In view of the fact that this paper is focused on calculation of partial output variance, the region  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$  is large enough. Hence, the 1-D surrogate model needs to be

accurate in the region  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$  so that quadratic interpolation could be an appropriate method to approximate the performance measure in the region.

Quadratic interpolation requires three samples, and the location of samples affects the accuracy of interpolation. The location of samples determined using the Chebyshev polynomial is known to give uniform error in the domain [34]. Because only the region  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$  is interested, the location of samples is determined as  $x_1 = \mu_X - 2.5981\sigma_X$ ,  $x_2 = \mu_X$  and  $x_3 = \mu_X + 2.5981\sigma_X$  using the Chebyshev polynomial. Since a random variable  $X$  may not be evenly distributed in its domain, providing uniform error does not necessarily mean that the calculated partial output variance is accurate. However, since no unique location of samples is best for accurate partial output variance, the sample location by the Chebyshev polynomial is used in this paper due to the fact that it yields reasonable result for various distribution types of random variable  $X$ . If the random variable  $X$  has a closed and bounded domain like  $[a, b]$ , the domain can be directly used for calculation of partial output variance, and the location of samples are  $x_1 = 0.93301a + 0.06699b$ ,  $x_2 = (a+b)/2$  and  $x_3 = 0.06699a + 0.93301b$ , using Chebyshev polynomials.

To check the performance of a selected location of samples, a nonlinear performance measure  $Y$  is used as

$$Y(X) = 0.3 + \sin(16X / 15 - 0.7) + \sin^2(16X / 15 - 0.7) \quad (24)$$

Assuming that random variable  $X$  follows  $N(0.5, 0.333^2)$ , three locations of samples are chosen to compare the accuracy of the partial output variance. The first location is  $\{0.167, 0.5, 0.833\}$ , which is  $\mu_X$  and  $\mu_X \pm \sigma_X$ , and the second location is from the Chebyshev polynomial as  $\{-0.365, 0.5, 1.365\}$ . The third location is wider, as  $\{-0.667, 0.5, 1.667\}$ , which is  $\mu_X$  and  $\mu_X \pm 3.5\sigma_X$ . Partial output variances are calculated using 100,000 realizations of  $X$ , and true partial output variance is calculated by Eq.(24) with the same realizations. To check the accuracy of the quadratic interpolation itself, mean square error (MSE) is calculated in the region of  $(-0.5, 1.5)$ , which is  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$  with 100 uniformly distributed points. The calculated result is shown in Table 1, and the shape of quadratic interpolations is shown in Figure 1, where asterisk marks (\*) represent the sample point. As shown in Table 1, the location of the sample using Chebyshev polynomials produces more accurate partial output variance compared to the true one and less MSE than the other cases.

Table 1: Quadratic Interpolation with Different Sample Locations

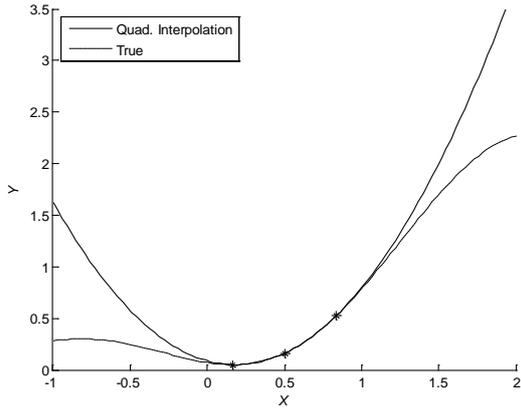
Location of samples		Quadratic interpolation			True
		$\{0.167, 0.5, 0.833\}$	$\{-0.365, 0.5, 1.365\}$	$\{-0.667, 0.5, 1.667\}$	
Partial output variance	Value	8.74E-02	7.84E-02	6.92E-02	7.84E-02
	Accuracy	111.53%	100.01%	88.33%	100.00%
MSE		1.17E-02	1.19E-03	9.00E-03	0

This example cannot represent all performance measures. When a highly nonlinear performance measure is expected, more sophisticated surrogate methods, such as the radial basis function (RBF), polynomial response surface (PRS), support vector regression (SVR), Kriging, and dynamic Kriging [2-10], are better. In any case, it is recommended to sample inside the region of  $(\mu_X - 3\sigma_X, \mu_X + 3\sigma_X)$  for the random variable  $X$  if the distribution has an infinite domain.

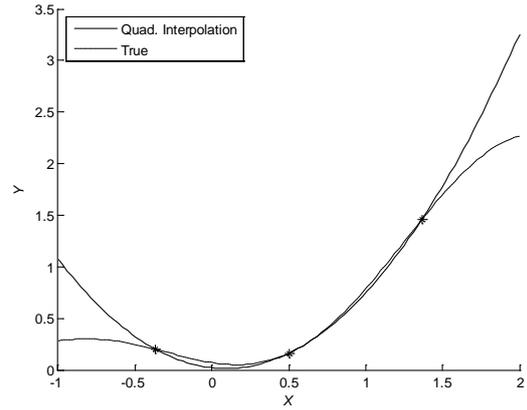
## 6. Numerical Examples

A car noise, vibration, and harshness (NVH) and crash safety problem is considered to demonstrate the performance and efficiency of the proposed method. The problem includes full frontal impact, 40% offset frontal impact, and NVH as constraints. There are a total of 11 performance measures as shown in Table 2: nine safety measures and two NVH measures.

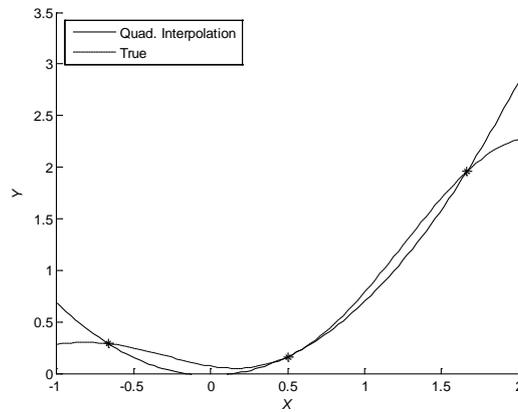
The 44 random variables shown in Table 3 are used to represent thicknesses of the body structure considered in the problem. All random variables follow normal distribution and are statistically independent. The design variable vector  $\mathbf{d}$  is the mean vector of the 44 random variables and there is no random parameter in this example. Among those random variables, six random variables ( $X_1$  to  $X_5$  and  $X_8$ ) are common variables for both safety and NVH measures, two ( $X_6$  and  $X_7$ ) are variables only for safety, and the other 36 random variables are only for NVH measures.



(a) Location of sample:  $\{0.167, 0.5, 0.833\}$



(b) Location of sample:  $\{-0.365, 0.5, 1.365\}$



(c) Location of sample:  $\{-0.667, 0.5, 1.667\}$

Figure 1: Quadratic Interpolation of  $Y$  with Different Locations of Samples

Table 2: Performance Measure Description

Mode		Function	Value	Feasibility decision
Safety	Full frontal impact	$G_1$	Chest G	$\leq Baseline_i$
		$G_2$	Crush displacement	
		$G_3$	Brake pedal	
	40% offset impact	$G_4$	Footrest	
		$G_5$	Left toepan	
		$G_6$	Center toepan	
		$G_7$	Right toepan	
		$G_8$	Left IP	
		$G_9$	Right IP	
NVH		$G_{10}$	Torsion mode	
		$G_{11}$	Vertical bending mode	

Table 3: Input Random Variables

RVs	Dist. Type	$\mathbf{d}$	STDV	$\mathbf{d}^L$	$\mathbf{d}^U$	RVs	Dist. Type	$\mathbf{d}$	STDV	$\mathbf{d}^L$	$\mathbf{d}^U$
$X_1$	Normal	1.9	0.05	1.5	2.3	$XN_1$	Normal	0.9	0.03	0.7	1.1
$X_2$	Normal	1.91	0.05	1.5	2.3	$XN_2$	Normal	1.1	0.03	0.8	1.4
$X_3$	Normal	2.51	0.06	2.0	3.0	$XN_3$	Normal	1.55	0.05	1.2	1.9
$X_4$	Normal	2.4	0.06	1.9	2.9	$XN_4$	Normal	0.9	0.03	0.7	1.1
$X_5$	Normal	2.55	0.06	2.0	3.1	$XN_5$	Normal	1.5	0.03	1.2	1.8
$X_6$	Normal	2.25	0.06	1.8	2.7	$XN_6$	Normal	1.2	0.03	0.9	1.5
$X_7$	Normal	2.25	0.06	1.8	2.7	$XN_7$	Normal	1.1	0.03	0.8	1.4
$X_8$	Normal	1.5	0.03	1.2	1.8	$XN_8$	Normal	1.52	0.05	1.2	1.9
$X_{10}$	Normal	1.28	0.03	0.9	1.6	$XN_9$	Normal	0.8	0.03	0.6	1.0
$X_{11}$	Normal	1.4	0.03	1.0	1.8	$XN_{10}$	Normal	0.8	0.03	0.6	1.0
$X_{12}$	Normal	1.1	0.03	0.8	1.4	$XN_{11}$	Normal	1.2	0.03	0.9	1.5
$X_{13}$	Normal	2.2	0.06	1.7	2.7	$XN_{12}$	Normal	0.75	0.03	0.6	0.9
$X_{14}$	Normal	1.5	0.03	1.2	1.8	$XN_{13}$	Normal	0.75	0.03	0.6	0.9
$X_{15}$	Normal	1.25	0.03	0.9	1.6	$XN_{14}$	Normal	0.75	0.03	0.6	0.9
$X_{16}$	Normal	2.5	0.06	2.0	3.0	$XN_{15}$	Normal	1.0	0.03	0.8	1.2
$X_{17}$	Normal	2.0	0.05	1.5	2.5	$XN_{16}$	Normal	1.14	0.03	0.9	1.4
$X_{18}$	Normal	1.4	0.03	1.1	1.7	$XN_{17}$	Normal	1.2	0.03	0.9	1.5
$X_{20}$	Normal	1.22	0.03	0.9	1.5	$XN_{18}$	Normal	1.4	0.03	1.1	1.7
$X_{23}$	Normal	0.75	0.03	0.6	1.0	$XN_{19}$	Normal	1.2	0.03	0.9	1.5
$X_{24}$	Normal	1.9	0.05	1.5	2.3	$XN_{20}$	Normal	1.4	0.03	1.1	1.7
$X_{25}$	Normal	0.65	0.03	0.5	0.8	$XN_{21}$	Normal	2.13	0.06	1.7	2.6
$X_{26}$	Normal	0.85	0.03	0.6	1.1						
$X_{27}$	Normal	0.85	0.03	0.6	1.1						

This problem requires three and a half hours for the impact dynamic analysis for crash safety and the modal analysis for NVH. Thus, the actual analysis takes too much time to test the proposed method thoroughly. Ford Motor Company provided full-dimensional global (considering the entire design domain) surrogate models so that we can use them to demonstrate the proposed method of variable screening. The full-dimensional surrogate models may not be accurate, since 44-D is too high to create accurate surrogate models, especially for RBDO. However, to test the proposed method of variable screening, the responses from the 44-D global surrogate models are treated as true responses in this paper. The maximum dimension at which accurate surrogate models can be generated depends on the computational power and nonlinearity of a given problem. In this paper, the dynamic Kriging (DKG) method [10] is used to generate an accurate surrogate model using the Iowa Reliability-Based Design Optimization (I-RBDO) code [35], and 18-D is targeted as the maximum degrees of freedom of DKG models. The I-RBDO code is also used to carry out RBDO in this paper.

### 6.1. Variable Screening

At the baseline design  $\mathbf{d}$ , which is the initial design as shown in Table 3, all 11 performance measures in Table 2 are active. That is, the value of every performance measure at the baseline design is the same as the baseline values, with  $G_i = \text{Baseline}_i$ ,  $i=1\sim 11$ . Therefore, the proposed variable screening method is performed for the problem at the baseline design. Parameters  $ns$ ,  $nr_0$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$  are set as 50, 50, 0.05, 0.05, and 1.0, respectively. Four hundred eighty-four (44 design variables  $\times$  11 performance measures) 1-D surrogate models with quadratic interpolation are generated using 89 analyses (i.e. simulation samples). It is noted that 11 values of performance measures are obtained from one analysis. The results of partial output variances  $\bar{v}_i$  are listed in Table 4 and Table 5 for every performance measure. The output variances of the important variables for each performance measure are marked with bold font. It is noted that only partial output variances of  $X_1\sim X_8$  are listed in Table 4 since  $G_1\sim G_9$  are only a function of  $X_1\sim X_8$  as the variable screening method identified the partial output variances to be zero for other random variables. In Table 5,  $X_6$  and  $X_7$  have zero partial output variances as  $G_{10}$  and  $G_{11}$  are not functions of  $X_6$  and  $X_7$ . In the last three rows of Table 4 and Table 5, the sums of partial output variances of selected variables; the sums of all partial output variances; and their ratios are listed. As all design variables are statistically independent of each other in this problem, this is the ratio between the output variance using selected variables and the total output variance. A minimum of 90.45% of the total output variance is captured by the output variances of selected variables. In total, 14 random variables:  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$ ,  $X_5$ ,  $X_6$ ,  $X_7$ ,  $X_8$ ,  $X_{10}$ ,  $X_{20}$ ,  $X_{23}$ ,  $X_{25}$ ,  $X_{26}$ , and  $XN_1$ , are selected as important variables. Accordingly, 14 design variables, which are the means of the selected random variables, are considered as important design variables.

Table 4: Partial Output Variances  $\bar{v}_i$  ( $G_1 \sim G_9$ )

RVs	$G_1$	$G_2$	$G_3$	$G_4$	$G_5$	$G_6$	$G_7$	$G_8$	$G_9$
$X_1$	<b>2.37E-02</b>	<b>3.75E+00</b>	<b>4.06E+01</b>	<b>1.11E+01</b>	<b>1.94E+00</b>	<b>1.73E+00</b>	<b>6.21E-01</b>	<b>3.45E-01</b>	<b>5.75E-01</b>
$X_2$	<b>1.88E-02</b>	<b>1.01E+00</b>	3.55E-01	<b>1.86E+00</b>	<b>2.07E+00</b>	<b>2.06E+00</b>	<b>1.09E+00</b>	<b>4.03E-01</b>	<b>2.69E-01</b>
$X_3$	5.08E-05	1.76E-02	<b>1.05E+01</b>	<b>9.74E+00</b>	<b>5.95E+00</b>	<b>4.38E+00</b>	<b>2.00E+00</b>	<b>2.37E-01</b>	2.24E-02
$X_4$	1.66E-04	9.27E-02	<b>3.91E+00</b>	<b>1.79E+00</b>	2.65E-01	6.65E-02	1.73E-03	4.78E-03	9.65E-03
$X_5$	1.12E-04	1.37E-04	<b>2.74E+00</b>	<b>2.23E+00</b>	<b>4.90E-01</b>	<b>9.11E-01</b>	<b>9.30E-01</b>	<b>1.79E-01</b>	<b>7.94E-02</b>
$X_6$	4.58E-05	<b>2.77E+00</b>	1.64E-01	8.14E-02	<b>8.18E-01</b>	1.60E-01	<b>1.70E-01</b>	<b>3.95E-02</b>	<b>3.27E-02</b>
$X_7$	<b>2.32E-03</b>	1.34E-01	1.91E-01	1.50E-01	<b>2.75E+00</b>	<b>1.71E+00</b>	<b>5.66E-01</b>	<b>1.10E-01</b>	<b>2.77E-01</b>
$X_8$	<b>1.23E-03</b>	6.28E-02	4.39E-02	3.02E-01	<b>5.35E-01</b>	<b>5.45E-01</b>	<b>3.27E-01</b>	<b>6.25E-02</b>	<b>4.87E-02</b>
Selected	4.60E-02	7.53E+00	5.78E+01	2.67E+01	1.45E+01	1.13E+01	5.70E+00	1.38E+00	1.28E+00
All	4.64E-02	7.83E+00	5.85E+01	2.72E+01	1.48E+01	1.16E+01	5.70E+00	1.38E+00	1.31E+00
Selec./All	99.19%	96.07%	98.71%	98.04%	98.21%	98.04%	99.97%	99.65%	97.56%

Table 5: Partial Output Variances  $\bar{v}_i$  ( $G_{10}$  and  $G_{11}$ )

RVs	$G_{10}$	$G_{11}$	RVs	$G_{10}$	$G_{11}$	RVs	$G_{10}$	$G_{11}$
$X_1$	<b>1.30E-03</b>	<b>1.48E-03</b>	$X_{17}$	4.24E-07	2.01E-04	$XN_8$	2.06E-05	5.45E-05
$X_2$	4.39E-05	3.44E-05	$X_{18}$	9.71E-06	3.33E-05	$XN_9$	1.61E-06	1.89E-05
$X_3$	<b>8.95E-04</b>	3.05E-04	$X_{20}$	<b>2.21E-03</b>	<b>1.84E-03</b>	$XN_{10}$	5.11E-06	1.83E-06
$X_4$	8.74E-05	2.63E-05	$X_{23}$	<b>2.67E-03</b>	<b>1.13E-03</b>	$XN_{11}$	1.56E-07	1.12E-07
$X_5$	3.06E-04	4.23E-05	$X_{24}$	2.35E-05	6.89E-05	$XN_{12}$	1.83E-06	3.10E-05
$X_6$	0	0	$X_{25}$	<b>4.85E-04</b>	<b>7.31E-03</b>	$XN_{13}$	1.69E-07	8.19E-07
$X_7$	0	0	$X_{26}$	<b>4.59E-03</b>	<b>1.06E-02</b>	$XN_{14}$	9.08E-08	3.75E-05
$X_8$	<b>5.56E-04</b>	1.89E-04	$X_{27}$	3.43E-05	3.76E-04	$XN_{15}$	1.84E-06	1.66E-05
$X_{10}$	<b>5.49E-04</b>	3.87E-04	$XN_1$	9.33E-05	<b>6.38E-03</b>	$XN_{16}$	4.35E-08	2.88E-08
$X_{11}$	3.32E-05	2.05E-04	$XN_2$	4.49E-07	7.48E-05	$XN_{17}$	3.24E-07	1.84E-06
$X_{12}$	9.35E-05	3.98E-04	$XN_3$	9.59E-08	2.00E-05	$XN_{18}$	1.04E-07	1.18E-07
$X_{13}$	7.23E-06	3.82E-04	$XN_4$	6.04E-08	8.36E-05	$XN_{19}$	3.03E-08	3.37E-07
$X_{14}$	5.76E-07	1.17E-05	$XN_5$	2.60E-07	1.42E-07	$XN_{20}$	2.06E-07	2.98E-06
$X_{15}$	1.16E-07	1.13E-05	$XN_6$	1.29E-07	2.22E-07	$XN_{21}$	2.44E-07	9.41E-06
$X_{16}$	9.14E-08	1.63E-06	$XN_7$	9.46E-08	7.94E-06			
Selected							1.33E-02	2.88E-02
All							1.40E-02	3.18E-02
Selec./All							94.53%	90.45%

Table 6: Effect of Input Standard Deviation and Design Sensitivity to Output Variance of  $G_1$ 

RVs	STDV ( $\sigma_{x_i}$ )	Design Sensitivity of $G_1$ ( $\alpha_i$ )	Induced Approximate Variance of $G_1$ ( $\alpha_i^2 \sigma_{x_i}^2$ )
$X_1$	<b>0.05</b>	<b>-3.0461</b>	<b>2.32E-02</b>
$X_2$	<b>0.05</b>	<b>-2.7135</b>	<b>1.84E-02</b>
$X_3$	0.06	-0.1170	4.92E-05
$X_4$	0.06	-0.2118	1.61E-04
$X_5$	0.06	-0.1746	1.10E-04
$X_6$	0.06	-0.0944	3.21E-05
$X_7$	<b>0.06</b>	<b>0.7934</b>	<b>2.27E-03</b>
$X_8$	<b>0.03</b>	<b>1.1577</b>	<b>1.21E-03</b>

The output variance is a combination of design sensitivity and variance of the input random variable. When a performance measure is linear to input variables ( $G = \sum \alpha_i X_i$ ), the output variance of  $G$  is  $\sum \alpha_i^2 \sigma_{X_i}^2$ . Hence, an important variable can be selected based on its variance as well as design sensitivity. For example, selected random variables for the constraint  $G_1$  are  $X_1, X_2, X_7$ , and  $X_8$  out of  $X_1 \sim X_8$ . Among  $X_1 \sim X_8$ , variables  $X_3 \sim X_7$  have the largest standard deviation of 0.06. However, only  $X_7$  is selected among them. This is because screened-out variables have small design sensitivity compared to others, as shown in Table 6. Design sensitivities are calculated at the design point using quadratic interpolation models. The variables in bold are selected as they have larger sensitivities even though they have smaller standard deviations, so that they induce large output variances.

Table 7: Global Sensitivity Indices  $S_i^{\text{tot}}$  ( $G_1 \sim G_9$ )

RVs	$G_1$	$G_2$	$G_3$	$G_4$	$G_5$	$G_6$	$G_7$	$G_8$	$G_9$
$X_1$	<b>4.16E-01</b>	<b>3.84E-01</b>	<b>5.23E-01</b>	<b>2.68E-01</b>	<b>1.74E-01</b>	<b>1.89E-01</b>	<b>2.07E-01</b>	<b>2.56E-01</b>	<b>4.35E-01</b>
$X_2$	<b>3.46E-01</b>	<b>1.18E-01</b>	<b>1.04E-01</b>	<b>1.88E-01</b>	<b>1.98E-01</b>	<b>2.09E-01</b>	<b>1.66E-01</b>	<b>2.62E-01</b>	<b>2.07E-01</b>
$X_3$	7.83E-03	5.75E-03	<b>3.28E-01</b>	<b>5.07E-01</b>	<b>5.18E-01</b>	<b>4.97E-01</b>	<b>4.92E-01</b>	<b>2.69E-01</b>	<b>4.96E-02</b>
$X_4$	1.89E-02	<b>2.23E-02</b>	<b>7.92E-02</b>	<b>1.34E-01</b>	<b>1.25E-01</b>	<b>1.38E-01</b>	<b>1.42E-01</b>	<b>7.38E-02</b>	<b>2.87E-02</b>
$X_5$	7.50E-03	-7.22E-04	<b>1.40E-01</b>	<b>2.10E-01</b>	<b>1.53E-01</b>	<b>1.76E-01</b>	<b>1.78E-01</b>	<b>1.90E-01</b>	<b>9.10E-02</b>
$X_6$	<b>5.65E-02</b>	<b>2.67E-01</b>	2.08E-02	<b>5.01E-02</b>	<b>1.11E-01</b>	<b>1.15E-01</b>	<b>1.67E-01</b>	<b>5.13E-02</b>	<b>2.74E-02</b>
$X_7$	<b>8.13E-02</b>	5.65E-03	1.77E-02	<b>4.52E-02</b>	<b>1.59E-01</b>	<b>1.60E-01</b>	<b>1.76E-01</b>	<b>6.73E-02</b>	<b>1.73E-01</b>
$X_8$	<b>3.23E-02</b>	<b>2.06E-02</b>	4.02E-03	2.33E-02	<b>3.82E-02</b>	<b>4.09E-02</b>	<b>5.21E-02</b>	<b>7.13E-02</b>	<b>6.14E-02</b>
Selected	9.32E-01	8.12E-01	1.17E+00	1.40E+00	1.48E+00	1.52E+00	1.58E+00	1.24E+00	1.07E+00
All	9.66E-01	8.23E-01	1.22E+00	1.43E+00	1.48E+00	1.52E+00	1.58E+00	1.24E+00	1.07E+00
Selec./All	96.46%	98.70%	96.50%	98.36%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 8: Global Sensitivity Indices  $S_i^{\text{tot}}$  ( $G_{10}$  and  $G_{11}$ )

RVs	$G_{10}$	$G_{11}$	RVs	$G_{10}$	$G_{11}$	RVs	$G_{10}$	$G_{11}$
$X_1$	<b>1.05E-01</b>	<b>7.09E-02</b>	$X_{17}$	3.30E-04	1.46E-02	$XN_8$	-8.85E-04	-1.34E-03
$X_2$	1.01E-03	2.38E-03	$X_{18}$	-4.77E-04	-2.50E-04	$XN_9$	8.54E-04	-1.71E-03
$X_3$	<b>5.72E-02</b>	1.01E-02	$X_{20}$	<b>2.07E-01</b>	<b>9.17E-02</b>	$XN_{10}$	1.28E-04	-1.07E-04
$X_4$	8.66E-03	2.65E-03	$X_{23}$	<b>1.34E-01</b>	<b>3.57E-02</b>	$XN_{11}$	-2.72E-04	-2.85E-05
$X_5$	<b>3.29E-02</b>	5.69E-03	$X_{24}$	-1.60E-03	-8.39E-04	$XN_{12}$	-1.51E-04	-1.94E-04
$X_6$	0	0	$X_{25}$	1.12E-02	<b>9.79E-02</b>	$XN_{13}$	-5.24E-05	1.59E-04
$X_7$	0	0	$X_{26}$	<b>2.96E-01</b>	<b>3.80E-01</b>	$XN_{14}$	-3.88E-05	-6.91E-05
$X_8$	<b>4.73E-02</b>	5.89E-03	$X_{27}$	6.04E-03	<b>2.38E-02</b>	$XN_{15}$	1.20E-04	4.59E-04
$X_{10}$	<b>5.98E-02</b>	1.99E-02	$XN_1$	5.62E-03	<b>1.65E-01</b>	$XN_{16}$	8.87E-05	2.85E-04
$X_{11}$	3.86E-03	1.67E-02	$XN_2$	-1.11E-04	2.31E-03	$XN_{17}$	-6.09E-05	-3.61E-04
$X_{12}$	1.34E-02	<b>2.88E-02</b>	$XN_3$	1.13E-04	1.33E-03	$XN_{18}$	-1.12E-05	-9.18E-05
$X_{13}$	3.20E-04	1.50E-02	$XN_4$	1.54E-04	-2.50E-03	$XN_{19}$	1.10E-04	1.23E-03
$X_{14}$	2.24E-04	-4.65E-04	$XN_5$	-2.96E-04	9.10E-04	$XN_{20}$	-1.67E-04	-8.38E-04
$X_{15}$	-1.33E-04	1.36E-03	$XN_6$	2.34E-04	-2.32E-05	$XN_{21}$	6.77E-05	-6.44E-04
$X_{16}$	-1.94E-04	2.42E-03	$XN_7$	-1.61E-05	-3.27E-04			
Selected							9.40E-01	8.94E-01
All							9.88E-01	9.88E-01
Selec./All							95.13%	90.52%

Since we have 44-D global surrogate models for this example, GSA can be carried out to verify effectiveness of the proposed method. Among various GSA methods, the global sensitivity index method, which identifies effect of interested variable to output on the domain, is used here. The main strength of the global sensitivity index method is that it can find interactions between all variables (not statistical correlation between random variables). All random variables are assumed to follow uniform distribution in their corresponding design domain of  $\mathbf{d}^L$  and  $\mathbf{d}^U$ , and global sensitivity indices are calculated using the Monte Carlo simulation (MCS) method with 1 million

MCS samples [29]. There are many global sensitivity indices in this 44-D problem; the total sensitivity index  $S_i^{\text{tot}}$  is used for variable ranking and screening. The total sensitivity index  $S_i^{\text{tot}}$  is a “total influence of  $i^{\text{th}}$  random variable” to the output. That is, it indicates the main effect plus interactions of  $i^{\text{th}}$  random variable with other random variables [20]. The results are listed in Table 7 and Table 8. To identify important random variables, the mean value of  $S_i^{\text{tot}}$  is calculated for each constraint and the random variable, which yields larger  $S_i^{\text{tot}}$  than the mean value, is selected as an important variable and marked in bold font in these tables. In Table 7, only  $S_i^{\text{tot}}$  for  $X_1 \sim X_8$  are listed as  $S_i^{\text{tot}}$  for other variables are zero. Also, the sum of  $S_i^{\text{tot}}$  for the selected random variables, the sum of  $S_i^{\text{tot}}$  for all random variables, and their ratios are listed in the last three rows, respectively.

Using the global sensitivity index method, 16 random variables selected as shown in Table 9. Those 16 random variables include all 14 random variables selected using the proposed method as shown in Table 9. Moreover, if we limit to 14 random variables to be selected,  $X_{12}$  and  $X_{27}$  will not be selected as they have least  $S_i^{\text{tot}}$  among the selected variables for  $G_{11}$  as shown in Table 8. Thus, the selected 14 random variables by both methods are identical. The ratio between the sensitivity indices of selected variables and all variables has no physical meaning. However, it is an indicator that shows how much variance is captured by the selected random variables. The results are quite similar to those of the proposed method as shown in Tables 4 and 5 and Tables 7 and 8, respectively. Hence, it is demonstrated that the proposed variable screening method is quite effective even though it does not require global surrogate models like the global sensitivity index method.

Table 9: Selected Random Variables

Method	Selected Variables
Proposed Method	$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_{10}, X_{20}, X_{23}, X_{25}, X_{26}, XN_1$ (14 RVs)
Global Sensitivity Index	$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_{10}, X_{12}, X_{20}, X_{23}, X_{25}, X_{26}, X_{27}, XN_1$ (16 RVs)

## 6.2 Reliability-based Design Optimization

For this example, RBDO is formulated as

$$\begin{aligned}
 & \text{minimize} && \text{Weight } (\mathbf{d}) \\
 & \text{subject to} && P[G_i(\mathbf{X}) > \text{Baseline}_i] \leq 10\%, \quad i = 1, \dots, 11 \\
 & && \mathbf{d}^L \leq \mathbf{d} \leq \mathbf{d}^U, \quad \mathbf{d} \in \mathbb{R}^{NDV}, \quad \text{and} \quad \mathbf{X} \in \mathbb{R}^{NRV}
 \end{aligned} \tag{25}$$

For comparison study, we considered three cases: (1) a set of 14 random variables are selected based on experience without using the proposed variable screening method, (2) another set of 14 random variables are selected using the proposed variable screening method as shown in Section 4.1, and (3) in addition to 14 random variables selected in case 2, four more random variables are selected using the cost function sensitivity for the total of 18 design variables to test the effectiveness of the proposed variable screening method and accuracy of the I-RBDO code [10]. The selected design variables are listed in Table 10.

Table 10: Selected Random Variables for RBDO

Cases	Common Selection	Different Selection
(1) Based on Experience	$X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_{20}, X_{25}, XN_1$ (11 RVs)	$XN_9, XN_{10}, XN_{11}$
(2) Variable Screening		$X_{10}, X_{23}, X_{26}$
(3) Variable Screening + Cost Function		$X_{10}, X_{23}, X_{26} + XN_4, XN_9, XN_{10}, XN_{11}$

Because the cost function, which is weight in this problem, is a function of design variables  $\mathbf{d}$ , not random variables  $\mathbf{X}$ , the function is deterministic. Therefore, the design sensitivity of the cost function with respect to the design variable is calculated by the finite difference method (FDM), and the four design variables (and related random variables  $XN_4, XN_9, XN_{10}$ , and  $XN_{11}$ ) that show the largest sensitivity among the unselected

design variables are chosen. RBDO is carried out with three sets of selected random variables. The optimum design results are summarized in Table 11. The bold font indicates chosen design variables, and others are fixed at the baseline design value. Also, probabilities of failure, cost function values, and design iteration details are listed in Table 12. All RBDOs are carried out using I-RBDO code with 500,000 MCS samples [10]. For the three cases, I-RBDO generates DKG surrogate models, using the sample responses obtained from the 44-D global surrogate models, which are treated as true responses, for RBDO. Since I-RBDO is able to carry out RBDO using the surrogate models generated by other methods (i.e., the surrogate models generated by Ford in this example) without having information on how the surrogate models are generated, the full-dimensional RBDO is performed as well using the 44-D surrogate models. The 44-D RBDO result is treated as the true RBDO optimum and used for the purpose of validation of RBDO results obtained for the three reduced-dimensional cases.

Indeed, the optimum design values for  $d_1 \sim d_8$  are very close to the full-dimensional 44-D case, as shown in Table 11, which shows that the DKG surrogate models generated in the I-RBDO code are accurate. In the three cases, the random variables  $X_1 \sim X_8$  are selected because they have large partial output variance for performance measures  $G_1 \sim G_9$ . This means that they contribute a large portion of the output variance. Hence, finding optimum values for them is the most effective way to reduce the probabilities of failure of  $G_1 \sim G_9$ . Similarly,  $d_{22}$  (corresponding to  $X_{26}$ ) moves to the upper bound of 1.1 when it is selected because it has the largest partial output variances of  $G_{10}$  and  $G_{11}$ . The design variables  $d_{27}$ ,  $d_{32}$ ,  $d_{33}$ , and  $d_{34}$  (corresponding to  $X_{N_4}$ ,  $X_{N_9}$ ,  $X_{N_{10}}$ , and  $X_{N_{11}}$ ), which are selected by the sensitivity of the cost function, move to their lower bounds of 0.7, 0.6, 0.6, and 0.9, respectively, to reduce the cost function without significantly affecting the reliability of the optimum design. On the other hand, even though some design variables are selected due to the partial output variances of some constraints, they move to their lower bounds. For example,  $d_{24}$  (corresponding to  $X_{N_1}$ ) moves to the lower bound of 0.7 because it has the largest sensitivity for the cost function, even though it has the third-largest partial output variance of  $G_{11}$ . That is, via the trade-offs in the optimization process, it is moved to the lower bound to minimize the cost function rather than to reduce the probability of failure.

In Table 12, the number of black box calls is listed. I-RBDO can deal with multiple samples simultaneously to create a surrogate model to utilize parallel computing. That is, one black-box call requests computational simulations at a number of sampling points; and thus the black box call indicates the clock time required for analyses for RBDO. Therefore, the number of black box calls represents actual computational cost more realistically than the number of CAEs when parallel computing is used. In this example, five samples are added to the DKG model at a time; therefore, the total number of CAEs is roughly five times the number of black box calls.

To verify once again that the surrogate model generated by DKG using I-RBDO is accurate, the same three cases are performed using I-RBDO by using responses from 44-D global surrogate model directly while fixing screened out variables at their baseline design point. As shown in Table 13, the optimums found using DKG and 44-D global surrogate models are very close to each other. Hence, it is confirmed that DKG in I-RBDO generated an accurate surrogate model. Moreover, it is also verified that RBDO can be conducted based on an accurate surrogate model even for a moderately large-dimensional problem (14 and 18 dimensions).

For the three reduced-dimensional cases, the probabilities of failure are calculated using only selected variables as random variables, since the other design variables are treated as deterministic as explained in Section 2.2 with fixed values at the baseline design. As shown in Table 12, it is noted that, in all cases, the target design constraints of 10% probability failure are closely satisfied as expected at these optimum designs since the RBDO considers only the selected variables as random variables. On the other hand, to check correct reliabilities, reliability analyses are carried out at these optimum designs treating all variables as random for the 44-D surrogate models and MCS with 1 million samples, as shown in Table 14. It is noted that probabilities of failure of full-dimensional optimum in Table 12 and Table 14 are different even though the discrepancy is negligible. Theoretically, they should be the same; however, they are not equal because different numbers of MCS samples (500,000 and 1 million) are used and MCS error is induced. At the baseline design, all constraints have approximately 50% of probability of failure and this is reasonable because all constraints are active at the baseline design. However, they are not exactly 50% because the constraint functions are nonlinear. Probabilistic constraint results corresponding to  $G_1 \sim G_9$  are active or feasible in both Table 12 and Table 14. Due to the fact that  $G_1$  to  $G_9$  are functions of  $X_1$  to  $X_8$ , and all of them are selected as important variables, the RBDO result with reduced-dimension and the reliability analysis result with full-dimension are very close to each other, considering MCS errors. The constraint  $G_{10}$  shows inactive results regardless of which selection design variable set is used.

All probabilities of failure for the constraint  $G_{11}$  in Table 12 satisfy the target probability of failure 10%, which makes it obvious that these are the reliability analyses results of reduced-dimensional problems. However, full-dimensional reliability analyses at optimum designs show quite different values. Selection based on experience shows 17.70% probability of failure, which violates the target reliability significantly. The variables selected

based on experience contain only 55.40% of the total output variance of  $G_{11}$ . Hence, it cannot find any safe design once dimension is reduced. On the other hand, the probabilities of failure for the proposed variable screening method and also the one considering cost function are close to the target probability of failure. The selected variables contain 93.55% and 93.88% of the total output variance of  $G_{11}$  for Case 2 and Case 3, respectively. Hence it can find a correct optimum even with reduced-dimension.

Table 11: RBDO Optimum Design

Design Variables	Corresponding RVs	Baseline Design	Based on Experience	Variable Screening	Variable Screening + Cost Function	Full Dimension
$d_1$	$X_1$	1.9	<b>1.8343</b>	<b>1.8366</b>	<b>1.8336</b>	<b>1.8491</b>
$d_2$	$X_2$	1.91	<b>2.1810</b>	<b>2.1804</b>	<b>2.1806</b>	<b>2.1692</b>
$d_3$	$X_3$	2.51	<b>2.8528</b>	<b>2.8561</b>	<b>2.8540</b>	<b>2.8803</b>
$d_4$	$X_4$	2.4	<b>1.9817</b>	<b>1.9810</b>	<b>1.9856</b>	<b>1.9507</b>
$d_5$	$X_5$	2.55	<b>2.7195</b>	<b>2.7228</b>	<b>2.7261</b>	<b>2.7296</b>
$d_6$	$X_6$	2.25	<b>2.2543</b>	<b>2.2497</b>	<b>2.2558</b>	<b>2.2408</b>
$d_7$	$X_7$	2.25	<b>2.3199</b>	<b>2.3185</b>	<b>2.3207</b>	<b>2.3284</b>
$d_8$	$X_8$	1.5	<b>1.7904</b>	<b>1.7966</b>	<b>1.7860</b>	<b>1.8</b>
$d_9$	$X_{10}$	1.28	1.28	<b>0.9</b>	<b>0.9</b>	<b>1.5336</b>
$d_{10}$	$X_{11}$	1.4	1.4	1.4	1.4	<b>1.0038</b>
$d_{11}$	$X_{12}$	1.1	1.1	1.1	1.1	<b>0.9868</b>
$d_{12}$	$X_{13}$	2.2	2.2	2.2	2.2	<b>1.7006</b>
$d_{13}$	$X_{14}$	1.5	1.5	1.5	1.5	<b>1.2</b>
$d_{14}$	$X_{15}$	1.25	1.25	1.25	1.25	<b>0.9</b>
$d_{15}$	$X_{16}$	2.5	2.5	2.5	2.5	<b>2.0</b>
$d_{16}$	$X_{17}$	2.0	2.0	2.0	2.0	<b>1.5033</b>
$d_{17}$	$X_{18}$	1.4	1.4	1.4	1.4	<b>1.1</b>
$d_{18}$	$X_{20}$	1.22	<b>0.9</b>	<b>0.9</b>	<b>0.9</b>	<b>0.9</b>
$d_{19}$	$X_{23}$	0.75	0.75	<b>0.6</b>	<b>0.6</b>	<b>0.6</b>
$d_{20}$	$X_{24}$	1.9	1.9	1.9	1.9	<b>1.5</b>
$d_{21}$	$X_{25}$	0.65	<b>0.6897</b>	<b>0.5429</b>	<b>0.5871</b>	<b>0.8</b>
$d_{22}$	$X_{26}$	0.85	0.85	<b>1.1</b>	<b>1.1</b>	<b>1.1</b>
$d_{23}$	$X_{27}$	0.85	0.85	0.85	0.85	<b>1.1</b>
$d_{24}$	$XN_1$	0.9	<b>0.7</b>	<b>0.7</b>	<b>0.7</b>	<b>0.7</b>
$d_{25}$	$XN_2$	1.1	1.1	1.1	1.1	<b>0.8</b>
$d_{26}$	$XN_3$	1.55	1.55	1.55	1.55	<b>1.2039</b>
$d_{27}$	$XN_4$	0.9	0.9	0.9	<b>0.7</b>	<b>0.7</b>
$d_{28}$	$XN_5$	1.5	1.5	1.5	1.5	<b>1.2</b>
$d_{29}$	$XN_6$	1.2	1.2	1.2	1.2	<b>0.9</b>
$d_{30}$	$XN_7$	1.1	1.1	1.1	1.1	<b>0.8</b>
$d_{31}$	$XN_8$	1.52	1.52	1.52	1.52	<b>1.6925</b>
$d_{32}$	$XN_9$	0.8	<b>0.6</b>	0.8	<b>0.6</b>	<b>0.6</b>
$d_{33}$	$XN_{10}$	0.8	<b>0.6</b>	0.8	<b>0.6</b>	<b>0.6</b>
$d_{34}$	$XN_{11}$	1.2	<b>0.9</b>	1.2	<b>0.9</b>	<b>0.9</b>
$d_{35}$	$XN_{12}$	0.75	0.75	0.75	0.75	<b>0.6</b>
$d_{36}$	$XN_{13}$	0.75	0.75	0.75	0.75	<b>0.6</b>
$d_{37}$	$XN_{14}$	0.75	0.75	0.75	0.75	<b>0.6</b>
$d_{38}$	$XN_{15}$	1.0	1.0	1.0	1.0	<b>0.8</b>
$d_{39}$	$XN_{16}$	1.14	1.14	1.14	1.14	<b>0.9083</b>
$d_{40}$	$XN_{17}$	1.2	1.2	1.2	1.2	<b>0.9</b>
$d_{41}$	$XN_{18}$	1.4	1.4	1.4	1.4	<b>1.1</b>
$d_{42}$	$XN_{19}$	1.2	1.2	1.2	1.2	<b>0.9</b>
$d_{43}$	$XN_{20}$	1.4	1.4	1.4	1.4	<b>1.1</b>
$d_{44}$	$XN_{21}$	2.13	2.13	2.13	2.13	<b>1.7</b>

Table 12: Cost, Probabilities of Failure at RBDO Optimum Design, and Optimization Details

Performance Measure	Based on Experience	Variable Screening	Variable Screening + Cost Function	Full Dimension
Cost	249.10	259.83	244.17	225.68
$G_1$	10.06%	9.94%	9.96%	9.96%
$G_2$	10.03%	10.11%	10.04%	9.99%
$G_3$	0.00%	0.00%	0.00%	0.00%
$G_4$	0.11%	0.11%	0.09%	0.10%
$G_5$	1.96%	1.95%	2.04%	1.91%
$G_6$	9.95%	10.01%	10.02%	9.99%
$G_7$	9.89%	9.93%	9.95%	10.01%
$G_8$	10.03%	9.98%	9.91%	9.07%
$G_9$	10.04%	9.99%	9.87%	9.92%
$G_{10}$	0.00%	0.00%	0.00%	0.00%
$G_{11}$	10.02%	9.97%	9.93%	10.02%
No. of Design Iter.	20	30	20	21
No. of CAEs	2,666	4,358	3,306	-
No. of Black Box Calls	473	727	573	-

Table 13: RBDO Optimum Design with I-RBDO and True Model

Design Variables	Corresponding RVs	Based on Experience		Variable Screening		Var. Screen. + Cost Fn.	
		I-RBDO	44-D Surrogate	I-RBDO	44-D Surrogate	I-RBDO	44-D Surrogate
$d_1$	$X_1$	1.8343	1.8371	1.8366	1.8359	1.8336	1.8425
$d_2$	$X_2$	2.1810	2.1799	2.1804	2.1807	2.1806	2.1771
$d_3$	$X_3$	2.8528	2.8578	2.8561	2.8576	2.8540	2.8654
$d_4$	$X_4$	1.9817	1.9807	1.9810	1.9851	1.9856	1.9525
$d_5$	$X_5$	2.7195	2.7222	2.7228	2.7233	2.7261	2.7209
$d_6$	$X_6$	2.2543	2.2488	2.2497	2.2501	2.2558	2.2464
$d_7$	$X_7$	2.3199	2.3173	2.3185	2.3169	2.3207	2.3265
$d_8$	$X_8$	1.7904	1.8	1.7966	1.7985	1.786	1.8
$d_9$	$X_{10}$	-	-	0.9	0.9	0.9	0.9
$d_{18}$	$X_{20}$	0.9	0.9	0.9	0.9	0.9	0.9
$d_{19}$	$X_{23}$	-	-	0.6	0.6	0.6	0.6
$d_{21}$	$X_{25}$	0.6897	0.6875	0.5429	0.5424	0.5871	0.5826
$d_{22}$	$X_{26}$	-	-	1.1	1.1	1.1	1.1
$d_{24}$	$XN_1$	0.7	0.7	0.7	0.7	0.7	0.7
$d_{27}$	$XN_4$	-	-	-	-	0.7	0.7
$d_{32}$	$XN_9$	0.6	0.6	-	-	0.6	0.6
$d_{33}$	$XN_{10}$	0.6	0.6	-	-	0.6	0.6
$d_{34}$	$XN_{11}$	0.9	0.9	-	-	0.9	0.9

## 7. Conclusion

A new efficient and effective variable screening method for RBDO is proposed in this paper. For the proposed methods, the output variance is used as a measure that can identify important design variables. Thus, a partial output variance based on the univariate DRM is proposed to calculate the output variance efficiently; and to identify the design variables that affect output variance more significantly than others. The univariate DRM and partial output variance only require multiple 1-D surrogate models, which is much more efficient than the full-dimensional surrogate models. Hence, the proposed method has great merit in efficiency as well as effectiveness. To reduce computational time and maintain a user-specified statistical error level, hypothesis testing is used in the variable screening process. Also, a required minimum number of samples for calculating the correct output

Table 14: Reliability Analysis Result Using Full-Dimensional Surrogate Model

Performance Measure	Baseline Design	Based on Experience	Variable Screening	Variable Screening + Cost Function	Full Dimension
Cost	269.47	249.10	259.83	244.17	225.68
$G_1$	48.25%	10.06%	9.96%	10.00%	10.05%
$G_2$	51.34%	10.02%	10.11%	10.04%	10.09%
$G_3$	54.14%	0.00%	0.00%	0.00%	0.00%
$G_4$	55.57%	0.10%	0.12%	0.09%	0.12%
$G_5$	58.94%	1.96%	1.93%	1.98%	1.91%
$G_6$	59.70%	10.08%	10.05%	10.05%	10.00%
$G_7$	59.86%	10.20%	10.04%	9.91%	10.06%
$G_8$	53.23%	10.02%	10.03%	9.97%	9.14%
$G_9$	51.15%	10.02%	9.96%	9.96%	9.96%
$G_{10}$	49.10%	0.00%	0.00%	0.00%	0.00%
$G_{11}$	52.46%	17.70%	11.23%	11.17%	10.05%

variance is proposed using the user-specified error level. In addition, the quadratic interpolation method is tailored to be applied to efficient partial output variance calculation.

A 44-D industrial example is used to verify the performance of the proposed method. Out of 44 design variables, 14 are selected by considering the output variances of 11 constraints. For comparison, another 14 design variables selected based on experience are used. In addition, 18 design variables are selected by adding four design variables, which affect the objective function significantly while not affecting the output variances much, to the 14 design variables previously selected with the proposed method. The selection based on experience shows a 7.6% reduced cost value, whereas the target probability of failure is violated by 77%. However, selection by the proposed method shows only a 12.3% violation of target value and a 3.6% reduced cost value. Moreover, the selection of 18 design variables shows 11.7% target violation as well as 9.4% reduced cost value. Therefore, the performance of the developed method is verified.

## 8. Acknowledgement

The research is jointly supported by ARO Project W911NF-09-1-0250 and the Automotive Research Center, which is sponsored by the U.S. Army TARDEC. This research was also partially supported by the World Class University Program through a National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (Grant Number R32-2008-000-10161-0 in 2009). These supports are greatly appreciated. The authors would also express appreciation to Lei Shi of Ford Motor Company for his modeling help and consultation.

## 9. Disclaimer

Reference herein to any specific commercial company, product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the Department of the Army (DoA). The opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or the DoA, and shall not be used for advertising or product endorsement purposes.

## 10. References

- [1] J.S. Arora, *Introduction to Optimum Design, 2nd ed.*, Elsevier/Academic Press, San Diego, 2004.
- [2] N.A.C. Cressie, *Statistics for Spatial Data*, John Wiley & Sons, New York, 1991.
- [3] R.R. Barton, Metamodeling: a State of the Art Review, *The 1994 Winter Simulation Conference*, J.D. Tew, S. Manivannan, D.A. Sadowski, A.F. Seila (Eds.), SCS, Orlando, USA, 1994.
- [4] R. Jin, W. Chen and T. Simpson, Comparative studies of metamodeling techniques under multiple modeling criteria, *Structural and Multidisciplinary Optimization*, 23 (1), 1-13, 2001.
- [5] T. Simpson, J. Poplinski and P. Koch, Metamodels for computer-based engineering design: survey and recommendations, *Engineering with Computers*, 17 (2), 129-150, 2001.
- [6] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan and P. Kevin Tucker, Surrogate-based analysis and optimization, *Progress in Aerospace Sciences*, 41 (1), 1-28, 2005.
- [7] G.G. Wang and S. Shan, Review of metamodeling techniques in support of engineering design optimization, *Journal of Mechanical Design*, 129 (4), 370-380, 2007.

- [8] A.I.J. Forrester, A. Sobester and A.J. Keane, *Engineering Design via Surrogate Modeling: A Practical Guide*, John Wiley & Sons, Chichester, 2008.
- [9] A.I.J. Forrester and A.J. Keane, Recent advances in surrogate-based optimization, *Progress in Aerospace Sciences*, 45 (1–3), 50-79, 2009.
- [10] L. Zhao, K.K. Choi and I. Lee, Metamodeling method using dynamic kriging for design optimization, *AIAA Journal*, 49 (9), 2034-2046, 2011.
- [11] W.J. Welch, R.J. Buck, J. Sacks, H.P. Wynn, T.J. Mitchell and M.D. Morris, Screening, predicting, and computer experiments, *Technometrics*, 34 (1), 15-25, 1992.
- [12] A.P. Duarte Silva, Efficient variable screening for multivariate analysis, *Journal of Multivariate Analysis*, 76 (1), 35-62, 2001.
- [13] H. Wang, Forward regression for ultra-high dimensional variable screening, *Journal of the American Statistical Association*, 104 (488), 1512-1524, 2009.
- [14] J. Eichler and U. Wille, Variable-screening model for the quasimolecular treatment of ion-atom collisions, *Physical Review A*, 11 (6), 1973-1982, 1975.
- [15] K.J. Craig, N. Stander, D.A. Dooge and S. Varadappa, Automotive crashworthiness design using response surface-based variable screening and optimization, *Engineering Computations*, 22 (1), 38-61, 2005.
- [16] Y. Wu and S. Mohanty, Variable screening and ranking using sampling-based sensitivity measures, *Reliability Engineering & System Safety*, 91 (6), 634-647, 2006.
- [17] K.K. Choi and N.H. Kim, *Structural Sensitivity Analysis and Optimization 1: Linear Systems*, Springer, New York, 2005.
- [18] K.K. Choi and N.H. Kim, *Structural Sensitivity Analysis and Optimization 2: Nonlinear Systems and Applications*, Springer, New York, 2005.
- [19] C.I. Reedijk, *Sensitivity Analysis of Model Output: Performance of various and Global Sensitivity Measures on Reliability Problems*, Master's Thesis, Delft University of Technology, 2000.
- [20] W. Chen, R. Jin and A. Sudjianto, Analytical variance-based global sensitivity analysis in simulation-based design under uncertainty, *Journal of Mechanical Design*, 127 (5), 875-886, 2005.
- [21] A. Haldar and S. Mahadevan, *Probability, Reliability and Statistical Methods in Engineering Design*, John Wiley & Sons, New York, 2000.
- [22] O. Ditlevsen and H.O. Madsen, *Structural Reliability Methods*, John Wiley & Sons, Chichester, 1996.
- [23] G.J.-W. Hou, *A Most Probable Point-Based Method for Reliability Analysis, Sensitivity Analysis, and Design Optimization*, NASA CR/ 2004-213002, 2004.
- [24] S. Rahman and D. Wei, Design sensitivity and reliability-based structural optimization by univariate decomposition, *Structural and Multidisciplinary Optimization*, 35 (3), 245-261, 2008.
- [25] I. Lee, K.K. Choi and D. Gorsich, Sensitivity analyses of FORM-based and DRM-based performance measure approach (PMA) for reliability-based design optimization (RBDO), *International Journal for Numerical Methods in Engineering*, 82 (1), 26-46, 2010.
- [26] I. Lee, K. Choi and L. Zhao, Sampling-based RBDO using the stochastic sensitivity analysis and dynamic kriging method, *Structural and Multidisciplinary Optimization*, 44 (3), 299-317, 2011.
- [27] I. Lee, K.K. Choi, Y. Noh, L. Zhao and D. Gorsich, Sampling-based stochastic sensitivity analysis using score functions for RBDO problems with correlated random variables, *Journal of Mechanical Design*, 133 (2), 021003, 2011.
- [28] M.D. McKay, J.D. Morrison and S.C. Upton, Evaluating prediction uncertainty in simulation models, *Computer Physics Communications*, 117 (1–2), 44-51, 1999.
- [29] I.M. Sobol, Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, *Mathematics and Computers in Simulation*, 55 (1–3), 271-280, 2001.
- [30] Y. Mack, T. Goel, W. Shyy and R. Haftka, Surrogate model-based optimization framework: A case study in aerospace design, *Studies in Computational Intelligence*, 51 323-342, 2007.
- [31] S. Bae, *Variable Screening Method Using Statistical Sensitivity Analysis in RBDO*, Master's Thesis, University of Iowa, 2012.
- [32] S. Rahman and H. Xu, A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics, *Probabilistic Engineering Mechanics*, 19 (4), 393-408, 2004.
- [33] B. Rosner, *Fundamentals of Biostatistics, 6th ed.*, Thomson Brooks/Cole, Belmont, 2006.
- [34] S.S. Rao, *Applied Numerical Methods for Engineers and Scientists*, Prentice Hall, Upper Saddle River, 2002.
- [35] K.K. Choi, N. Gaul, H. Song, H. Cho, I. Lee, L. Zhao, Y. Noh, D. Lamb, J. Sheng, D. Gorsich, R.-J. Yang, L. Shi, B. Epureanu and S.-K. Hong, Iowa Reliability-Based Design Optimization (I-RBDO) Code and Its Applications, *ARC Conference*, Ann Arbor, USA, 2012.