Chinese Journal of Aeronautics, 2013,26(1): 112–121

Chinese Society of Aeronautics and Astronautics
& Beihang University

Chinese Journal of Aeronautics

cja@buaa.edu.cn
www.sciencedirect.com

# Stochastic structural optimization using particle swarm optimization, surrogate models and Bayesian statistics

Jongbin Im [a], Jungsun Park [b,*]

[a] Korea Aerospace Research Institute, Daejeon 305-333, Republic of Korea
[b] Department of Aerospace and Mechanical Engineering, Korea Aerospace University, Koyang 412-791, Republic of Korea

Received 18 January 2012; revised 20 March 2012; accepted 18 July 2012
Available online 16 January 2013

KEYWORDS

Bayesian statistics;
Kriging;
Particle swarm optimization (PSO);
Response surface method (RSM)

**Abstract** This paper focuses on a method to solve structural optimization problems using particle swarm optimization (PSO), surrogate models and Bayesian statistics. PSO is a random/stochastic search algorithm designed to find the global optimum. However, PSO needs many evaluations compared to gradient-based optimization. This means PSO increases the analysis costs of structural optimization. One of the methods to reduce computing costs in stochastic optimization is to use approximation techniques. In this work, surrogate models are used, including the response surface method (RSM) and Kriging. When surrogate models are used, there are some errors between exact values and approximated values. These errors decrease the reliability of the optimum values and discard the realistic approximation of using surrogate models. In this paper, Bayesian statistics is used to obtain more reliable results. To verify and confirm the efficiency of the proposed method using surrogate models and Bayesian statistics for stochastic structural optimization, two numerical examples are optimized, and the optimization of a hub sleeve is demonstrated as a practical problem.



## 1. Introduction

Various optimization algorithms have been presented over the past few decades and have continuously improved in efficiency. Nowadays, stochastic optimization algorithms such as the genetic algorithm (GA), simulated annealing (SA) and particle swarm optimization (PSO) have been applied to many design fields as computing power has improved.[1–3] Stochastic optimization algorithms have their own mechanism to find the global optimum. However they need lots of evaluations to reach the global optimum. Therefore, to reduce the evaluation costs inherent in stochastic algorithms, approximation methods such as response surface method (RSM) and Kriging have been applied to structural optimization.

In this paper, a methodology for structural optimization is proposed that combines PSO, RSM, Kriging, and Bayesian statistics. PSO is based on a simplified social model that is closely tied to swarming theory. A physical analogy might be a swarm of bees or ants searching for a good food source. In this mechanism, each bee or ant is called a particle in PSO, which uses its own information as well as collective knowledge gained by the swarm to find the best available food source. PSO has lots of advantages over other optimization algorithms. First, PSO is generally easy to program. Second, PSO can efficiently

* Corresponding author. Tel.: +82 2 300 0283.
E-mail addresses: jongbinim@gmail.com (J. Im), jungsun@kau.ac.kr (J. Park).

Peer review under responsibility of Editorial Committe of CJA.

ELSEVIER Production and hosting by Elsevier

1000-9361 
http://dx.doi.org/10.1016/j.cja.2012.12.022

make use of large numbers of processors and does not require the initial values of design variables. In addition, PSO is generally better suited for finding the global solution and is ideally suited for solving discrete, continuous and/or combinatorial types of optimization problems.[3,4] Zhao et al.[5] presented a visual modeling method describing a particle's dimensional vector behavior. Based on an analysis of visual modeling, they explained the reason for premature convergence and diversity loss in PSO, and also proposed a new modified algorithm to ensure the rational flight of every particle's dimensional component. Also, they suggested the addition of two parameters of particle-distribution-degree and particle-dimension-distance to the proposed algorithm in order to avoid premature convergence. Simulation results of the new PSO algorithm showed that it has a better ability to find the global optimum and still keeps a rapid convergence, as with the standard PSO.

The response surface method (RSM) has been put to practical use in the field of quality engineering for purposes such as product process optimization and variation reduction.[6] The RSM is a type of surrogate model that applies an approximation technique to functions such as objective and constraints in an optimization problem. For approximation, it uses a function called a response surface, which is a function that approximates a problem with design variables and state quantities, using several analysis or experimental results. In general, the design of experiments is used for analysis or parameter setting of experiment point, and the least square method is used for function approximation.[6]

Kriging was originally developed and used in mining engineering and geostatistics. Kriging is an approach for curve fitting. With a statistical technique based on geostatistics, Kriging is useful to estimate spatial interaction of various data.[7,8] Diverse approximation models by Kriging are known to be suitable for non-linear models. The Kriging model is influenced by sampling data. In the Kriging model, the interpolation error is affected by the distance between the sample points. Consequently, efficient sampling processes are needed to obtain maximum information with a minimum number of design experiments.[9]

RSM and Kriging have different mechanisms to generate approximation models, and they do not produce the same approximated values of design problems, especially in nonlinear cases. To consider the different approximation abilities of the two models, a hybrid using both RSM and Kriging is developed for the evaluations of PSO. And we will discuss the effects of using hybrid approach. In this paper, Bayesian statistics is also adopted in order to suggest more reliable optimum values. Bayesian statistics uses the data of design variables generated by PSO during the optimization process. To verify the proposed method using surrogate models and Bayesian statistics for stochastic structural design problems, two- and four-bar truss problems are selected from Refs. [10,11] as numerical examples and the optimization of a hub sleeve is demonstrated as a practical problem.

## 2. PSO

PSO is a parallel population-based computation technique first proposed by Eberhart and Kennedy.[3] It is motivated by the behavior of certain organisms, such as the schooling of fish and the flocking of birds. PSO can solve a variety of difficult optimization problems. It uses the physical movements of the particles in the swarm and has a flexible and well-balanced mechanism to enhance and adapt to global and local search abilities. Another advantage of PSO is its simplicity in coding and consistency in performance. The global optimization model proposed by Shi and Eberhart is as follows[12]:

$$\boldsymbol{V}_{i+1} = W\boldsymbol{V}_i + C_1 r_1 (\boldsymbol{P}_{\text{best}} - \boldsymbol{X}_i) + C_2 r_2 (\boldsymbol{G}_{\text{best}} - \boldsymbol{X}_i) \tag{1}$$

$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i + \boldsymbol{V}_{i+1} \tag{2}$$

where $\boldsymbol{V}_i$ is the $i$th particle velocity, $\boldsymbol{X}_i$ the particle position, $W$ the inertial weight; $C_1$ and $C_2$, which affect the convergence of optimization, are positive constant parameters; $r_1$ and $r_2$ are random values in the range [0,1] obtained by a random number generator; $\boldsymbol{P}_{\text{best}}$ is the best position of $i$th particle, and $\boldsymbol{G}_{\text{best}}$ is the best position among all particles in the swarm.

### 2.1. PSO parameters

The parameters $W$, $C_1$, and $C_2$ in Eq. (1) are important to find the optimum value and achieving good convergence. The basic PSO algorithm has three problem-dependent parameters, $W$, $C_1$ and $C_2$. Ref. [3] proposes using $C_1 = C_2 = 2$. Additionally, Shi and Eberhart suggest using $0.8 < W < 1.4$, starting with larger $W$ values (a more global search behavior) that are dynamically reduced (a more local search behavior) during the optimization.[12] The scheme, which dynamically adjusts the $W$ value, is as follows[12]:

$$W = w_{\max} - \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}} \text{iter} \tag{3}$$

where $w_{\max}$ is 0.9 and $w_{\min}$ 0.4. iter denotes the current iteration number, and $\text{iter}_{\max}$ is the maximum number of iterations, which is a user-defined value.

### 2.2. Neighborhood topology

In the swarm, all the particles are neighbors of each other. The position of the best particle ($\boldsymbol{G}_{\text{best}}$) in the swarm is used in the velocity vector Eq. (1). It is assumed that the best particle converges quickly, as all the particles are attracted simultaneously to the best part of the search space. However, if we use single $\boldsymbol{G}_{\text{best}}$ for all particles, and $\boldsymbol{G}_{\text{best}}$ does not lead to the global optimum but a local optimum, it may be impossible for the swarm to explore other areas; this means that the swarm can be trapped in local optimum. To address this limitation, neighborhood topology is used. In this case, only a specific number of particles can affect the velocity of a given particle. The swarm will converge more slowly but has a greater chance of locating the global optimum.[13] In this study two kinds of neighborhood topologies are used: circular topology and random topology (Fig. 1).

## 3. Surrogate models

### 3.1. RSM

The response surface is an approximation of the relational expression of the response $y$ predicted from variables $x_i$ (where $i = 1,2,\cdots,n$). The response surface can be expressed by

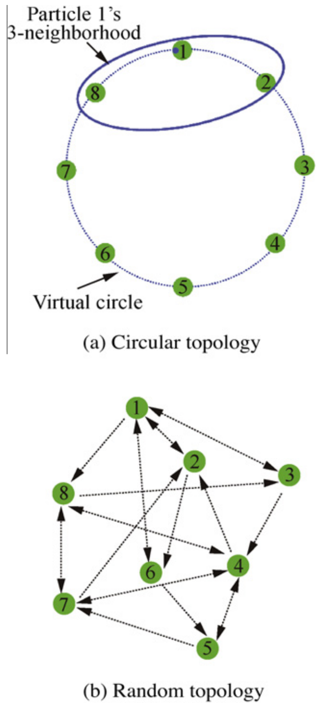


**Fig. 1** Neighborhood topology.

$$y = f(x_1, x_2, \dots, x_n) + \varepsilon \tag{4}$$

where $\varepsilon$ is an error assumed to have a zero mean.

In general, for this function $f$, a polynomial is often used because it is easier to handle. For response surfaces, linear functions are advantageous because their coefficients can be determined easily using the least square method, and statistical evaluation can be conducted on them once their coefficients have been determined. For this reason, function approximations using the least square method are used most often with the RSM. If a quadratic polynomial is expressed as a response function, the response surface can be given by

$$y = \beta_0 + \sum_{i=1}^{n}\beta_i x_i + \sum_{i=1}^{n}\beta_{ii}x_i^2 + \sum_{i<j}^{n}\beta_{ij}x_i x_j \tag{5}$$

Simple representation of Eq. (5) with two variables can be written as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 \tag{6}$$

If we replace the second degree terms with single variables $\left(x_1^2 = x_3,\ x_2^2 = x_4,\ x_1 x_2 = x_5\right)$ in Eq. (6), respectively, this expression is converted into a multi-variable, linear expression. Such conversion is applicable to a third or higher-degree polynomial. If linearization is performed in this way, a linear regression model can be represented by Eq. (7), assuming that the number of experiment points is $n$ and the number of design variables is $k$.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{7}$$

where

$$\boldsymbol{y} = [\,y_1 \quad y_2 \quad \cdots \quad y_n\,]^{\mathrm{T}}$$

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = [\,\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n\,]^{\mathrm{T}}$$

$$\boldsymbol{\varepsilon} = [\,\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n\,]^{\mathrm{T}}$$

By minimizing the sum of squares of error $\boldsymbol{\varepsilon}$, the unbiased estimator $\boldsymbol{b}$ of the coefficient $\boldsymbol{\beta}$ can be calculated using Eq. (8).

$$\boldsymbol{b} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y} \tag{8}$$

In this work, the quadratic polynomial is used for numerical examples and practical problem because it is efficient compared to linear or higher-degree polynomials. It can predict the nonlinear behavior of response function rather than linear polynomials, and it can reduce computing time compared to higher-degree polynomials.

### 3.2. Kriging

Kriging, an approximation model, was originally developed and used in mining engineering and geostatistics. In the 1980s, statisticians developed design and analysis of computer experiments (DACE) for deterministic computer-generated data based on Kriging. The Kriging model used in this work, which is described as follows, is based on the DACE approach.[14]

Kriging is represented as a combination of a global model and departures as shown in the following equation:

$$\boldsymbol{Y}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{Z}(\boldsymbol{x}) \tag{9}$$

where $\boldsymbol{Y}(\boldsymbol{x})$ is an unknown function of interest, $\boldsymbol{F}(\boldsymbol{x})$ a known regression model, and $\boldsymbol{Z}(\boldsymbol{x})$ a realization model of a stationary random process with mean zero and variance, $\sigma^2$. While $\mathbf{F}(\boldsymbol{x})$ globally approximates the design space, $\boldsymbol{Z}(\boldsymbol{x})$ creates localized deviations so that the Kriging model interpolates the sampled data points, $n_s$. The covariance matrix of $\boldsymbol{Z}(\boldsymbol{x})$ is given by

$$\mathbf{Cov}[\boldsymbol{Z}(\boldsymbol{x}^i), \boldsymbol{Z}(\boldsymbol{x}^{i+1})] = \sigma^2 \boldsymbol{R}[R(x^i, x^{i+1})] \tag{10}$$

where $\boldsymbol{R}[\cdot,\cdot]$ is the correlation matrix and $R(x^i, x^{i+1})$ the correlation function between two points of the sampled data points, $x^i$ and $x^{i+1}$. $\boldsymbol{R}[\cdot,\cdot]$ is an $n_s \times n_s$ symmetric matrix with diagonal elements equal to 1.0. The correlation function $R(x^i, x^{i+1})$ is provided by the user.

The predicted values, $\widehat{\boldsymbol{Y}}$ of the response $\boldsymbol{Y}$ at the untried values of $\boldsymbol{x}$, are given by

$$\widehat{\boldsymbol{Y}} = \hat{\boldsymbol{\beta}} + \boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{Y} - \hat{\boldsymbol{\beta}}\boldsymbol{F}) \tag{11}$$

where $\boldsymbol{Y}$ is the column vector of length $n_s$ which contains the response at each sample point, and $\boldsymbol{F}$ is a column vector of length $n_s$ that has elements equal to 1.0, when $\boldsymbol{F}(\boldsymbol{x})$ is taken as a constant. In Eq. (11), $\boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})$ is the correlation vector of length $n_s$ between an untried $\boldsymbol{x}$ and the sampled data points

$[\boldsymbol{x}^1\boldsymbol{x}^2\cdots\boldsymbol{x}^{n_s}]$. $\hat{\boldsymbol{\beta}}$ is estimated using least squares regression in Eq. (12).

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F})^{-1}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{Y}) \tag{12}$$

In this work, the exponential correlation function (ECF) is used. When the distance between two points is short, the ECF is more efficient than a Gaussian correlation function that is popular to be used in Kriging. An ECF can be represented by

$$R(x^i, x^{i+1}) = \prod_{k=1}^{N_{\mathrm{dv}}} \exp(-\theta_k |d_k|) \tag{13}$$

where $d_k$ is the distance between already known sampled points, $N_{\mathrm{dv}}$ the number of design variables, and $\theta_k$ the unknown correlation parameters. The values of $\theta_k$ for each response are obtained by solving an unconstrained optimization problem and can be expressed as

$$\mathrm{Max}\, f(\theta_k) = -\frac{n_{\mathrm{s}} \ln \hat{\sigma}^2 + \ln|\boldsymbol{R}|}{2} \tag{14}$$

where $\hat{\sigma}^2$ is the estimated variance between the global model $\boldsymbol{F}(\boldsymbol{x})$ and actual response $\widehat{\boldsymbol{Y}}$.

The estimate of the variance is given by

$$\hat{\sigma}^2 = \frac{(\boldsymbol{Y} - \hat{\boldsymbol{\beta}}\boldsymbol{F})^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{Y} - \hat{\boldsymbol{\beta}}\boldsymbol{F})}{n_{\mathrm{s}}} \tag{15}$$

## 4. Sequential approximation method

When surrogate models are used in optimization, the concept of sequential approximation is considered because it is difficult to express the real phenomenon of response such as the objective function and constraint functions using one approximation model for the whole design space. In a sequential approach, surrogate models are regenerated for each iteration using a reduced design space called a trust region. The trust region is the small part of the design space including the best value of the objective function in the current iteration of the optimization process. We call the process to determine the trust region the move limit strategy.

The choice of the move limit strategy has a great influence on the behavior of the sequential optimization process. Especially if the surrogate models are less accurate, the move limit strategy is usually of vital importance. A move limit strategy has to determine the size of the search trust region at the start of each new design cycle. A correct choice of the search trust region is important for a good convergence of the optimization process. Large move limits can cause the solution process to oscillate, while small move limits may slow down the convergence. The effect of the move limit strategy is directly related to the quality of the approximations. Poor approximations have a greater need for the support of the move limit strategy during the optimization, compared to high quality approximations. The proposed move limit strategy to determine the upper and lower bounds of the trust region in this paper is expressed in the following equations.

The lower bound of trust region is

$$x_{k+1}^{\mathrm{l}} = \mu_k - z_{\alpha/2}\sigma_k \tag{16}$$

The upper bound of trust region is

$$x_{k+1}^{\mathrm{u}} = \mu_k + z_{\alpha/2}\sigma_k \tag{17}$$

where $\mu_k$ indicates the optimum values of design variables, and $\sigma_k$ is the standard deviations of design variables that are obtained from optimization results by PSO. $z_{\alpha/2}$ is the value obtained from standard normal distribution table. In this paper, the confidence interval 95 % is used, which means $z_{\alpha/2} = 1.96$ from the standard normal distribution table. Fig. 2 shows the move limit strategy used in this paper.

## 5. Bayesian statistics

When surrogate models and the penalty function method, which will be discussed in the next section, are used, there are some errors between exact values and approximated values. Sometimes, these errors affect the reliability of optimum values. Therefore, to reduce the effects of the errors and to suggest the reliable value of optimum that is safer, in this work, Bayesian statistics is adopted.

Bayesian methods start with the prior knowledge of known data obtained from experiments or simulations. As new data are obtained from other experiments or simulations, the prior knowledge is updated, and the new data is used to obtain the posterior knowledge. Bayesian methods[15] can be used for the statistical analysis of data. There are two types of Bayesian methods that are used in statistical analysis. The first type deals with a random variable $X$ that is normally distributed with a known standard deviation. The mean value of the random variables is of interest and is estimated using Bayesian methods. In the second type, the random variable $X$ is also normally distributed, but its standard deviation is unknown. In this type, both the mean value and the variance of the random variable are of interest and are estimated using Bayesian methods. In optimization problems, we can estimate mean value from the best or optimum value, but we do not know the variance of the estimated mean value. Therefore, in this study the second type of Bayesian method is used.

In this type of Bayesian statistics, the random variable $X$ is considered to be normally distributed with an unknown mean value $\mu$ and unknown variance $\sigma^2$. Both the mean value and variance of the random variable are of interest but are unknown. The prior joint distribution of the unknown mean and unknown variance is assumed to be normal-gamma, which is defined as the product of a normal distribution for the mean and a gamma distribution for the variance. The prior informa-

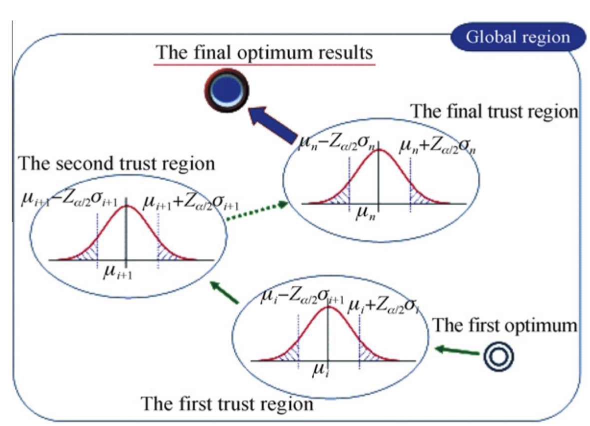


**Fig. 2** Move limit strategy.

tion about the mean and variance is based on a sample of size $N$ with sample mean and variance of $\bar{\mu}_0$ and $\sigma_0^2$, respectively. New information is obtained by a sample of size $q$. The mean value and variance based on the sample $n$ are $\bar{\mu}_{\mathrm{new}}$ and $\sigma^2_{\mathrm{new}}$, respectively. We are interested in determining the posterior distribution of the mean and variance. It can be shown that the posterior distribution is also normal-gamma. The posterior mean $\mu_{\mathrm{pos}}$ and posterior variance $\sigma^2_{\mathrm{pos}}$ can be shown as[16]

$$\mu_{\mathrm{pos}} = \frac{N\bar{\mu}_0 + q\bar{\mu}_{\mathrm{new}}}{m} \tag{18a}$$

and

$$\sigma^2_{\mathrm{pos}} = \frac{(N-1)\sigma_0^2 + N\bar{\mu}_0^2 + (q-1)\sigma^2_{\mathrm{new}} + q\bar{\mu}^2_{\mathrm{new}} - m\mu^2_{\mathrm{pos}}}{m-1} \tag{18b}$$

where

$$m = N + q \tag{18c}$$

The resulting values of Eqs. (18a) and (18b) are the posterior mean and variance of the unknown mean and variance, respectively. The posterior mean is the reliable value of optimum results obtained from PSO. We use this Bayesian method during optimization process in order to obtain reliable optimum results.

## 6. Numerical examples

Two test problems, two- and four-bar truss, were selected from the literature to verify the ability and effect of the proposed stochastic structural optimization. Each example is described along with its constraints, bounds, and objective function. In this paper the same parameters are used in PSO. The maximum number of iterations is 1000, and the swarm size is 30 for each problem. The procedure of optimization proposed in this paper is shown in Fig. 3.

To generate surrogate models, sample data are needed, and they are important to obtain a reliable approximation function. Therefore, in this paper, central composite design (CCD) is used as a design of experiments (DOE) to generate simulation points of design variables.[17] There are more efficient DOEs than CCD such as adaptive sampling for reducing the number of samples. That will be considered in future work. After generating sample data using DOE that is the first step, we make the surrogate models for an objective function and constraint functions of optimization that is the second step. The third step of the proposed optimization process is to achieve an optimization to find the minimum value of the generated objective function. The fourth step is to perform Bayesian statistics, which is an optional step to obtain reliable results. We obtain only PSO results when the fourth step is not used in the process. This procedure is repeated until the objective function value converges. Move limit strategy as mentioned in the previous section is used in the step of generating sample data using DOE.

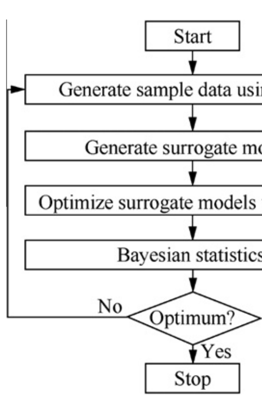


**Fig. 3** Flowchart of the proposed optimization process.

### 6.1. Penalty function method

PSO is for unconstrained optimization problems, therefore some methods are required to change constrained problem to unconstrained. The most widely used approach to handle constraints in optimization algorithms that cannot deal with constrained problems is to use penalties. The idea of the penalty function method is to change the formulation of constrained optimization problems to that of unconstrained problems by adding or subtracting a certain value from the objective function based on the amount of constraint violation present in a certain solution.[11] The general formulation of the exterior penalty function is as follows[18]:

$$\phi(x) = f(x) \pm \left( \sum_{i=1}^{n} r_i G_i + \sum_{j=1}^{p} c_j L_j \right) \tag{19}$$

where $\phi(x)$ is the new objective function to be optimized, called a pseudo-objective function, $G_i$ and $L_j$ are functions of the constraints $g_i(x)$ and $h_j(x)$, respectively, and $r_i$ and $c_j$ are positive constants normally called "penalty parameters".

The most common forms of $G_i$ and $L_j$ are

$$G_i = \max(0, g_i(x))^{\beta} \tag{20}$$

$$L_j = |h_j(x)|^{\gamma} \tag{21}$$

where $\beta$ and $\gamma$ are normally 1 or 2. In this study a value of 2 is selected. In PSO, the unconstrained function $\phi(x)$ is optimized.

### 6.2. Hybrid using of RSM and Kriging

In this paper, two surrogate models are presented, RSM and Kriging, and numerical examples are optimized using the two models, respectively. And a hybrid using both RSM and Kriging is developed for the evaluations of PSO. This approach is expressed in Fig. 4 (in which $K$ is the number of neighbors). For example, there are 20 particles (this means that the swarm size is 20). When the particle has an odd number, RSM is used to evaluate the objective function, and Kriging is used to evaluate the objective function when the particle has an even number.

### 6.3. Two-bar truss

The two-bar truss problem is introduced by Morris and tested by Rao.[10] The example shown in Fig. 5 is subject to a vertical load $2P$ and is to be designed for the minimum weight. The members have a tubular section with mean diameter $d$ and wall thickness $t$. The maximum permissible stress in each member ($\sigma_0$) is equal to 413.68 MPa. In this problem, design variables

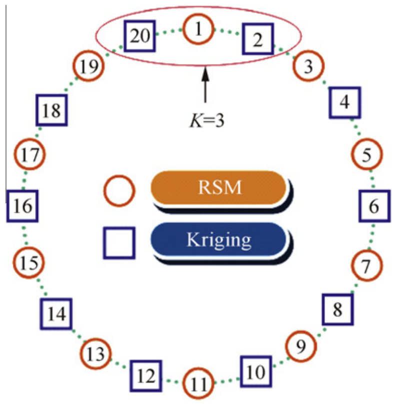


**Fig. 4** Function evaluations using two surrogate models.

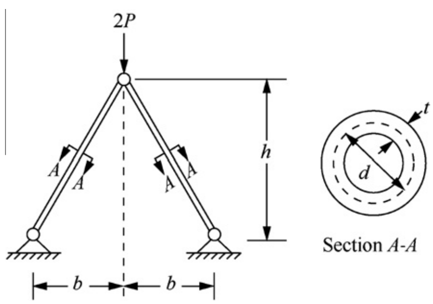


**Fig. 5** Two-bar truss.

are $h$ and $d$. This problem has data that are $P = 147$ kN, $t = 0.54$ mm, $b = 762$ mm, $\sigma_0 = 413.68$ MPa, and density $\rho = 8303.97$ kg/m$^3$. Optimization formulation is described as

$$\begin{aligned} \text{Min } f(d,h) &= 2\rho\pi dt\sqrt{b^2+h^2} \\ &= 2(0.3)\pi d(0.1)\sqrt{900+h^2} \\ &= 0.188d\sqrt{900+h^2} \end{aligned} \tag{22}$$

$$\text{s.t.} \quad G_1 = 1.75\frac{\sqrt{900+h^2}}{dh} - 1 \leqslant 0 \tag{23}$$

As seen in Table 1, the best optimum value of the objective function is 8.758 kg, which is obtained using RSM with circular topology in PSO. All the results using RSM have errors compared to the exact value that is calculated with Eqs. (22) and (23). In Table 1, "Rand' represents the 'random topology and 'Cir' the 'circular topology' as shown in Fig. 1.

The reason RSM has errors is believed to result from the fact that the equation of the objective function has a non-linear term, $h^2$. Kriging also seems to have errors compared to the exact values, but the error is substantially less than that using RSM. This result confirms once again that Kriging is more useful to express non-linear functions than RSM. The hybrid approach is also applied in this work. As shown in Table 1, the error between the exact objective function value and the estimated value calculated from surrogate models shows us that the hybrid approach has smaller errors than the approach using RSM or Kriging. From this result, the hybrid using RSM and Kriging is more useful than using only one of the surrogate models such as RSM and Kriging in this problem.

The constraint values when the optimization is achieved are shown in Table 2. As shown in Table 2 the constraint ($G_1$) of using RSM, Kriging, and the hybrid exceeds the limit bound, which is $G_1 \leqslant 0$. Consequently, these optimum values are not acceptable for designers, and it is believed that this error comes from the approximation model's error and penalty parameter's error. Therefore, to obtain reliable and acceptable optimum, the penalty parameters are changed and the optimization should be achieved again. This means we need extra costs for stochastic structural optimization. In this paper, Bayesian statistics are also performed to obtain more reliable optimum. The optimum value of the objective function by Bayesian statistics is 9.405 kg as seen in Table 1. The value of the objective function is bigger by 0.606 kg than the optimum of the hybrid that is 6.44% over. However, the constraint ($G_1$) seems more reliable than the hybrid's value, as seen in Table 2, because it does not violate the limit condition ($G_1 \leqslant 0$). Therefore, according to the safety of constraint, the results of Bayesian statistics give the designer more reliability without extra optimization costs.

Figs. 6 and 7 show the variation of mean and variance of design variables obtained by Bayesian statistics during the optimization process. From the figures, the probability density function (PDF) of mean of design variables decreases, and the variance of design variable increases as the optimum converges. It means that Bayesian gives us more robust results compared with the optimum results by PSO to the extent that the previous values of design are considered.

### *6.4. Four-bar truss*

The four-bar truss problem is designed to minimize the weight of the four bars, and the configuration is shown in Fig. 8.[11] The constraints are stresses in the members and a displacement constraint at the tip joint of the truss. For simplicity, it is assumed that Members 1–3 have the same cross-section area

**Table 1** Optimization results of two-bar truss.

| Method | PSO | | | | | | | | Bayesian |
|---|---|---|---|---|---|---|---|---|---|
| | RSM | | Kriging | | RSM & Kriging | | Rand | Cir | |
| | Rand | Cir | Rand | Cir | Rand | Cir | | | |
| Objective function (kg) | 8.767 | 8.758 | 8.814 | 8.812 | 8.799 | 8.814 | 8.816 | 8.816 | 9.405 |
| Design variables $h$ (mm) | 760.349 | 762.762 | 781.101 | 791.286 | 780.923 | 765.023 | 762.000 | 762.254 | 845.668 |
| $d$ (mm) | 60.249 | 60.147 | 60.350 | 59.766 | 61.163 | 60.858 | 60.960 | 60.960 | 62.509 |
| Exact objective function (kg) | 8.571 | 8.573 | 8.706 | 8.677 | 8.821 | 8.687 | | | |
| Error (%) | 2.238 | 2.115 | 1.227 | 1.536 | 0.242 | 1.445 | – | – | – |

**Table 2** Constraint value at the optimum.

| Method | PSO | | | | | | | | Bayesian |
|---|---|---|---|---|---|---|---|---|---|
| | RSM | | Kriging | | RSM & Kriging | | Rand | Cir | |
| | Rand | Cir | Rand | Cir | Rand | Cir | | | |
| Constraint ($G_1$) | 0.0462 | 0.0451 | 0.0289 | 0.0322 | 0.0154 | 0.0307 | 0.0312 | 0.0310 | −0.0428 |
| Exact constraint | 0.0447 | 0.0445 | 0.0288 | 0.0327 | 0.0154 | 0.0308 | – | – | – |

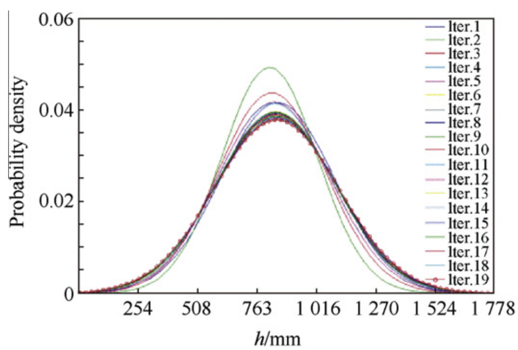


**Fig. 6** History of design variable ($h$) of Bayesian statistics.

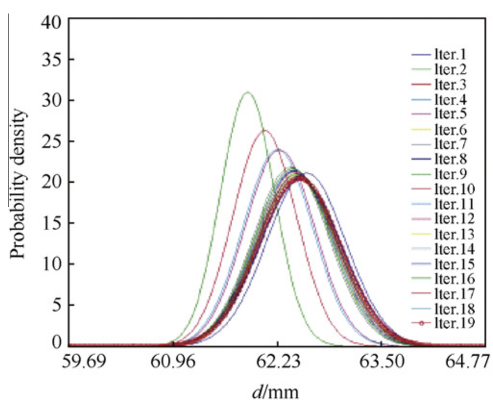


**Fig. 7** History of design variable ($d$) of Bayesian statistics.

$A_1$, and Member 4 has the cross-section area $A_2$. The minimum weight design optimization subject to the constraints of stresses and displacements can be formulated in terms of nondimensional design variables. Therefore, design variables are $x_1 = 0.001A_1E/p$ and $x_2 = 0.001A_2E/p$, in which $E$ is elastic modulus and $p$ the load. Allowable stresses in tension and compression are assumed to be $7.73\times10^{-4}E$ and $4.833\times10^{-4}E$, respectively, and the vertical tip displacement is constrained to be no greater than be $3\times10^{-3}L$. The problem of the minimum weight design subject to stress and displacement constraints can be formulated in terms of the non-dimensional variables as follows:

$$\text{Min}\quad f(x_1,x_2)=3x_1+\sqrt{3}x_2 \tag{24}$$

$$\text{s.t.}\quad G_1=\frac{18}{x_1}+\frac{6\sqrt{3}}{x_2}-3\leqslant 0 \tag{25}$$

$$G_2=5.73-x_1\leqslant 0 \tag{26}$$

$$G_3=7.17-x_2\leqslant 0 \tag{27}$$

As seen in Table 3, the optimum value of the objective function is 44.760 which is obtained using RSM with circular topology in PSO, and the optimum value of the objective function is 44.601, which is obtained using Kriging with circular topology. The optimum value of Kriging is better than RSM's optimum. As shown in Table 3, the error between the exact objective function value and the estimated value calculated from surrogate models with the circular topology in PSO shows us that the hybrid approach has smaller errors than the approach using RSM or Kriging. However in case of using the random topology in PSO with Kiging, the error between the exact objective function value and the estimated value is smaller than the error of using the hybrid approach. However the optimum value of objective function using the hybrid approach with random topology is smaller than the value using Kriging with random topology. From this result, it can be seen that the hybrid approach is more useful than using RSM, and is competent compared with using Kriging.

The constraint values when the optimization is achieved are shown in Table 4. As shown in Table 4 the constraint ($G_1$) exceeds the limit bound, $G_1 \leqslant 0$, and other constraints ($G_2$, $G_3$) satisfy the limit condition. The violation of the constraint ($G_1$) comes from the approximation model's error and penalty parameter's error. Therefore, when surrogate models and the

**Table 3** Optimization results of four-bar truss.

| Method | | PSO | | | | | | | | Bayesian |
|---|---|---|---|---|---|---|---|---|---|---|
| | | RSM | | Kriging | | RSM & Kriging | | Rand | Cir | |
| | | Rand | Cir | Rand | Cir | Rand | Cir | | | |
| Objective | | 44.776 | 44.760 | 44.608 | 44.601 | 44.577 | 44.719 | 44.602 | 44.602 | 47.567 |
| Design variables | $x_1$ | 9.394 | 9.377 | 9.383 | 9.409 | 9.377 | 9.660 | 9.384 | 9.388 | 10.074 |
| | $x_2$ | 9.478 | 9.503 | 9.527 | 9.352 | 9.405 | 9.011 | 9.395 | 9.387 | 10.015 |
| Exact objective function | | 44.597 | 44.589 | 44.648 | 44.426 | 44.420 | 44.588 | | | |
| Error (%) | | 0.399 | 0.383 | 0.091 | 0.394 | 0.354 | 0.291 | – | – | – |

**Table 4** Constraint values at the optimum.

| Method | | PSO | | | | | | | | Bayesian |
|---|---|---|---|---|---|---|---|---|---|---|
| | | RSM | | Kriging | | RSM & Kriging | | Rand | Cir | |
| | | Rand | Cir | Rand | Cir | Rand | Cir | | | |
| Constraints | $G_1$ | 0.0244 | 0.0239 | 0.0222 | 0.0243 | 0.0245 | 0.0208 | 0.0244 | 0.0245 | −0.1755 |
| | $G_2$ | −3.6637 | −3.6464 | −3.6586 | −3.6686 | −3.5749 | −0.9302 | −3.6537 | −3.6577 | −4.3436 |
| | $G_3$ | −2.3079 | −2.3331 | −2.2371 | −2.1997 | −2.3459 | −0.8413 | −2.2245 | −2.2169 | −2.8447 |
| Exact constraints | $G_1$ | 0.0127 | 0.0133 | 0.0093 | 0.0243 | 0.0247 | 0.0166 | – | – | – |
| | $G_2$ | −3.6637 | −3.6464 | −3.6525 | −3.6789 | −3.6468 | −0.9302 | – | – | – |
| | $G_3$ | −2.3079 | −2.3331 | −2.3566 | −2.1824 | −2.2346 | −0.8413 | – | – | – |

penalty method are used, one of the significant issues is how to reduce errors between approximated values and exact values.

The optimization of the four-bar truss problem is also achieved using Bayesian statistics. The optimum value of the objective function using Bayesian statistics is 47.567 as seen in Table 3. Compared with PSO's optimum using surrogate models, the value of the objective function is bigger by 2.990, which is 6.285% over. As seen in Table 4, the constraint ($G_1$) seems more reliable than PSO's values using surrogate models because it does not violate the limit condition ($G_1 \leqslant 0$). Therefore, according to the safety of constraint, the Bayesian's results give the designer more robust data.

Figs. 9 and 10 show the variation of mean and variance of design variables obtained by Bayesian statistics during optimization process.

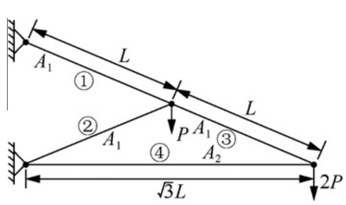


**Fig. 8** Four-bar truss.

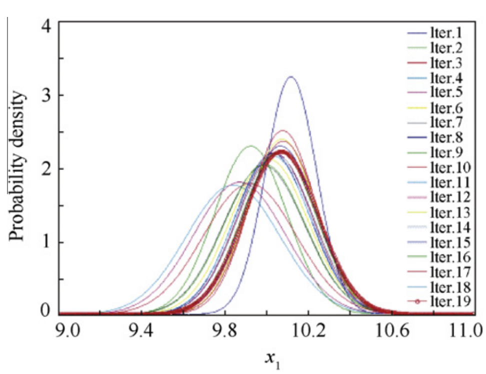


**Fig. 9** History of design variable ($x_1$) of Bayesian statistics.

## 7. Optimization of a hub sleeve of helicopter

In this work, the hub sleeve of a helicopter is optimized considering static failure. The optimization method is the proposed approach using PSO, with the hybrid consisting of RSM and Kriging, and Bayesian statistics. The sleeve attaches the rotor

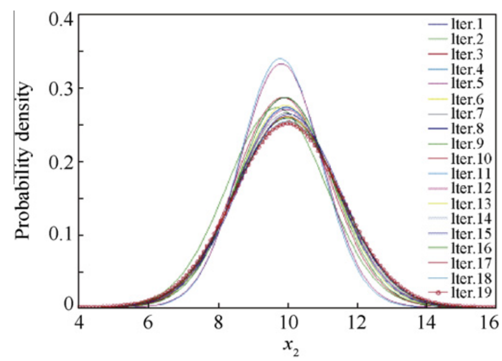


**Fig. 10** History of design variable ($x_2$) of Bayesian statistics.

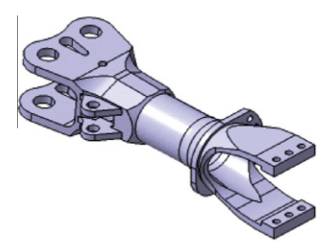

**Fig. 11** Hub sleeve model.

blade to the rotor hub. It also provides attachments for the pitch lever and flapping stops. The flapping stops limit the blade angles. The rotor torque is transmitted to the rotor drive system via the rotor mast (by bushings). The mast is hollow to allow for internal routing of the instrumentation cables. A hub sleeve model is shown in Fig. 11.

### 7.1. Design variables and constraint

The objective is to find the minimum weight of the sleeve subject to a stress constraint. There are six design variables as shown in Fig.12. The optimization formulation is expressed in Eqs. (28)–(33).

$$\text{Min} \quad \text{weight}, \tag{28}$$

$$\text{s.t.} \quad g_1 = \frac{\sigma_V}{1.5(\sigma_V)_u} - 1 \leqslant 0 \tag{29}$$

$$40.0\ \text{mm} \leqslant R_1, R_2, R_3 \leqslant 59.0\ \text{mm} \tag{30}$$

$$85.0\ \text{mm} \leqslant L_1 \leqslant 100.0\ \text{mm} \tag{31}$$

$$70.0\ \text{mm} \leqslant L_2 \leqslant 85.0\ \text{mm} \tag{32}$$

$$50.0\ \text{mm} \leqslant L_1 \leqslant 73.0\ \text{mm} \tag{33}$$

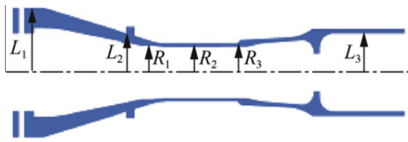


**Fig. 12** Cross section of sleeve.

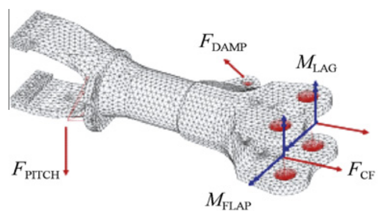


**Fig. 13** Load conditions.

**Table 5** Applied load values.

| Load | Value |
|---|---|
| $F_{CF}$ (N) | 450000 |
| $M_{FLAP}$ (N · mm) | $0.5 \times 10^7$ |
| $M_{LAG}$ (N · mm) | $1.0 \times 10^7$ |
| $F_{PITCH}$ (N) | 30000 |
| $F_{DAMP}$ (N) | 50000 |

where $\sigma_V$ is a constraint for von Mises stress in Eq. (29), and the side constraints are shown in the Eqs. (30)–(33). $\sigma_u$ is 1430 MPa, which is the ultimate strength of sleeve material that is Ti-10V-2Fe-3Al. In this optimization, finite element (FEA) by ANSYS is used to get the weight and the maximum stress. In FEA, 10-node tetrahedral structural solid element is used and the number of nodes is 35212 (DOF = 105636). We use autorotation load condition for sleeve FEA and the load positions are expressed in Fig. 13. Table 5 shows the values of loads applied to this static analysis. Fig. 14 shows the stresses contour calculated by FEM.

### *7.2. Optimum results and discussion*

As shown in Table 6, the objective function (weight) of 15.00 kg in the initial design is reduced to 12.66 kg in the optimum design. And the constraint $g_1 = 0$. In the results, the design variable, $L_1$ is changed the most. From the standard deviations in Table 6, $R_1$ and $R_2$ have greater search regions than others, and $L_1$ has the smallest search region.

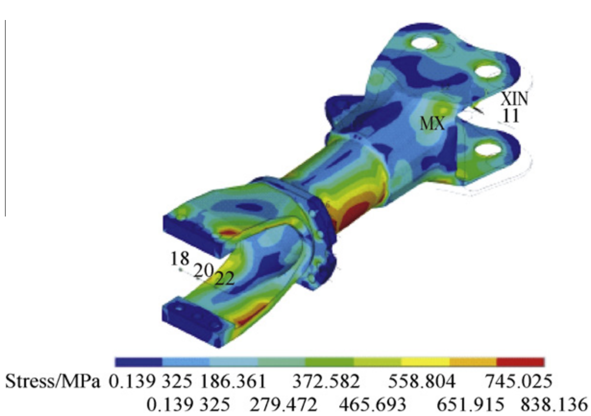


**Fig. 14** FEM result.

**Table 6** Optimum results of hub sleeve.

| Paramater | | Initial | Optimum |
|---|---|---|---|
| Objective function (kg) | | 15.00 | 12.66 |
| Design variables (mm) | $R_1$ | 57.10 | 58.63 |
| | $R_2$ | 55.10 | 58.86 |
| | $R_3$ | 55.30 | 58.50 |
| | $L_1$ | 116.50 | 93.19 |
| | $L_2$ | 73.00 | 84.16 |
| | $L_3$ | 70.30 | 67.03 |

## 8. Conclusions

In stochastic structural optimization using surrogate models, an original function is approximated in order to reduce the numerical calculation cost. In this paper, an optimization procedure using surrogate models with PSO is proposed. In addition, to obtain more reliable optimal results, Bayesian statistics is used. Two test problems are selected from the literature to verify the ability of the proposed method. From the results of the optimization of two test problems, we find that the difference of efficiency between random topology and circular topology is not distinct. It depends on optimization problems. Therefore we should consider using both of the topology types that are random and circular in PSO. Also we find the hybrid using RSM and Kriging is more competent than using only one of the surrogate models such as RSM and Kriging in this problem.

When surrogate models and a penalty method are used in stochastic optimization, there are some errors, which come from the difference between the exact values and the expected values of the surrogate model. The errors make the constraints exceed their limit condition. In case of the violation of the constraint limits, the optimum results obtained from Bayesian statistics suggest reliable optimum values with robust constraint values without a re-optimization process. According to the Bayesian statistic's results, the optimum values are bigger than the original optimum results obtained by PSO. However, the reliability of the optimum values increases. This is caused by the fact that Bayesian statistics uses the entire data generated by PSO during the optimization process. Consequently, Bayesian statistics allows the optimum results to have greater stability and reliability based on the values of the constraints. In other words, Bayesian's results give designers a choice to use more reliable data for their designs without extra optimizations.

## References

1. Goldberg DE. *Genetic algorithms in search, operation, and machine learning*. Boston (MA): Addison-Wesley Publishing Company, Inc.; 1989.
2. Emile A, Jan K. *Simulated annealing and boltzmann machine*. Chichester, UK: John Wiley & Sons; 1997.

3. Eberhart RC, Kennedy J. A new optimizer using particles swarm theory. In: *Sixth international symposium on micro machine and human science*; 1995. p. 39–43.
4. Jie J, Zeng J, Han C, Wang Q. Knowledge-based cooperative particle swarm optimization. *Appl Math Comput* 2008;**205**(2):861–73.
5. Zhao Y, Zu W, Zeng H. A modified particle swarm optimization via particle visual modeling analysis. *Comput Math Appl* 2009;**57**(11–12):2022–9.
6. Myers R, Montgomery DC. *Response surface methodology – process a product optimization using designed experiments*. Canada: John Wiley & Sons; 1995.
7. Martin JD, Simpson TW. Use of Kriging models approximate deterministic computer models. *AIAA J* 2005;**43**(4):853–63.
8. Giunta AA, Watson TL. A comparison of approximation modeling techniques: polynomial versus interpolating models. In: *7th AIAA/USAF/NASA/ISSOMO symposium on multidisciplinary analysis and optimization*; 1998. p. 392–404.
9. Sacks J, Schiller SB, Welch J. Designs for computer experiments. *Technometrics* 1989;**31**:41–7.
10. Rao SS. *Engineering optimization*. New York: Wiley Interscience; 1996.
11. Haftka RT, Gurdal Z. *Element of structural optimization*. Boston (MA): Kluwer Academic Publishers; 1992.
12. Shi YH, Eberhart RC. A modified particle swarm optimizer. In: *Proceedings of the 1998 IEEE international conference on evolutionary computation;* 1998.
13. Maurice C. *Particle swarm optimization*. Newport Beach (CA): ISTE USA; 2006.
14. Simpson TW. A concept exploration method for product family design [dissertation]. Atlanta (GA): Georgia Institute of Technology; 1998.
15. Heckerman D, Geiger D, Chickering M. Learning Bayesian networks: the combination of knowledge and statistical data. Redmont (WA): Microsoft Corporation; 1994. Report No: WS-94-03.
16. Bilal MA, Richard HM. *Probability, statistics, and reliability for engineers and scientists*. Boca Raton (FL): Chapman Hall/CRC; 2003.
17. Ahmadi M, Vahabzadeh F, Bonakdarpour B, Mofarrah E, Mehranian M. Application of the central composite design and response surface methodology to the advanced treatment of olive oil processing wastewater using Fenton's peroxidation. *J Hazard Mater* 2005;**123**(1–3):187–95.
18. Ozgur Y. Penalty function methods for constrained optimization with genetic algorithms. *Math Comput Appl* 2005;**10**(1):45–56.

**Jongbin Im** is a senior researcher at Korea Aerospace Research Institute, Republic of Korea. He received the Ph.D. degree from Korea Aerospace University in 2008. His current research interests are reliability-based design optimization, laser peening process, and multiscale analysis.

**Jungsun Park** is a professor at Department of Aerospace and Mechanical Engineering, Korea Aerospace University, Republic of Korea. He received the Ph.D. degree from the University of Michigan in 1993. His current research interests are multidisciplinary design optimization, fatigue analysis, and radar absorbing structure.