10th World Congress on Structural and Multidisciplinary Optimization
May 19 -24, 2013, Orlando, Florida, USA

Constrained Global Design Optimization Using a Multi-fidelity Model

Seok-Ho Son[1], Do-Hyun Park[2], Kyeong-Jun Cha[3], Dong-Hoon Choi[4]

[1] Hanyang University, Seongdong-gu, Seoul, Republic of Korea, seokho.son@gmail.com
[2] Hanyang University, Seoul, Republic of Korea, hqwer21@hanmail.net
[3] Hanyang University, Seoul, Republic of Korea, kjcha@hanyang.ac.kr
[4] Hanyang University, Seoul, Republic of Korea, dhchoi@hanyang.ac.kr

1. Abstract
In many engineering optimization problems, the number of function evaluations is severely limited due to time or cost for obtaining the optimal solution when the time for one function evaluation is long. In this study, we present a constrained global optimization method using a multi-fidelity model. A multi-fidelity model employed in this study is a surrogate that combines two models of different fidelities. The multi-fidelity model has been used to obtain unconstrained optimal solution with less number of function evaluations than using just a high fidelity model.

The multi-fidelity model we used in this study is that based on Bayesian prediction. For sequential sampling taking the constrained global optimization into account, we formulate an optimization problem to find an infill sampling point that maximizes the generalized expected improvement (GEI), a generalized form of the expected improvement. The GEI balances a local exploitation and a global exploration with a cooling scheme. To solve the formulated optimization problem, we employ the pattern search method to find the constrained global optimal solution. The sequential sampling terminates when stopping criteria are met.

In order to demonstrate the effectiveness and usefulness of the proposed method, several mathematical problems are solved and its performance is compared to that of constrained global optimization methods using a high fidelity model only.

2. Keywords: constrained global optimization, multi-fidelity model, generalized expected improvement

3. Introduction
In many engineering problems, the amount of the high fidelity data is not enough to represent the entire design space because the evaluation of high fidelity data is very expensive. To avoid this problem, the multi-fidelity model is used for optimization process without directly using high-fidelity simulations. Among many proposed multi-fidelity modeling, a good performance model is known to be a high fidelity model constructed by a low fidelity Gaussian process multiplied by a scale parameter plus a correction Gaussian process. We investigate the performance of this multi-fidelity model and present it in Section 4. For sequential sampling taking the global optimization into account, a constrained global optimization problem is formulated to find an infill sampling point that maximizes the generalized expected improvement while satisfying the constraints. In this work, we use two steps of optimization process in order to guarantee the exact solution. The formulated optimization process is stopped when the stopping criteria is met. In order to demonstrate the effectiveness and usefulness of the proposed method, we investigate its performance using several mathematical problems and represent it in Section 5.

4. Performances of Existing Multi-fidelity Models
In this section, we study the characteristics of existing multi-fidelity models: how to build a multi-fidelity model, how to estimate parameters of the multi-fidelity model, and how to predict at a new point of interest. Then, we apply the four models to mathematical test problems and compare the accuracy of four models with root mean square errors.

4.1. Building methods of the multi-fidelity models
Kennedy and O'Hagan [1] and Forrester et al. [2] built multi-fidelity models as follows:

$$y_h(\mathbf{x}) = \rho \cdot y_l(\mathbf{x}) + \delta(\mathbf{x}), \tag{1}$$

where ρ (scale parameter) was defined as a constant. Using auto regressive model, the multi-fidelity model is created as a Gaussian process $y_l(\mathbf{x})$ multiplied by a scale parameter ρ plus a Gaussian process $\delta(\mathbf{x})$.

Qian et al. [3] employed $\rho(\mathbf{x})$ represented as a linear regression function instead of a constant as follows:

$$\rho(\mathbf{x}) = \rho_0 + \sum_{j=1}^{nsat} \rho_j x_{ij}, \ i = 1, ... nexp, \tag{2}$$

where nsat is the number of saturated points and nexp is number of experimental (or sampling) points for the high fidelity model. $y_l(\mathbf{x})$ and $\delta(\mathbf{x})$ was assumed to be a stationary Gaussian process with means (β_l, β_δ) , variances $(\sigma_l^2, \sigma_\delta^2)$, and correlation parameters $(\boldsymbol{\theta}_l, \boldsymbol{\theta}_\delta)$.

Qian et al. [4] also built a multi-fidelity model that employed a stationary Gaussian process for the scale factor as:

$$y_h(\mathbf{x}) = \rho(\mathbf{x}) \cdot y_l(\mathbf{x}) + \delta(\mathbf{x}), \tag{3}$$

where $\rho(\mathbf{x})$ is a stationary Gaussian process with mean β_ρ , variance σ_ρ^2 , and correlation parameters $\boldsymbol{\theta}_\rho$.

4.2 Estimation Methods of Parameters

Kennedy and O'Hagan [1] estimated all parameters with the Bayes' inference and maximum likelihood estimator. In their study, they considered data set from low fidelity simulation (LF) and high fidelity simulation (HF) together. Therefore, mean parameters $\boldsymbol{\theta}_\rho$ with a non-informative prior were estimated by their posterior mean. In contrast, the parameters (σ_l^2, $\boldsymbol{\theta}_l$) of LF model were estimated by maximizing the log-likelihood function. In addition, the parameters $\boldsymbol{\theta}_\rho$ of the HF model were estimated by maximizing the log-likelihood function of high fidelity model data.

Forrester et al. [2] also estimated all parameters by maximizing the concentrated log-likelihood function.

In Qian et al. [3], the parameters $\boldsymbol{\beta}_l, \sigma_l^2, \boldsymbol{\theta}_l$ of the LF model were calculated by maximizing the likelihood function of $\mathbf{y}_l$, and the estimates $\boldsymbol{\alpha} = (\boldsymbol{\beta}_\delta, \rho_o, \rho_{1, ...} \rho_{nsat})$ and $\boldsymbol{\theta}_\rho$ were obtained by maximizing the likelihood function of $\mathbf{y}_h$.

Qian et al. [4] employed the Bayesian approach to estimate the mean parameters $(\boldsymbol{\beta}_l, \boldsymbol{\beta}_\rho, \boldsymbol{\beta}_\delta)$, the variance parameters $(\sigma_l^2, \sigma_\rho^2, \sigma_\delta^2)$, and the correlation parameters $(\boldsymbol{\theta}_l, \boldsymbol{\theta}_\rho, \boldsymbol{\theta}_\delta)$ with conjugate prior distribution.

4.3 Prediction Method of the Multi-fidelity Model

In this section, we briefly mention predictive methods at the new interested point based on the multi-fidelity data. Kennedy and O'Hagan [1] applied the Bayesian approach to predict the function value at the new interested point. Forrester et al. [2] predicts a new interested point by using co-Kriging which consists of both high and low fidelity data.

Qian et al. [3] defined the prediction model using the BLUP for LF model $\hat{\mathbf{y}}(\mathbf{x}^*)$ multiplied by a linear regression scale parameter $\hat{\rho}(\mathbf{x}^*)$ plus BLUP predictor of correction model $\hat{\boldsymbol{\delta}}(\mathbf{x}^*)$ at the untried point $\mathbf{x}^*$.

Qian et al. [4] established a predictor using the predictive distribution of the untried point given the high and low fidelity data.

4.4 Results

To investigate the accuracy of each multi-fidelity model introduced above, we compare accuracies of four multi-fidelity models listed in Table 1 using ten mathematical functions. The multi-fidelity models are built using 4*ndv* high fidelity points and 10*ndv* low fidelity points, where *ndv* is the number of design variables. Sample points are generated using the optimal Latin hyper-cube design (OLHD). The accuracy of a Kriging model generated using a high fidelity model only is also listed in Table 1 as a reference. The Kriging model is built using 4*ndv* high fidelity points. We selected ten mathematical functions from a variety of papers. Function 1 is a one-dimensional function from Forrester et al. [2]; Functions 2 and 3 are a one-dimensional function and the modified Branin function, respectively, from Xiong et al. [5]; Functions 4 and 5 are two-dimensional Rosenbrock functions from Eldred et al. [6]; Functions 6 through 8 are the three-dimensional Hartman function, the four-dimensional Wood's function, and the six-dimensional Hartman function, respectively, from Rajnarayan et al. [7]; Functions 9 and 10 are the three-dimensional Harman function and the five-dimensional Ackley's function, respectively, from Huang et al. [8].

We coded four multi-fidelity models by ourselves to compare their accuracy using all ten problems because the previous studies solved only a part of the ten test functions. The root mean square error of approximate function values from the exact values at 1,000 test points is used as accuracy measure. The test points are also generated using the OLHD. From the comparison results shown in Table 1, we observe the following:

- All four multi-fidelity models are better than the Kriging model (HF) for all test functions except for Function 8 (Qian et al. [3] is worse than the Kriging model (HF)) and Function 10 (Forrester et al. [2] and Qian et al. [4] are worse than the Kriging model (HF)).
- Kennedy and O'Hagan [1] model shows the best accuracy among four multi-fidelity models for six test functions

and slightly worse accuracy than the other multi-fidelity models for Functions 3, 4, 6, and 8.

Table 1: Accuracy comparison among existing multi-fidelity models

Test Function	Kriging (HF)	Kennedy and O'Hagan [1]	Forrester et al. [2]	Qian et al. [3]	Qian et al. [4]
Function 1 [2]	5.366	0.017	0.017	1.611	1.675
Function 2 [7]	0.470	0.296	0.296	0.308	0.301
Function 3 [7]	38.506	25.642	26.195	22.971	29.077
Function 4 [8]	1119.164	4.562	4.155	14.016	355.844
Function 5 [8]	631.643	73.335	74.017	83.519	363.152
Function 6 [9]	0.862	0.493	0.494	0.492	0.454
Function 7 [9]	82.917	28.387	28.555	37.476	43.553
Function 8 [9]	0.352	0.326	0.325	0.420	0.316
Function 9 [10]	0.862	0.236	0.238	0.249	0.258
Function 10 [10]	0.765	0.408	0.803	0.409	0.805

5. Global Optimization Using a Multi-Fidelity Model

In this section, we present a constrained global optimization procedure using a multi-fidelity model. The multi-fidelity model we adopt in this study is the Kennedy and O'Hagan model because of its excellent accuracy as described in Section 4.4. In all test problems, the initial sampling points for building the multi-fidelity model employ 3*ndv* high fidelity points and 10*ndv* low fidelity points generated using OLHD.

5.1. Global Optimization Procedure

The proposed global optimization procedure is depicted in Figure 1. The procedure includes two optimization process and two stopping criteria. First, in Optimization Process I, an approximate optimization problem is formulated to maximize the approximate generalized expected improvement (GEI) with satisfying approximate constraints. If converged due to the GEI value in Optimization Process I, then we proceed to Optimization Process II to efficiently obtain an accurate optimization result. This convergence criterion is presented in detail in Section 5.3. In Optimization Process II, an approximate optimization problem is formulated to maximize the approximate objective function with satisfying approximate constraints.

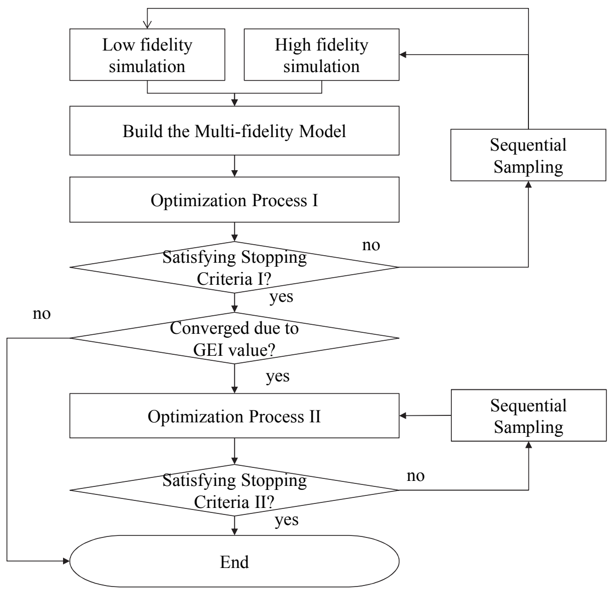


Figure 1: Proposed global optimization procedure

5.2. Sequential Sampling Method

To efficiently find the global optimum using approximate models, it is well known that balancing between the local exploitation and the global exploration is important in determining the next sample point. The local exploitation is needed to improve the accuracy in the region of the optimum predicted by approximate models to find the global optimum quickly [9]. The global exploration is also needed to enhance the overall accuracy of the approximate models using the MSE in case of Kriging models. As such a sequential sampling method, we adopt the generalized expected improvement (GEI) function [10] of the objective function with a cooling scheme. The GEI uses a positive integer parameter g to implement a cooling scheme. When the g value is high, maximizing the GEI finds the next sample point in the region where the prediction error is high. In contrast, maximizing the GEI finds the next sample point in the region of the predicted optimum when the g value is low. In this study, we adopt the cooling scheme proposed by Sasena [11] as shown in Table 2.

Table 2: Cooling scheme

Iteration number	g value
1 – 4	20
5 – 6	10
7 – 9	5
10 – 11	2
12 – 14	1
≥ 15	0

5.3. Stopping Criteria

In Optimization Process I, we use two different stopping criteria according to the g value. If the g value is larger than 1, the stopping conditions are

$$\left(\left[E\left(I^{g}\right)\right]^{1/g}\right)^{i} \leq 1.0e-4 \;\; or \;\; \left\|\mathbf{x}^{i}-\mathbf{x}_{B}\right\| \leq 1.0e-4 \;\; or \;\; \left|\frac{y_{B}-y^{i}}{y_{B}}\right| \leq 1.0e-3 \;\; or \;\; \left|y_{B}-y^{i}\right| \leq 1.0e-4 \quad , \tag{4}$$
$$and \quad g_{j}^{\,i} \leq 3.0e-3, \;\; j=1,2,\ldots,m,$$

where superscript i indicates the current iteration, subscript B indicates the current best (or minimum) value, and m denotes the number of constraints [10]. If the g value is equal to zero, the stopping conditions are

$$\left\|\mathbf{x}^{i}-\mathbf{x}_{B}\right\| \leq 1.0e-4 \;\; or \;\; \left|\frac{y_{B}-y^{i}}{y_{B}}\right| \leq 1.0e-3 \;\; or \;\; \left|y_{B}-y^{i}\right| \leq 1.0e-4 \quad . \tag{5}$$
$$and \quad g_{j}^{\,i} \leq 3.0e-3, \;\; j=1,2,\ldots,m.$$

If Optimization Process I converges due to the condition on the GEI value in Eq. (4), then we proceed to Optimization Process II. The stopping conditions are the same as Eq. (5). Note that the condition on satisfaction of constraints should be always satisfied for convergence of the constrained problem to the optimum solution as stated by "*and*" in Eqs. (4) and (5). For unconstrained problems, we simply omit the condition on satisfaction of constraints.

5.4. Results of Unconstrained Test Problems

To investigate the efficiency and accuracy of the proposed global optimization procedure for unconstrained problems, we solve six unconstrained mathematical problems. We build ten different multi-fidelity models using ten different OLHDs to take the randomness of OLHD. Table 3 lists the exact, maximum, minimum, average, and standard deviation values of the objective function at the global optimum, and the maximum, minimum, and average numbers of HF analyses for each problem. Problem 1 is the modified Branin function from Xiong et al. [5]; Problem 2 is the two-dimensional Rosenbrock function from Eldered et al. [6]; Problem 3 is the four-dimensional Wood's function from Rajnarayan et al. [7]; Problem 4 is the six-dimensional Hartman function from Rajnarayan et al. [7]; Problem 5 is the three-dimensional Hartman function from Huang et al. [8]; and Problem 6 is the five-dimensional Ackley's function from Huang et al. [8]. In order to assess the performance of the proposed global optimization procedure employing the multi-fidelity model of Kennedy and O'Hagan (denoted as MF model in Table 3), we also solved six test problems employing the Kriging model using 3*ndv* high fidelity points only (denoted as Kriging 3*ndv* in Table 3) and the Kriging model using 10*ndv* high fidelity points only (denoted as Kriging 10*ndv* in Table 3).

From Table 3, we observe that the MF model shows comparable accuracy to that of Kriging 10*ndv* with much higher efficiency for all six unconstrained problems, and superior accuracy and efficiency to those of Kriging 3*ndv* for all six unconstrained problems except the accuracy of Problem 5 (even though the difference in accuracy is very small).

5.5. Results of Constrained Test Problems

To investigate the efficiency and accuracy of the proposed global optimization procedure for constrained problems, we solve three two-dimensional constrained mathematical problems presented in Sasena [11]. However, Sasena [11] solved these problems employing high fidelity Kriging models only. Thus, we make low fidelity Kriging models for each problem to build the multi-fidelity models. Just like unconstrained test problems, we build ten different multi-fidelity models using ten different OLHDs to take the randomness of OLHD. The objective and constraint functions of Problems 1, 2, and 3 are depicted in Figures 2, 3, and 4, respectively, and the numerical results of employing the MF model, Kriging 3*ndv*, and Kriging 10*ndv* for Problems 1, 2, and 3 are listed in Tables 4, 5, and 6, respectively.

From Tables 4-6, we observe that the MF model shows comparable accuracy to that of Kriging 10*ndv* or Kriging 3*ndv* with higher efficiency for all three constrained problems.

Table 3: Optimization results of six unconstrained optimization problems

Problem Number		MF Model		Kriging 3*ndv*		Kriging 10*ndv*	
		Opt.Y	No. of Eval.	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.
Problem 1 [5]	Global Opt.	-19.291					
	max	-19.279	20	-19.259	27	-19.288	29
	average	-19.290	18.4	-19.287	23.8	-19.291	27.4
	S.T.D	0.0060	-	0.0102	-	0.0016	-
	min	-19.292	16	-19.292	21	-19.292	26
Problem 2 [6]	Global Opt.	0.					
	max	1.8E-04	19	4.0E-03	29	2.1E-03	32
	average	4.8E-05	17.9	9.1E-04	26.9	2.8E-04	30.5
	S.T.D	6.5E-05	-	1.2E-03	-	6.6E-04	-
	min	2.0E-06	17	2.1E-07	25	6.0E-07	30
Problem 3 [7]	Global Opt.	0.					
	max	4.0E+00	28	1.7E+01	30	4.0E+00	56
	average	8.0E-01	24.7	8.2E+00	26.5	1.2E+00	51.8
	S.T.D	1.7E+00	-	5.6E+00	-	1.9E+00	-
	min	7.9E-08	15	3.6E-10	17	1.4E-08	42
Problem 4 [7]	Global Opt.	-3.322					
	max	-3.018	42	-2.285	42	-2.659	77
	average	-3.155	35.5	-2.837	35.1	-3.132	74.7
	S.T.D	0.087	-	0.370	-	0.203	-
	min	-3.256	33	-3.256	33	-3.318	70
Problem 5 [8]	Global Opt.	-3.863					
	max	-3.799	24	-3.849	25	-3.844	46
	average	-3.853	20.6	-3.854	22.2	-3.858	42.6
	S.T.D	0.0189	-	0.0030	-	0.0059	-
	min	-3.863	18	-3.859	20	-3.863	40
Problem 6 [8]	Global Opt.	0.					
	max	3.8E-02	33	1.6586	115	7.8E-02	73
	average	2.1E-02	31.5	0.3262	56.9	2.4E-02	67.4
	S.T.D	1.1E-02	-	0.065	-	3.1E-02	-
	min	4.4E-03	31	0.0038	34	4.9E-04	65

6. Conclusion

In this study, we proposed a constrained global optimization procedure using a multi-fidelity model with sequential sampling. First, we compared the accuracies of four existing multi-fidelity models to adopt the best one in the proposed procedure. The multi-fidelity model adopted was that proposed by Kennedy and O'Hagan employing a Gaussian process low fidelity model multiplied by a constant scale parameter plus a Gaussian process correction model. The proposed procedure included two optimization process and two stopping criteria. In Optimization Process I, an approximate optimization problem was formulated to maximize the approximate generalized expected improvement (GEI) with satisfying approximate constraints to find a sequential sampling point. If converged due to the GEI value in Optimization Process I, then we proceeded to Optimization Process II to efficiently obtain an accurate optimization result.

We compared the performance of the proposed procedure employing the MF model to those of the procedure employing Kriging models using 3*ndv* and 10*ndv* high fidelity points only for six unconstrained and three constrained problems. Comparison results for unconstrained problems revealed that the MF model had comparable accuracy to that of Kriging 10*ndv* with much higher efficiency for all six unconstrained problems, and superior accuracy and efficiency to those of Kriging 3*ndv* for all six unconstrained problems except the accuracy of

Problem 5. Comparison results for constrained problems revealed that the MF model shows comparable accuracy to that of Kriging 10*ndv* or Kriging 3*ndv* with higher efficiency for all three constrained problems. From the comparison results, we can observe that the numerical performance of the proposed procedure employing the multi-fidelity model is better than those employing Kriging models using high fidelity points only.

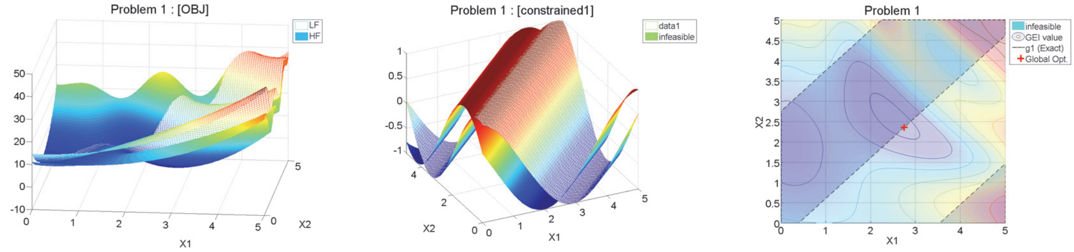


Figure 2: The objective function and a constraint function of Problem 1

Table 4: Optimization results of Problem 1

	MF Model		Kriging 3*ndv*		Kriging 10*ndv*	
	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.
Global Opt.	-1.1743					
max	-1.1620	25	4.1972	26	-1.1740	39
average	-1.1727	22.1	0.5672	20.5	-1.1742	35.5
S.T.D	0.0040	-	2.5203	-	0.0001	-
min	-1.1760	20	-1.1743	15	-1.1743	33

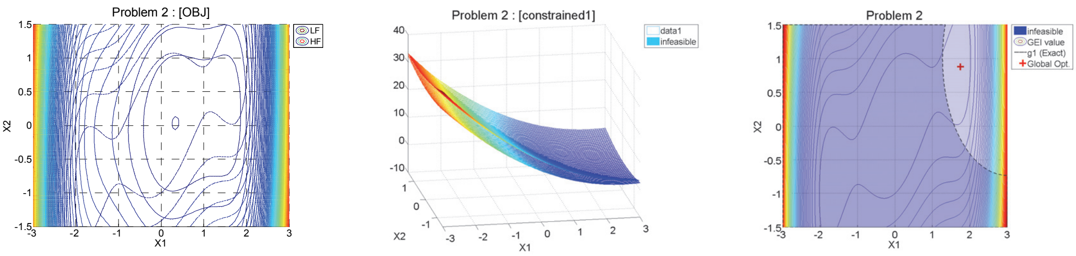
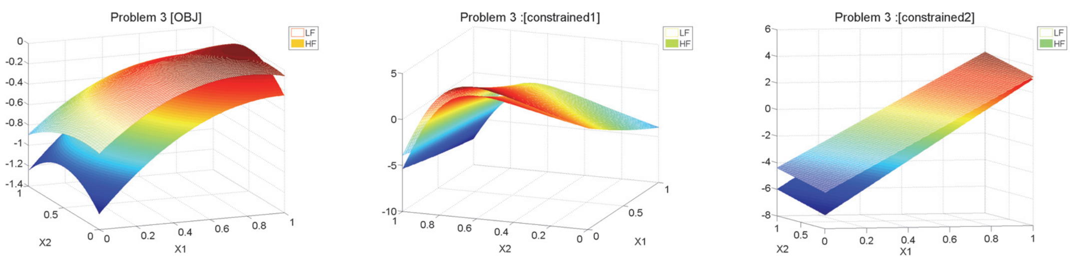


Figure 3: The objective function and a constraint function of Problem 2

Table 5: Optimization results of Problem 2

	MF Model		Kriging 3*ndv*		Kriging 10*ndv*	
	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.
Global Opt.	0.2986					
max	0.3062	17	0.2987	21	0.2994	31
average	0.3002	14.2	0.2987	19.3	0.2988	28.9
S.T.D	2.7E-03	-	4.2E-05	-	2.4E-04	-
min	0.2986	12	0.2986	18	0.2986	27



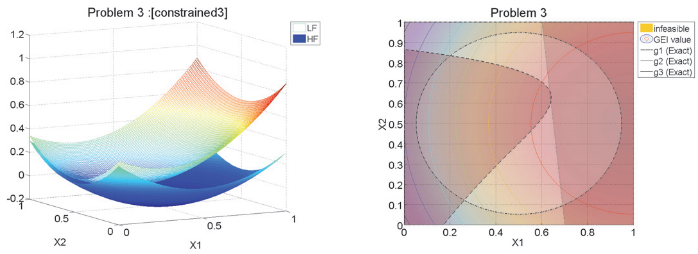


Figure 4: The objective function and three constraint functions of Problem 3

Table 6: Optimization results of Problem 3

	MF Model		Kriging 3*ndv*		Kriging 10*ndv*	
	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.	Opt.Y	No. of Eval.
Global Opt.	-0.7483					
max	-0.7433	12	-0.7414	16	-0.7092	40
average	-0.7475	9.5	-0.7475	13.5	-0.7381	25.2
S.T.D	1.7E-03	-	2.2E-03	-	1.6E-02	-
min	-0.7483	8	-0.7483	11	-0.7483	23

7. Acknowledgements

This work was supported by grants from "Development of the Prototyping Ball Bearings for a Rocket Turbopump" project of Ministry of Education Science and Technology (MEST) (No. 2012-0000629) and the National Research Foundation of Korea (NRF) grant funded by the Korea government(MEST) (20120005530).

8. References

[1] Kennedy M. C. and O'Hagan A., Predicting the output from a complex computer code when fast approximations are available, *Biometrika*, 87 (1), 1-13, 2000.

[2] Alexander I. J. Forrester, Andras Sobester and Andy J. Keane, Multi-fidelity optimization via surrogate modeling, *Proceedings of the Royal Society A*, 463, 3251-3269, 2007.

[3] Zhiguang Qian, Carolyn Conner Seepersad, V. Rhoshan Joesph, Janet K. Allen and C. F. Jeff Wu, Building surrogate models based on detailed and approximate simulations, *Journal of Mechanical Design*, 128, 668-677, 2006.

[4] Zhiguang Qian and C. F. Jeff Wu, Bayesian hierarchical modeling for integrating low-accuracy and high-accuracy experiments, *Technometrics*, 50, 192-204, 2008.

[5] Ying Xiong, Wei Chen and Kwok-Leung Tsui, A new variable fidelity optimization framework based on model fusion and objective-oriented sequential sampling, *Journal of Mechanical Design*, 130, 111401-1 – 111401-9, 2008.

[6] M. S. Eldred, A. A. Giunta and S. S. Collis, Second-Order corrections for surrogate-based optimization with model hierarchies, *Proceedings of the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Albany, New York, 2004.

[7] Dev Rajnarayan, Alex Hass and Ilan Kroo, A Multifidelity gradient-free optimization method and application to aerodynamic design, *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Victoria, British Columbia, Canada, 2008.

[8] D. Huang, T. T. Allen, W. I. Notz and R. A. Miller, Sequential Kriging optimization using multiple-fidelity evaluations, *Structural and Multidisciplinary Optimization*, 32, 369-382, 2006.

[9] Donald R. Jones, Matthias Schonlau and William J. Welch, Efficient Global Optimization of Expensive Black-Box Functions, *Journal of Global Optimization*, 13, 455-492, 1998.

[10] Matthias Schonlau, William J. Welch and Donald R. Jones, Global versus local search in constrained optimization of computer models, *New Developments and Applications in Experimental Design IMS Lecture Notes*, Monograph Series 34, 1998.

[11] Michael James Sasena, *Flexibility and efficiency enhancements for constrained global design optimization with Kriging approximation*, Ph. D. thesis, University of Michigan, 2002.