10$^{\text{th}}$ World Congress on Structural and Multidisciplinary Optimization
May 19 - 24, 2013, Orlando, Florida, USA

Genetic Algorithm Optimization for Reduced Order Problem Based on Kriging Modeling with Restricted Maximum Likelihood Criterion

Guang Dong

CAE group, Tesla Motors Inc., Fremont, CA 94538 Email:gdong@teslamotor.com

1. Abstract
Complex and computationally intensive modeling and simulation of real-world engineering systems can include a large number of design variables in the optimization of such systems. Consequently, it is desirable to conduct variable screening to identify significant or active variables so that a simpler, more efficient, and accurate optimization process can be achieved. This paper employs a variable screening method based on Kriging modeling with Restricted Maximum Likelihood criterion to reduce the design space, and the GA method is applied to optimize the re-defined problem with reduced order design space afterwards. The Kriging metamodeling method is more reliable for highly nonlinear systems, such as the complex engineering systems, than the traditional response surface method. Meanwhile, the Restricted Maximum Likelihood criterion makes the variable screening process more efficient. The Improved Distributed Hypercube Sampling method is applied at the first sampling stage in this study.
The strategy with the combination of variable screening method based on a Kriging modeling with Restricted Maximum Likelihood criterion and GA optimization method is evaluated using a 20 variables standard nonlinear benchmark function. This optimization strategy then is applied to a rubber material model optimization problem with 18 design variables. After reducing the design space to a less dimension using the variable screening method, the optimal rubber material model is obtained by using GA optimization. These two examples show that the optimization strategy proposed in this paper can solve the problem both efficiently and effectively.
2. Keywords: Genetic Algorithm, Kriging Metamodeling, Restricted Maximum Likelihood Criterion.

3. Introduction
Computational simulation and analysis are widely used in a great number of different engineering applications. Although computational power and speed grow continuously, complicated high-fidelity engineering models still have relatively high computational cost, especially when modeling parameters having uncertainties; thus design optimization for such computational intensive engineering system is limited. Therefore, numerous statistical approximation methods and approximation-based optimization are becoming widely used to minimize the computational expense [1]. A simple analytical model, which is used to approximate the computation-intensive engineering model, is denoted as a metamodel, and the process of generating a metamodel is called metamodelling [2]. It is important to note that deterministic computer experiments differ from physical experiments, which have random error. Three fundamental principles need to be considered for physical experiments: replication, randomization and blocking. These are generally not applicable to the computer experiments because the same input in a computer experiment gives rise to the same output [3].
In many cases, the complex engineering system includes a large number of design variables in the optimization process, and it is reasonable to expect some of these variables to be insignificant, or much less important than others. Thus, it is desirable to conduct a variable screening to identify the important variables so that a simpler metamodel and better interpretation can be achieved, such that further optimization can be conducted to determine the problem optimal solution more efficiently.
Kriging is a spatial correlation modeling method evolved in the field of geostatistics [4]. The first application of Kriging to computer experiments was introduced by Sacks et. al. [5]. Although the Response Surface (RS) methodology works well for small scale problems with simple curvature [6, 7, 8, 9], Kriging provides flexibility to approximate many complex response functions [10]. Kriging assumes some form of spatial correlation between points in the multi-dimensional input space, and uses this correlation to predict response values between the observed points. The resulting estimated surface can interpolate the observed responses [11], consequently, it is good for metamodeling. It is important to note that the estimated Kriging model correlation parameters are critical for the performance of the model.

In this paper, the following notations are employed: D is the experimental design space; d is the number of input variables, which corresponds to the dimension of D; $\mathbf{X}$ is the set of design points chosen in D ; n is the number of design points in $\mathbf{X}$, which corresponds to the number of observations of the response variable; $\mathbf{x}_i$ denotes the ith design point and $\mathbf{X} = [\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots,\ \mathbf{x}_n]$ with $\mathbf{x}_i \in \mathbb{R}^d,\ i = 1, 2, \ldots, n$. $y(\mathbf{x}_i)$ denotes the ith observation of the response variable and $\mathbf{Y}$ is the vector of response observations $\mathbf{Y} = [y(\mathbf{x}_1),\ y(\mathbf{x}_2),\ \ldots,\ y(\mathbf{x}_n)]^T$ the response variable $y(\mathbf{x}_i)$ could also be a q dimensional vector. Kriging model $\mathbf{Y}(\mathbf{x}) \in \mathbb{R}^q$ is the deterministic response for a d dimensional input $\mathbf{x} \in D \subseteq \mathbb{R}^d$ as a realization of a regression model F and a random function. The general Kriging approximation has the form:

$$Y_l(\mathbf{x}) = F(\beta_{:,l}, \mathbf{x}) + Z_l(\mathbf{x}) \quad l = 1, 2, \ldots, q \tag{1}$$

The regression model F is assumed as a linear combination of p chosen functions of $f_j(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$

$$F(\boldsymbol{\beta}, \mathbf{x}) = \beta_{1,l} f_1(\mathbf{x}) + \beta_{2,l} f_2(\mathbf{x}) + \cdots + \beta_{p,l} f_p(\mathbf{x}) \tag{2}$$

where $\boldsymbol{\beta} = [\beta_{1,l},\ \beta_{2,l},\ \ldots,\ \beta_{p,l}]^T$. The random process $Z_l(\mathbf{x})$ is assumed to have mean zero and covariance

$$Cov[Z_l(\mathbf{w}) Z_l(\mathbf{x})] = \sigma_l^2 R_l(\mathbf{w}, \mathbf{x}, \theta_{:,l}, \eta_{:,l}) \tag{3}$$

between $Z_l(\mathbf{w})$ and $Z_l(\mathbf{x})$ at two input vectors $\mathbf{w}$ and $\mathbf{x}$, where σ_l^2 is the process variance and $R_l(\mathbf{w}, \mathbf{x}, \theta_{:,l}, \eta_{:,l})$ is the correlation model with parameters $\theta_{:,l}$ and $\eta_{:,l}$, which depends on the relative location of two design points, $\mathbf{w}$ and $\mathbf{x}$.
A commonly used correlation model has the form,

$$R_l(\mathbf{w}, \mathbf{x}, \theta_{:,l}, \eta_{:,l}) = \Pi_{i=1}^{d} exp(-\theta_{i,l} \mid w_i - x_i \mid^{\eta_{i,l}}) \tag{4}$$

where $\theta_{:,l} \geq 0$ and $1 \leq \eta_{:,l} \leq 2$. The parameter $\eta_{:,l}$ can be interpreted as an indicator of increasing the smoothness of the response surface; thus larger $\eta_{:,l}$ indicates greater nonlinearity. It was pointed out that $\theta_{:,l}$ seems to be the more important than $\eta_{:,l}$ [12]. In this study the parameter $\eta_{:,l}$ was fixed at a value of 2, as Martin and Simpson [13] pointed out that $\eta_{:,l} = 2$ is the best suited to smooth functions and is the most commonly used value in engineering applications. Therefore, a Gauss exponential correlation model Eq. (5) is employed to reduce the complexity of variable screening algorithm in this study.

$$R_l(\mathbf{w}, \mathbf{x}, \theta_{:,l}, \eta_{:,l}) = \Pi_{i=1}^{d} exp(-\theta_{i,l} \mid w_i - x_i \mid^2) \tag{5}$$

Welch [12] proposed a variable screening method which combines the screening process with the selection of better model parameter sets. Welch performed the screening by building a Kriging metamodel based on a Latin Hypercube Sampling set [14, 15], and identified the important variables using the criterion of Maximum Likelihood Estimation (MLE) [16]. They proposed an algorithm that maximizes the MLE by considering the contribution of individual variables sequentially. In each loop of the algorithm, the most significant variable is selected from the initial set until only unimportant variables remain. A metamodel that only contains significant variables is constructed based on the results.
The Improved Distributed Hypercube Sampling (IHS) methods is employed in the variable selection step. It was developed by Beachkofski [17], based on the Distributed Hypercube Sampling (DHS) method [18], which adds another constraint by distributing sample points evenly as projected on to a two-dimensional face of the hypercube. Since LHS makes the set evenly distributed on the edge and DHS makes the set evenly distributed on the surface of the hypercube. IHS makes the set evenly distributed on the volume of the hypercube.(source code of J. Burkardt http://people.sc.fsu.edu/ jburkardt/)
Genetic Algorithms (GA), as a popular global optimization method, is an important member of the class of Evolutionary Algorithms (EA), which is inspired by the phenomenon of Darwins concept of survival of the fittest. The algorithm generates solutions to the optimization problems by using the techniques inspired by the natural evolution, such as selection, crossover, and mutation. GA has been widely used to solve a variety of nonlinear optimization problems [19]. In GA, the crossover and mutation operators are two basic operators used to generate the offsprings from parents to explore the design space. These processes ultimately result in the next generation population that is different from the initial generation. Generally the average fitness/cost will be improved by these two operators for the population, because only the best candidates from the previous generation are selected by the selection operator.

4. Restricted Maximum Likelihood Kriging Method for Variable Screening

4.1. Maximum Likelihood Estimation (MLE) Parameters

Using the Best Linear Unbiased Predictor (BLUP) approach, for given correlation parameters θ_i and η_i of the Kriging metamodel in Eq. 4, the predictor of y at an arbitrary point $\mathbf{x}$ can be shown as follows [12]:

$$\hat{y}(\mathbf{x}) = F(\hat{\boldsymbol{\beta}}, \mathbf{x}) + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}}) \tag{6}$$

where $\mathbf{F} = \begin{bmatrix} f_1(\mathbf{x}_1) & f_2(\mathbf{x}_1) & \cdots & f_p(\mathbf{x}_1) \\ f_1(\mathbf{x}_2) & f_2(\mathbf{x}_2) & \cdots & f_p(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\mathbf{x}_n) & f_2(\mathbf{x}_n) & \cdots & f_p(\mathbf{x}_n) \end{bmatrix}_{n \times p}$; $\mathbf{r}(\mathbf{x})$ is $n \times 1$ vector of correlations $\mathbf{R}(\mathbf{x}, \mathbf{x}_i)$ for $i = 1, 2, \ldots, n$ between covariance at arbitrary design point $\mathbf{X}$ and at each sampled points; and $\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimator of $\boldsymbol{\beta}$, given by,

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{R}^{-1}\mathbf{y} \tag{7}$$

The maximum likelihood estimator of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{n}(\mathbf{y} - \mathbf{F}\beta)^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\beta) \tag{8}$$

The correlation parameters θ_i and η_i , which determine the characteristics of the approximation between sample points, can be computed using the MLE approach. Martin and Simpson [13] concluded that the MLE approach is better than the CV method for selecting Kriging model parameters. The MLE approach is an unconstrained nonlinear optimization process in the space of parameters (θ_i, η_i) which tries to maximize the log-likelihood in Eq. (9).

$$Log\{L(\boldsymbol{\theta}, \boldsymbol{\eta}, \beta, \sigma^2)\} = -\frac{1}{2}[nLog\sigma^2 + Log(det(\mathbf{R})) + \frac{(\mathbf{y} - \mathbf{F}\boldsymbol{\beta})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\boldsymbol{\beta})}{\boldsymbol{\sigma}^2}] \tag{9}$$

4.2. Restricted Maximum Likelihood Estimation (REML) Parameters

The REML method is not based on a maximum likelihood fit of all the information, but instead employs a likelihood function calculated from transformed data, and it can produce unbiased estimates of variance and covariance parameters in contrast to the MLE [20]. In addition, the MLE estimator of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ are not involved in the optimization problem for the correlation parameter in the Kriging model, so it is not necessary to calculate the maximum likelihood estimator of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$, which, in some cases, can be difficult to obtain.

If we have n observations $\mathbf{Y}$ following a multivariate Gaussian distribution,

$$\mathbf{Y} \sim \mathbf{N}(\mathbf{F}\boldsymbol{\beta}; \mathbf{Z}) \tag{10}$$

then the restricted likelihood can be expressed in term of $\mathbf{Y}$, $\mathbf{F}$ and $\mathbf{Z}$ only as

$$2Log\{L(\boldsymbol{\theta}, \boldsymbol{\eta}, \beta, \sigma^2)\} = cons. - Log\{\mid \mathbf{Z} \mid\} - Log\{\mid \mathbf{F}^T\mathbf{Z}^{-1}\mathbf{F} \mid\} - \mathbf{Y}^T\{\mathbf{Z}^{-1} - \mathbf{Z}^{-1}\mathbf{F}(\mathbf{F}^T\mathbf{Z}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{Z}^{-1}\}\mathbf{Y} \tag{11}$$

The variance-covariance matrix for Kriging method is $\mathbf{Z} = \sigma^2\mathbf{R}$, which is a function of $\boldsymbol{\theta}$ and $\boldsymbol{\eta}$. If the regression model $F(\boldsymbol{\beta}, \mathbf{x})$ can be replaced by an unknown constant β , then $\mathbf{F} = \mathbf{1}$ and $\mathbf{F}\boldsymbol{\beta} = \mathbf{1}\beta$, where $\mathbf{1}$ is a column vector of 1's. For this case, the REML is an unconstrained nonlinear optimization problem in the space of parameters $(\theta_i, \eta_i, \sigma)$ which tries to maximize the log-likelihood in Eq. (11). Cholesky factorization for the covariance matrix $\mathbf{R}$ can handle the singularity issue in Eq. (11). Since we assume Gaussian correlation, the design space is reduced to the $n + 1$ dimensional space (θ_i, σ) from the $2n$ dimensional space (θ_i, η_i) without requiring the MLE estimator of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$.

4.3. Variable Screening Scheme

The basic idea of the algorithm is similar to Welch's method [12], but simpler. At first, the correlation parameters in the Kriging model are set as $\theta_1 = \theta_2 = \cdots = \theta_d = \theta$ for the correlation function in Eq. (5), then the numerical maximization of restricted likelihood only over two variables θ and σ . At each stage, let S denote the set of indexes of variables under the constraint of sharing common values of correlation parameter θ_i, while the remaining variables are free to have their own θ_i. Starting with $S = \{\theta_1, \theta_2, \ldots, \theta_d\}$, the algorithm iterates as follows. For each i in S , we relax the constraint $\theta_i = \theta$

and maximize the restricted likelihood in Eq. (11) subject to $\theta_j = \theta$ for all j in $S - \{i\}$. The variable x_i which results in the largest restricted likelihood is removed from S. The iterations terminate when none on the variables in S makes a large increment in the restricted likelihood relative to previous iteration. The variable screening algorithm using the Kriging method based on the criterion of REML can be encapsulted in the flowchart in Figure 1 [21].

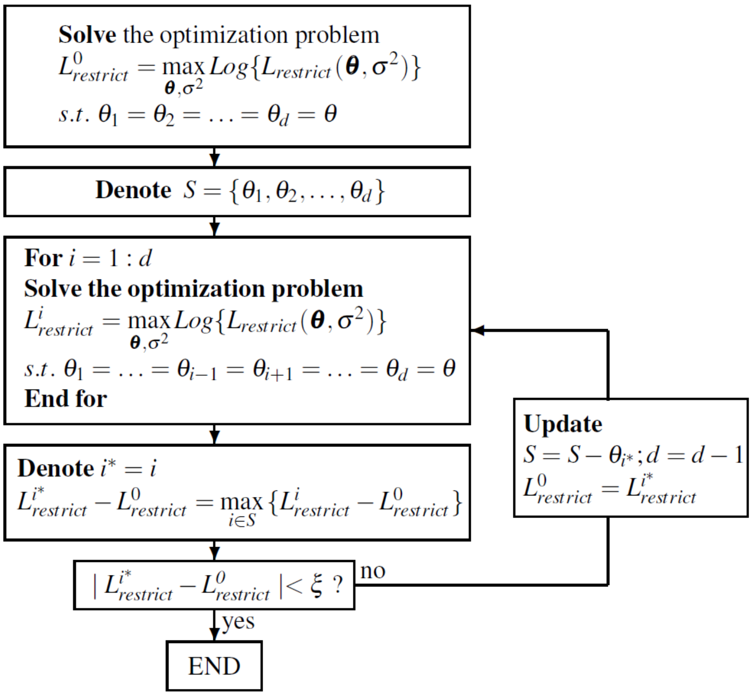

Figure 1: Flowchart of Kriging variable screening method

Therefore, the spirit of this algorithm is similar to the forward selection method of regression variables. The few most important variables can be screened out at first due to demanding their own θ_i and can produce larger values of θ_i in maximization of the restricted likelihood. The value of common θ_i for the variables remaining in S decreases and the restricted likelihood increases when more important variables are screened out and removed from the set S . If all the variables are either exceptionally active or exceptionally inactive, the value of common θ_i for set S would be zero after few iterations. However, if the variables in set S still have minor effects, the value of common θ_i may not trend to zero eventually, or may start to oscillate after all the important variables are screened out. In this case, we force the algorithm to stop if there is no substantial increment in the restricted likelihood relative to previous stage [22].

Based on the significant variables selected by the restricted maximum likelihood Kriging method, the problem dimension is reduced and the optimization algorithm is employed to solve the reduced problem. On one hand, the reduced order problem has less dimension searching space, the problem can be solved more efficiently. On the other hand, the reduced order problem keeps all the significant variables, the original problem can be solved effectively. After the design space is shrunk to a low dimensional space through the proposed nonlinear variable selection method, the GA method is then applied to find the optimal solution. Since GA is a stochastic optimization algorithm, reducing the problem dimension, or eliminating the insignificant variables can also reduce the results standard deviation and give better optimization performance.

5. Numerical Example

The test function proposed in this study is shown in Eq. (12) with the range of $\mathbf{x} \in [-0.5, 0.5]^{20}$,

$$\begin{aligned} f(\mathbf{x}) = {} & 5x_{12}\sin(1+x_1) + 5(x_4 - x_{20})^2 + x_5 + 40x_{19}^3 - 5x_{19} + 0.05x_2 + 0.08x_3 - 0.03x_6 + 0.03x_7 \\ & - 0.09x_9 - 0.01x_{10} - 0.07x_{11} + 0.25x_{13}^2 - 0.04x_{14} + 0.06x_{15} - 0.01x_{17} - 0.03x_{18} \end{aligned} \tag{12}$$

This function is obtained based on the benchmark test function proposed by Welch [12] with a modification of changing the first term from $\frac{5x_{12}}{1+x_1}$ to $5x_{12}\sin(1+x_1)$. The modification is to make the function is well bounded in the domain $[-0.5, 0.5]^{20}$. The function is strongly nonlinear and contains two interactions term, which is very challenging for both variable screening and optimization. The IHS sampling methods is employed in the variable selection step to generate 50 sampling points in the space. By keeping only the significant variables selected by the REML criterion Kriging variable screening method, the problem dimension can be reduced to 6 and the variable screening results are shown in Table 1.

Table 1: Variable screening results

Selected variables	θ_j for factors in S and $(\hat{\sigma}^2)^*$	Restricted Log-likelihood
19	(0.0436,1.0891)	26.8623
19,12	(0.0178,1.2686)	18.9545
19,12,4	(0.0123,1.2019)	12.2613
19,12,4,20	(0.0055,1.4510)	-42.2098
19,12,4,20,1	(0.0000,7.8997)	-49.2648
19,12,4,20,1,5	(0.0000, 26.6134)	-80.8847
19,12,4,20,1,5,15	(0.0000, 37.4303)	-81.7505

The reduced objective function $f'(\mathbf{x})$ in Eq. (13) only includes the significant variables x_1,x_4,x_5,x_{12},x_{19}, and x_{20}.

$$f'(\mathbf{x}) = 5x_{12}\sin(1+x_1) + 5(x_4 - x_{20})^2 + x_5 + 40x_{19}^3 - 5x_{19} \tag{13}$$

In the GA minimization, the selection operation is stochastic universal sampling method, the crossover operation is single point crossover, and the mutation rate is 0.1 [23]. There are 15 individuals for each generation and maximum generation number is 500. The GA optimization for both problems run 200 times, and the statistical results are shown in Table 2. It is known that the optimum of $f(\mathbf{x})$ is -5.7437, the optimum of $f'(\mathbf{x})$ is -5.4937 in the range of $x_i \in [-0.5, 0.5]$.

Table 2: Optimization results

Results	Mean	(Min,Max)	Standard deviation	Time (s)
original problem $f(\mathbf{x})$	-5.0202	(-5.7437,-3.2413)	0.9156	245
reduced order problem $f'(\mathbf{x})$	-5.2455	(-5.4937,-3.6169)	0.6285	100

From the optimization results in Table 2, it is concluded that the optimal solution for the reduced order problem with only significant variables is close to the original problem, but the computational time reduced significantly. Since the original problem has more variables, the stand deviation of the 200 GA optimization results is higher than the reduced order problem, and the mean value of the reduced order problem 200 optimization results is even better than the original problem results. Therefore, the GA optimization can obtain more benefits from reducing the problem dimension than the deterministic optimization methods.

6. Rubber Material Optimization Problem

A practical engineering example of using the proposed method is to improve the material computational model correlation with respect to the test data. The rubber material coupon tensile test normalized stress-strain results is shown in Figure 2 as "Test". Ogden material model is selected as a hyperelastic model to describe this rubber material in simulation, because it is widely used to describe the non-linear stress-strain behavior of complex materials such as rubbers, polymers, and biological tissue. This material model is proposed by Ogden in the year of 1972[24].

The simulation commercial code used in this study is LS-DYNA[26]. The design space of this optimization problem is defined as 18 parameters for defining the material card in the LS-DYNA code, which are listed in Table 3. More details about the material model equations can be found in the publication by Ogden [26]. After the variable screening based on Kriging method using the REML criterion, 6 significant variables are detected, μ_1, μ_2, α_1, α_2, G, and ξ. The Ogden material model is modified to only keep these variables for the optimization process. The design objective is the Least Square Residue (LSR) between the test curve and simulation curve [27]:

$$f(\mu_i, \alpha_i, G, \xi) = \sqrt{\sum_{i=1}^{N} \left(\frac{x_1 - y_i}{\lambda_i} \right)^2} \tag{14}$$

where x_i is the ith data point from the test curve, and y_i is the ith data point from the simulation curve, N is the number of total data points, λ_i is the scale factor for normalization or weighting factor for the ith data point.

Before the optimization, the normalized stress-strain simulation curve using the baseline Ogden material model is shown in Figure 2 as "before optimization". There is a clear discrepancy between the test stress-strain curve and the simulation curve by using the baseline Ogden material model.

Table 3: Ogden material card in LS-DYNA

Variables	Comments	Range
$\mu_i,\ i = 1, 2, \ldots, 8$	the shear modulus	[0, 3]
$\alpha_i,\ i = 1, 2, \ldots, 8$	the exponent parameters	[0, 4]
G	the shear relaxation modulus	[0, 80]
ξ	the decay constant	[0, 2000]

After the optimization, the normalized stress-strain simulation curve using the optimized Ogden material model is shown in Figure 2 as "after optimization". It shows that the rubber material model is improved significantly after the optimization and has much better correlation to the test data comparing to the baseline material model. Since the number of the design variables is reduced from 18 to 6, this optimization problems is solved more efficiently. Therefore, the rubber material model is improved by the proposed optimization strategy in this example.

7. Discussion

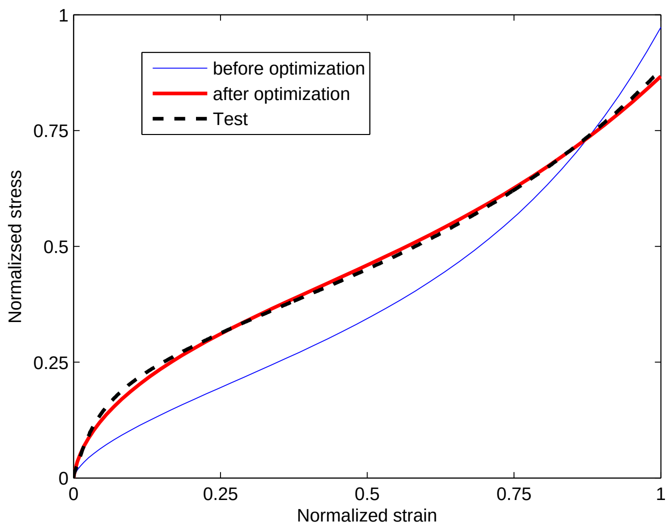

Figure 2: THE OGDEN MATERIAL MODEL

This study employs a variable screening method for complex and computational intensive engineering systems based on Kriging meta-models using the REML criterion. This approach is able to select important variables in a system without any linearity or additivity assumption. The nonparametric Kriging metamodel treats the deterministic computer experiments results as the realization of a stochastic process, and this model can automatically adapt to nonlinear and interaction effects in the data. Therefore, this variable screening method is more suitable for the highly nonlinear functions. Moreover, reducing the problem dimension size gives more benefits to the GA optimization since it is a stochastic algorithm, and there is randomness of the optimization results.

8. Conclusion

A variable screening method based on Kriging with REML criterion was employed to identify significant variables in the system, and then to establish an effective and simplified metamodel can be optimized more efficiently by using the GA optimization. Gaussian exponential correlation model and REML method are adopted to reduce complexity and improve the variable screening method based on Kriging metamodel. The reduced order problem with only the selected significant variables can be optimized by the GA method more efficiently. The effectiveness and efficiency of the developed optimization method can be shown using a 20 dimensional benchmark function and a rubber material model improvement example.

9. Acknowledgements

Thanks to Mr. John Cooper for the helps in preparing the paper and valuable suggestions to this study. The computer implementation of Prof. Burkardt's work (http://people.sc.fsu.edu/ jburkardt/) is used for IHS sampling method and thank him for sharing his source code.

10. References

[1] T. W. Simpson and D. K. Lin and W. Chen, Sampling Strategies for Computer Experiments: Design and Analysis, *International Journal of Reliability and Applications*, 2 (3), 209-240, 2001.

[2] G. Wang and S. Shan, Review of Metamodeling Techniques in Support of Engineering Design Optimization, *Journal of Mechanical Design*, 129 (4), 370-380, 2007.

[3] J. Wu and M. Hamada, *Experiments Planning, Analysis and Optimization 2nd Edition*, Wiley, New York, 2009.

[4] G. Matheron, Principles of Geostatistics, *Economic Geology*, 58 (8), 1246-1266, 1963.

[5] J. Sacks and W.J. Welch and T.J. Mitchell and H.P. Wynn, Design and Analysis of Computer Experiments, *Statistical Science*, 4 (4), 409-423, 1989.

[6] G. Box and N. Draper, *Empirical Model Building and Response Surface*, Wiley, New York, 1987.

[7] R. Myers and D. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Design Experiments*, Wiley, New York, 1995.

[8] K. J. Craig and N. Stander and D. A. Dooge and S. Varadappa, Automotive Crashworthiness Design Using Response Surface-Based Variable Screening and Optimization, *Engineering Computations: International Journal for Computer-Aided Engineering and Software*, 22 (1), 38-61, 2005.

[9] N. Stander and K. Craig, On the Robustness of the Successive Response Surface Method for Simulation Based Optimization, *Engineering Computations*, 19 (4), 431-450, 2003.

[10] R. Jin and W. Chen and T.W. Simpson, Comparative Studies of Metamodeling Techniques under Multiple Modeling Criteria, *8th AIAA/NASA/USAF/ISSMO Symposium Multidisciplinary Analysis and Optimization*, Long Beach, CA, 2000.

[11] V.C.P. Chen and K.L. Tsui and R.R. Barton and J.K. Allen, A Review on Design, Modeling and Applications of Computer Experiments, *IIE Transactions*, 38 (4), 273-291, 2010.

[12] W.J. Welch and R.J. Buck and J. Sacks and H.P. Wynn and T.T. Mitchell and M.D. Morris, Screening, Predicting, and Computer Experiments, *Technometrics*, 34 (1), 15-25, 1992.

[13] J. Martin and T. Simpson, On the Use of Kriging Models to Approximate Deterministic Computer Models, *ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Salt Lake City, UT, September 2004.

[14] M. D. McKay and W.J. Conover and R.J. Beckman, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code, *Technometrics*, 42 (1), 239-245, 1979.

[15] M. Stein, Large Sample Properties of Simulation Using Latin Hypercube Sampling, *Technometrics*, 29 (2), 143-150, 1987.

[16] D. Harville, Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems, *Journal of the American Statistical Association*, 72 (358), 320-338, 1977.

[17] B. Beachkofski, Improved Distributed Hypercube Sampling, *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Material Conference and Exhibit*, Denver, CO, April 2002.

[18] R. Manteufel, Distributed Hypercube Sampling Algorithm, *42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Material Conference and Exhibit*, Seattle, WA, April 2001.

[19] M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, 1996.

[20] M. Kenward and J.H. Roger, Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood, *Biometrics*, 53 (1), 983-997, 1997.

[21] G. Dong and Z. -D. Ma and G. M. Hulbert and N. Kikuchi, Variable Screening Using Restricted Maximum Likelihood Kriging Method with Application to Gunner Joint Stiffness Variables, *the ASME 2011 International Design Engineering Technical Conferences/Computers and Information in Engineering Conference*, Washington, D.C., August, 2011.

[22] G. Dong, Topology Optimization for Multi-Functional Components in Multibody Dynamics Systems, *Ph.D. Dissertation*, University of Michigan, Ann Arbor, January, 2012.

[23] A.J. Chipperfield and P. J. Fleming, The MATLAB Genetic Algorithm Toolbox, *Applied Control Techniques Using MATLAB, IEE Colloquium on*, London, UK, January, 1995.

[24] R. W. Ogden, Large Deformation Isotropic Elasticity - On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 326 (1567), 565-584, 1972.

[26] LSTC, *LS-DYNA Keyword User's Manual - Version 971*, Livermore Software Technology Corporation, 2007.

[26] R. W. Ogden, *Nonlinear Elastic Deformation*, Ellis Horwood Ltd., Chichester, Great Britain, 1984.

[27] S. I. Yi and P. K. Mohan and C. D. Kan and G. J. Park, Finite Element Modelling of a Hybrid III Dummy and Material Identification for Validation, *Proceeding of the Institute of Mechanical Engineers, Part D: Journal of Automotive Engineering*, 225 (54), 54-73, 2011.