Springer

Journal of Mechanical Science and Technology 23 (2009) 1175~1181

Journal of Mechanical Science and Technology

www.springerlink.com/content/1738-494x
DOI 10.1007/s12206-008-1201-3

Comparison study on the accuracy of metamodeling technique for non-convex functions †

Byeong-Soo Kim[1], Yong-Bin Lee[1] and Dong-Hoon Choi[2,*]

[1]*Department of Mechanical Engineering, Hanyang University Graduate school, Seoul, 133-791, Korea*
[2]*The Center of Innovative Design Optimization Technology (iDOT), Hanyang University, Seoul, 133-791, Korea*

(Manuscript Received June 4, 2008; Revised October 15, 2008; Accepted December 4, 2008)

Abstract

In order to increase the efficiency of design optimization, many efforts have been made on studying the metamodel techniques for effectively representing expensive and complex models. In this study, a comparison is conducted on the accuracy of several widely used meta-model techniques - moving least squares (MLS), Kriging, support vector regression (SVR) and radial basis functions (RBF) - which are able to approximate non-convex functions well.. RMSE (root mean squared error) value is identified as a measure of the accuracy for this comparison. Each metamodel technique is used to approximate the six well-known mathematical functions and a resign of experiment (DOE) is generated by using the Latin hypercube design (LHD), which is also performed for each resulting metamodel. The results show that Kriging and MLS can create a more accurate metamodel than SVR and RBF with the mathematical functions tested.

Keywords: Metamodel; Moving least squares; Kriging; Radial basis function; Support vector regression

1. Introduction

In order to improve the efficiency of design optimization, various researches on metamodels are in progress. One of the representative metamodeling techniques is response surface methodology (RSM) [1] which employs the method of least squares. However, the widely used quadratic RSM has a shortcoming in that it is not able to represent non-convex functions. In this paper, the accuracy of four metamodeling techniques is compared, including moving least squares (MLS), Kriging, radial basis functions (RBF) and, support vector regression (SVR), which are known to be suitable for representing non-convex functions as well as convex functions.

Jin et al. [2] claimed that RBF is a better metamodeling technique compared to the other three techniques: Kriging, RSM, and multivariate adaptive regression splines (MARS). However, because the problems are lopsided in two extreme cases, i.e., cases with 2 or 3 variables and cases with over 10 variables, it is not possible to predict the trend for cases with 4-10 variables. Krishnamurthy [3] also compared the accuracy of Kriging, MLS, RSM and RB; however, the conclusion drawn from his work is highly biased because too few application problems were studied and the sample problems dealt with only one variable.

In the study presented in this paper, the accuracy of four metamodeling techniques is compared with the use of the metamodels with various numbers of sampling points. Furthermore, the accuracy of the metamodeling techniques is measured in terms of the number of variables by using the Rosenbrock function in which an arbitrary number of variables can be set. As a proof of implementation, representative mathematical sample problems are applied, and the Latin hypercube design method is used to create the

† This paper was recommended for publication in revised form by Associate Editor Dae-Eun Kim

*Corresponding author. Tel.: +822 22900478, Fax.: +822 22914070
E-mail address: dhchoi@hanyang.ac.kr

design of experiments for every metamodeling technique. The root mean square error (RMSE) and ranking are identified as the measure of accuracy.

In order to measure the accuracy of the metamodeling techniques in terms of the number of sampling points and the number of variables, 225 metamodels are constructed for comparison purpose. Especially for MLS, two different weight functions are utilized in order to analyze the effect of the weight function on the accuracy.

2. Metamodeling methodologies

2.1 Moving Least Squares (MLS)

Typically, the quadratic RSM predicts convex functions relatively well; however, large errors for highly non-linear data is also seen. This issue can be resolved by introducing the method of least squares, usually used in a mesh free method, to the RSM

Multiplying weight functions to the coefficient matrices, MLS zeros out the data except for those around a particular region. The response surface constructed in this way can provide very accurate local predictions. The mathematical statement of the RSM with MLS is given by Eq. (1) [4-5].

$$\tilde{y} = \mathbf{P}^{\mathrm{T}} \cdot \mathbf{a}$$

$$\mathbf{P} = \begin{bmatrix} 1 & x & \cdots & x^m \end{bmatrix}^{\mathrm{T}}$$

$$\mathbf{a} = \left(\mathbf{B}^{\mathrm{T}}\mathbf{W}\mathbf{B}\right)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{W}\mathbf{f}$$

$$where,\ \mathbf{B} = \begin{bmatrix} 1 & x^0 & \cdots & x_m^0 \\ 1 & x^1 & \cdots & x_m^1 \\ & & \vdots & \\ 1 & x^n & \cdots & x_m^n \end{bmatrix},\ \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix},\ \mathbf{f} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}$$

$$\mathbf{W} = diag\left[w_0(\mathbf{x}), w_1(\mathbf{x}), \cdots, w_n(\mathbf{x}) \right] \tag{1}$$

where, $m+1$ is the number of basis functions and $n+1$ is the number of sampling points.

2.2 Kriging

The Kriging model [6] addresses the computer experiment information at the sampling points with the summation of the global model and the local deviation, as given by Eq. (2):

$$y(x)=\mathbf{f}(\mathbf{x})\cdot\boldsymbol{\beta}+Z(\mathbf{x}) \tag{2}$$

where, $Z(\mathbf{x})$ is a normal distribution with mean of 0 and variance of σ^2. The deviation at each sampling point has the correlation defined in Eq. (3):

$$Cov[Z(\mathbf{x}^i), Z(\mathbf{x}^j)]=\sigma^2\mathbf{R}[R(\mathbf{x}^i,\mathbf{x}^j)],\ i,j=1,\ldots,n \tag{3}$$

$$R(\mathbf{x}^i,\mathbf{x}^j) = \exp\left[-\sum_{k=1}^{N} \theta_k \left| x_k^{\,i} - x_k^{\,j} \right|^2 \right]$$

where, n and N are the number of sampling points and the number of design variables, respectively. $x_k^{\,i}$ is the kth design variable at the ith sampling point. $R(\mathbf{x}^i,\mathbf{x}^j)$ is a function that represents the correlation between two sampling points. $\mathbf{x}^i$ and $\mathbf{x}^j$, and is generally depicted by Gaussian correlation function. In order to construct a Kriging model, the correlation coefficients θ_k are determined by using the most likely maximum estimate method. However, since Kriging requires a global optimization process to produce the metamodel, this method leads expensive computation for the problems with many design variables and sampling points. Another shortcoming is that it may fail to generate a proper metamodel if the global optimization is not successful.

2.3 Radial Basis Function (RBF)

RBF [7] is often used to perform the interpolation of scattered multivariate data. The metamodel appears in a linear combination of Euclidean distances, as shown in Eq. (4).

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{n} w_i \phi_i(\mathbf{x}, x_i) \tag{4}$$

where, n is the number of sampling points, w_i is the weight determined by the least-squares method and $\phi_i(\mathbf{x}, x_i)$ is the ith basis function determined at the sampling point $\mathbf{x}_i$. Various symmetric radial functions used as the basis function are listed in Table 1.

The RBF method is known to be as good an interpolator as Kriging. However, such an interpolation method has shortcomings in that the appearance of a metamodel varies significantly with the type of basis function and its internal parameters. In this paper a Gaussian basis function with the fixed parameter r^2=1.5 is used.

Table 1. Radial function for RBF model.

Name	Radial function
Gaussian	$h(\mathbf{x}) = \exp\left(-\frac{(\mathbf{x}-\mathbf{c})^T(\mathbf{x}-\mathbf{c})}{r^2}\right)$
multiquadric	$h(\mathbf{x}) = \sqrt{1+\frac{(\mathbf{x}-\mathbf{c})^T(\mathbf{x}-\mathbf{c})}{r^2}}$
inverse multiquadric	$h(\mathbf{x}) = \frac{1}{\sqrt{1+\frac{(\mathbf{x}-\mathbf{c})^T(\mathbf{x}-\mathbf{c})}{r^2}}}$
Cauchy	$h(\mathbf{x}) = \frac{1}{1+\frac{(\mathbf{x}-\mathbf{c})^T(\mathbf{x}-\mathbf{c})}{r^2}}$

2.4 Support Vector Regression (SVR)

In SVR [8] the linear regression model appears in the form defined by Eq. (5):

$$f(\mathbf{x}) = \langle \mathbf{w} \cdot \mathbf{x} \rangle + b \tag{5}$$

where, $\langle \mathbf{w} \cdot \mathbf{x} \rangle$ is a dot product of **w** and **x**. The distance between the models and the support vector is denoted as $1/\|\mathbf{w}\|$. The final model is determined by solving an optimization problem given by Eq. (6), minimizing $\|\mathbf{w}\|^2$ within a specified range of ε.

$$\begin{aligned} &\textit{Minimize} \ \ \frac{1}{2}|w|^2 \\ &\textit{subject to} \begin{cases} y_i - \langle w \cdot x_i \rangle - b \le \varepsilon \\ \langle w \cdot x_i \rangle + b - y_i \le \varepsilon \end{cases} \end{aligned} \tag{6}$$

With the use of LaGrange' theory, the problem given by Eq. (6) can be transformed into a dual problem as shown in Eq. (7).

$$\begin{aligned} \textit{Maximize} \ & -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle \mathbf{x_i} \cdot \mathbf{x_j} \rangle \\ & -\varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \\ \textit{subject to} \ & \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ & \alpha_i \ , \ \alpha_i^* \in [0, C] \end{aligned} \tag{7}$$

The final regression model can be given by Eq. (8).

$$f(\mathbf{x}) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)\langle \mathbf{x_i} \cdot \mathbf{x} \rangle + b. \tag{8}$$

The Gaussian kernel function, which is widely used in SVR, is defined by Eq. (9).

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right) \tag{9}$$

Using the Gaussian kernel function, the input parameters of the SVR model become ε, C and σ, which is related to Gaussian function. The metamodel is constructed with fixed input parameter: ε=0.0, C=10^5, σ=0.5.

3. Comparative study method

3.1 Design of Experiments (DOE)

Developed by McKay et al. [9], the Latin hypercube design (LHD) [6] distributes sampling points uniformly within a range of each single design variable. It rearranges the sampling points into an array with n_{level} rows and N columns, and then assigns each column as sampling points. Where, n_{level} is the number of the level of a design variable and N is the number of design variables. Each column is a permutation vector where the levels $1,2,\cdots,n_{level}$ of a design variable are arbitrarily allocated to each row. Therefore, in the case of one-dimensional projections, n_{level} levels are obtained, and there is no overlap of sampling points for each design variable, which makes it suitable for deterministic computer experiments. In addition, LHD is easily implemented because the levels of the design variables can be allocated arbitrarily, thereby making it convenient to control the number of sampling points.

3.2 Comparative study parameters

In this work, in order to measure the accuracy of each metamodel created by different techniques in terms of the number of variables N, metamodels are constructed for cases with 2, 4, 6 and 8 variables.

Additionally, in order to investigate the effect of the number of sampling points on the accuracy, sample sets with $3N$, $5N$ and $7N$ sampling points are used.

For MLS, in order to measure the accuracy of the metamodels with respect to weight functions, two widely used weight functions are considered. They are given by Eqs. (10)-(11), respectively [5].

Table 2. Number of test points.

Number of variables	2	4	6	8
Number of test points	21^2 =441	15^4 =50,625	9^6 =531,441	6^8 =1,679,616

$$w_j(\mathbf{x}) = \frac{\exp\left(-0.1\left\|\mathbf{x}-\mathbf{x}_j\right\|^2\right)}{\left\|\mathbf{x}-\mathbf{x}_j\right\|^4 + 10^{-7}} \tag{10}$$

$$w_j(\mathbf{x}) = \frac{1}{\left\|\mathbf{x}-\mathbf{x}_j\right\|^2 + 10^{-7}} \tag{11}$$

3.3 Measure of performance

The RMSE value is used as the measure of accuracy for evaluation of metamodels and is given by Eq.(12).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n_{test}} (\tilde{y}_i - y_i)}{n_{test}}} \tag{12}$$

where, n_{test} is the number of sampling points used in the test. In this study n_{test} is developed through the grid points as shown in Table 2.

It is not prudent to compare the accuracy by simply adding the RMSEs of 225 metamodels constructed based on the number of variables and the number of sampling points. Therefore, a quantitative measure is needed to combine and compare that data in a new way. The study presented in this paper suggests using a ranking method to rank the metamodels in an ascending order. In other words, the higher the ranking is the greater the RMSE value the model has. In addition, if the difference between the RMSEs of two metamodels is greater than 40%, the metamodel with larger RMSE gets +1 promotion in ranking, and if the difference is greater than 200%, it gets another +1 in ranking.

3.4 Sample problem

The sample problem consists of 6 mathematical problems, as shown in Table 3. The sum of square function [10], Rosenbrock function [10] and Giunta's function [11] can be extended to have more than 2 variables. In this work, the functions are extended with the capability of dealing with 2, 4, 6 and 8 variables to construct the metamodels. The PDF shape function, Goldstein and Price function [10] and sinusoidal function [2] are used to construct the metamodels with only 2 variables. Therefore, there are 6 sample problems used for 2 variables, and 3 sample problems for more than 4 variables.

Table 3. Mathematical sample problems.

Function name	equation	bound
Sum of squares function	$f(\mathbf{x}) = \sum_{i=1}^{n} i x_i^2$	$-5 \le x_i \le 10$
Normal PDF shape function	$f(\mathbf{x}) = \frac{1}{1 + x_1^4 + 5x_2^4 + 2x_2^2}$	$-3 \le x_i \le 3$
Rosenbrock function	$f(\mathbf{x}) = \sum_{i=1}^{n}\left(\frac{3}{10}\sin\left(\frac{16}{15}x_i - 1\right) + \sin^2\left(\frac{16}{15}x_i - 1\right)\right)$	$-5 \le x_i \le 10$
Giunta's function	$f(\mathbf{x}) = \sum_{i=1}^{n+1}\left[100\left(x_i^2 - x_{i+1}\right)^2 + (x_i - 1)^2\right]$	$-1 \le x_i \le 1$
Goldstein and Price function	$f(\mathbf{x}) = [1 + (x_1 + x_2 + 1)^2$ $\times(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$ $\times[30 + (2x_1 - 3x_2)^2$ $\times(18 - 32x_1 + 12x_1^2 - 48x_2 - 36x_1x_2 + 27x_2^2)]$	$-2 \le x_i \le 2$
Sinusoidal nonlinear function	$f(\mathbf{x}) = \sin(x_1 + x_2) + (x_1 - x_2)^2$ $-1.5x_1 + 2.5x_2 + 1$	$-10 \le x_1 \le 10$ $-20 \le x_2 \le 20$

4. Results of accuracy comparison

According to the comparison study method suggested in section 3, the accuracy of metamodels created by different metamodeling techniques is compared and described in terms of the measure of performance discussed in section 3.3 for the sample problems presented in section 3.4.

Fig. 2 shows all the metamodels ranked and sorted according to the number of variables. On the other hand, Fig. 4 shows the metamodels sorted according to the number of sampling points.

A lower ranking means that the metamodel is relatively more accurate than the others, i.e., the RMSE is small. Overall, it can be seen that MLS and Kriging have a relatively higher accuracy than RBF and SVR.

Fig. 3 depicts the RMSE of the metamodels that approximate the Rosenbrock function depending on the number of variables. Similar results with the RMSE are found by the other functions. It is noted that this is a very similar trend to the ranks shown in Fig. 2.

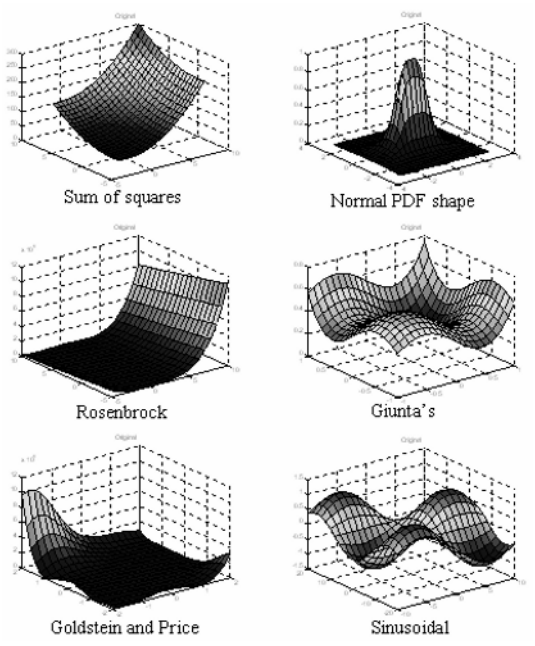

Fig. 1. Mathematical sample problems for 2 variables.

Fig. 5 illustrates the RMSE of each metamodel that approximates the Rosenbrock function depending on the number of sampling points. As the number of sampling points increases, the accuracy of the metamodel also increases, leading to a gradual decrease in the RMSE value of the Rosenbrock function. A noteworthy thing is that the slope of the RMSE produced using SVR is less than the others. This agrees with the fact that the more sampling points there are, the higher the ranking is, as shown in Fig. 4.

Furthermore, the MLS #2 with the weight function defined by Eq. (11) has a higher accuracy than the MLS #1 with the weight function given by Eq. (10), while there is no significant difference between the RMSE values. Therefore, by adopting the widely used weight functions, designers may minimize the number of input parameters controlled.

Kriging constructs a much more accurate metamodel for 2-variable cases, as compared to the other methods. However, for the cases with 4 or more variables, results show that it had similar or worse accuracy than MLS. This is due to the fact that increasing the computational loads results in the increase in the number of variables. In addition, its shortcomings, such as calculating inverse matrices and global optimization, cause more computational loads than the other techniques.

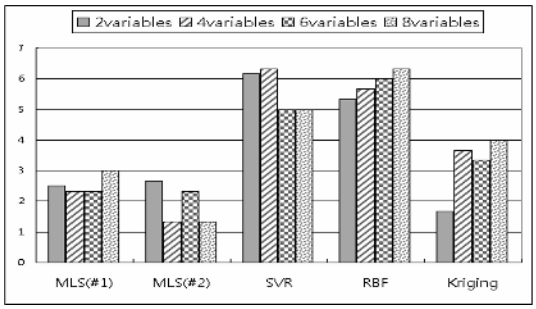

Fig. 2. Ranking (The Number of Variables).

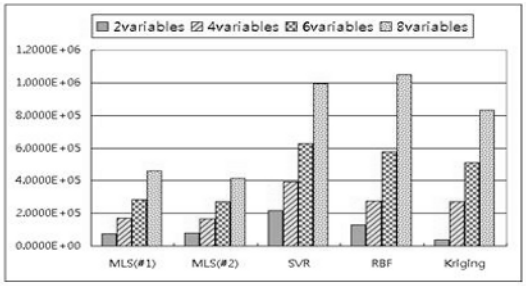

Fig. 3. RMSE of Rosenbrock function (The Number of Variables).

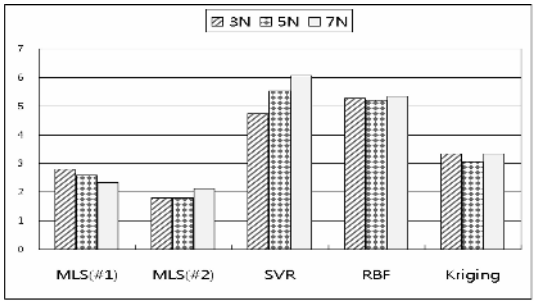

Fig. 4. Ranking (The Number of DOE).

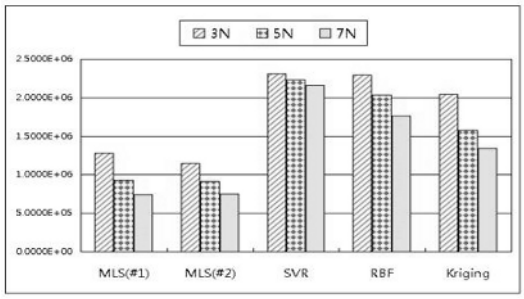

Fig. 5. RMSE of Rosenbrock function (The Number of DOE).

Overall, SVR constructed the most inaccurate metamodels, while it showed fast convergence of RMSE according to the number of sampling points. In other words, it showed a trend that increasing the number of sampling points did not improve the accuracy.

RBF exhibits an abrupt improvement in accuracy over the other metamodeling methods for some cases with an increased number of sampling points. However, it has also shown the opposite trend of accuracy due to the increase in the number of sampling points. This method appeared the most unstable.

In brief, SVR and RBF provide less robust results than Kriging and MLS in terms of accuracy. Kriging and MLS exhibit similar RMSE value trends without having significant differences for the same problem, while SVR and RBF have severe deviations. Namely, for the same problem, sometimes they constructed the most accurate metamodels for cases with fewer sampling points, while sometimes the RMSE value increases for cases with more sampling points. On the other hand, having a higher RMSE increases the number of sampling points abruptly dropped by the RMSE and thus changing the ranking.

5. Concluding remarks

In conclusion, Kriging and MLS constructed more accurate metamodels than RBF and SVR. Kriging constructed the most accurate metamodels for the cases with 2 variables, while MLS can create the most accurate metamodel for the cases with more variables.

Especially, the metamodels that were developed by using MLS show the highest accuracy. Higher accuracy can also be obtained with more sampling points, as compared to the other metamodeling methods.

This study shows that a metamodel that represents an actual model could not be constructed for the small number of sampling points, e.g. 3N, 5N and 7N, because of the strong non-linearity. Therefore, follow-up researches with sufficient sampling points are required for constructing metamodels that closely resemble the actual model.

Acknowledgment

This research was performed in support of The Center of innovative Design Optimization Technology (iDOT), Seoul R&BD program, and the International Collaborative Technology Development Program from the Korea Ministry of Commerce, Industry and Energy. The authors would like to express gratitude for their support.

References

[1] R. H. Myers, and D. C. Montgomery, Response Surface Methodology –Process and Product Optimization Using Designed Experiments, John Wiley & Sons, New York, USA, (1995).

[2] R. Jin, W. Chen and T. W. Simpson, Comparative studies of metamodeling techniques under multiple modeling criteria, *Journal of the International Society for Structural and Multidisciplinary optimization*, 23 (1) (2001) 1-13.

[3] T. Krishnamurthy, Comparison of Response Surface Construction Methods for Derivative Estimation Using Moving Least Squares, Kriging and Radial Basis Functions, *Proc. of 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference,* AIAA-2005-1821, (2005) 1-28.

[4] P. Lanscater and K. Salkauskas, *Curve and Surface Fitting: An Introduction*, Academic Press, (1986).

[5] G. G. Maisurdze and D. L. Thompson, Interpolating Moving Least Squares Methods for Fitting Potential Energy Surfaces: Illustrative Approaches and Applications, *Journal of physical chemistry. A, Molecules, spectroscopy, kinetics, environment & general theory*, 107 (37) (2003) 7118-7124.

[6] Y. Shin, Y. Lee and J. Ryu, Sequential Approximate Optimization Using Kriging Metamodels, *Transactions of the KSME*, 29 (9) (2005) 1199-1208.

[7] M. Powell, Radial Basis Functions for Multivariable Interpolation: A review, Algorithms for Approximation (Mason, J. C. and Cox, M. G., eds.), Oxford University Press, (1987) 143-167.

[8] S. M. Clarke, H. Griebsch and T. W. Simpson, Analysis of Support Vector Regression For Approximation of Complex Engineering Analysis, *ASME Journal*, 127 (6) (2005) 1077-1087.

[9] M. D. McKay, R. J. Beckman and W. J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics* 21 (2) (1979) 239-245.

[10] http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO_files/Page364.htm.

[11] A. A. Giunta and L. T. Watson, A Comparison of

Approximation Modeling Techniques: Polynomial Versus Interpolating Models, *Proc. of 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis, AIAA-98-4758*, (1998) 392-404.

Byeong-Soo Kim received B.S and M.S. degree in Mechanical Engineering from Hanyang University in 2006 and 2008, respectively. Mr. Kim is currently a Research Engineer at LG Electronics.

Yong-Bin Lee received a B.S. and M.S. degree in Mechanical Engineering from Hanyang University in 2002 and 2004, respectively. He is currently a Ph.D. student in Hanyang University. Mr. Lee's research interests are in the area of optimization, approximation, and design of experiments.

Dong-Hoon Choi received a B.S. degree in Mechanical Engineering from Seoul National University in 1975. He then went on to receive his M.S. from KAIST in 1977 and Ph. D. degree from University of Wisconsin-Madison in 1986, respectively. Dr. Choi is currently a Professor at the School of Mechanical Engineering at Hanyang University in Seoul, Korea. He is currently the director of iDOT(the center of innovative design optimization technology). Prof. Choi's research interests are in the area of optimization techniques: developing MDO methodology, developing optimization techniques to ensure a reliability of optimum solution, and developing approximation optimization technique, etc.