

# Algorithm for joint optimization of machine learning for upgrading a finite difference model

*Natalya Ivakhnenko*<sup>1\*</sup> and *Maxim Badekin*<sup>2</sup>

<sup>1</sup> Russian State Agrarian University – Moscow Timiryazev Agricultural Academy (RSAU – MTA), Department of Physics, 127434, Timiryazevskaya st., 49, Moscow, Russian Federation

<sup>2</sup> Donetsk State University (DonSU), Department of Theoretical Physics and Nanotechnology, 283001 Universitetskaya st. 24, Donetsk, Russian Federation

**Abstract.** Modernization of the finite element model is currently the basic tool for refining the numerical solution of modeling problems by adjusting the numerical response to the observed empirical behavior of the system. Recently, the modification of the model in some cases is carried out using the maximum likelihood method. Following this approach, the update problem can be transformed into a multiobjective optimization problem. Due to the non-trivial non-linear behavior of the desired objective functions, metaheuristic optimization algorithms are usually used to solve such an optimization problem. However, despite the fact that recently this method has proven itself quite well, nevertheless, there are two significant drawbacks that must be eliminated for the practical application of this method. Among which are the length of time spent on the calculation of the problem and the uncertainty associated with choosing the most optimal updated model among all Pareto optimal solutions. To circumvent these limitations, this paper proposes to apply a new joint algorithm that takes advantage of the joint relationship between two optimization algorithms, a machine learning method, and a statistical toolkit. As a result of the study, two main advantages of the newly proposed algorithm were revealed: it leads to a clear reduction in simulation time; and also allows you to make a reliable choice of the best updated model.

## 1 Introduction

Modernization of the model based on the finite difference method in recent years is a widely used method for refining numerical models in various fields of science and technology. Currently, modified finite element models are a popular tool for assessing the performance parameters of various structures [1-3]. Modified finite element models refine the numerical simulation of the structure's response to external factors by making changes to the numerical results in accordance with empirical data. Consequently, they form the basis of most monitoring systems [4].

The method of modified finite difference models as applied to civil engineering appeared in the 90s as an approach with the help of which it became possible to reduce the

---

\* Corresponding author: [yulduz19.77@mail.ru](mailto:yulduz19.77@mail.ru)

discrepancy between model and empirical data [5]. In some cases, modified vibration models were used. [6]. Empirical design properties can be determined using empirical modal analysis or operational modal analysis. [7].

An overview of the whole variety of methods for upgrading a finite difference model is available in the scientific literature, for example, in [8], the following approaches can be used to upgrade a finite difference model of structures: using direct methods, in this case, the matrices characterizing the inertial-strength characteristics of the numerical model are direct modification for accurate reproduction of empirical modal data; iterative methods, in which I discretely change a number of model parameters in order to reduce the discrepancy between numerical and empirical data.

Despite the fact that direct methods were carefully used in early works devoted to the modernization of the finite difference model, they have three significant drawbacks [9]: associated with the matrix of the system, due to inaccuracies in the relationships between elements; the appearance of factors negatively affecting the result; loss of physical meaning in the modified model. Subsequently, methods have been developed and implemented to overcome these shortcomings. To determine the objective function, two different problems are usually considered: the first is the optimization problem according to one criterion and the second is the problem of multi-criteria optimization. The first type of optimization problems is characterized by the presence of one objective function, which is defined in terms of the sum of weighted residuals. Weight values are usually obtained by trial and error. Alternatively, the use of multiobjective optimization makes it possible to avoid the process described above. When solving a single-objective optimization problem, the result is a set of unique parameters. Both the first and the second approaches accept parameters that have a significant impact on its dynamic behavior. The choice of these parameters is usually carried out using sensitivity analysis [10]. Based on the foregoing, it follows that iterative methods make it possible to improve the numerical model. Iterative methods can be divided into three main groups: non-linear programming algorithms [11]; artificial intelligence algorithms [12]; and Bayesian algorithms [13].

In classical works, the problem of upgrading a finite difference model was solved using classical algorithms of nonlinear programming [14-16]. The performance of these algorithms was sufficient to solve the problem of upgrading simple models. However, as the numerical models became more complex, the time required to solve the problem increased, which prevented the use of these classical algorithms. Later, in order to reduce the calculation time, various artificial intelligence algorithms were considered. These methods offer a set of meta-heuristic rules that can guide an algorithm in finding a solution to an optimization problem. From the whole variety of different artificial intelligence algorithms, one can single out the bee algorithm or genetic algorithms. Relatively recently, machine learning methods have also been used to solve the problem of upgrading a finite difference model. Despite this, these computational methods still have two limitations: although they are more efficient than classical non-linear programming methods, they still require a lot of computation time to complete the update process; and in the case of multiobjective optimization, an additional decision problem needs to be solved.

Bayesian algorithms are based on estimating the probability density function of model parameters. These algorithms use Bayes' theorem to estimate the probability density function of model parameters. The likelihood function determines the differences between the numerical and empirical modal properties of the structure. The probability density function characterizes the previous assumption about the statistical behavior of the model parameters. Bayesian algorithms are typically implemented by applying sampling techniques. While there are several advantages to using Bayesian methods to update a finite difference model, the simulation time required to complete the model update is even greater

than for the aforementioned AI algorithms. This fact does not allow applying this approach to solving many practical problems.

In order to reduce the computation time when upgrading the finite difference model, most researchers consider the possibility of using joint or hybrid artificial intelligence algorithms. Algorithms of this type take advantage of two or more AI algorithms such that combining those results in a new overall algorithm that shows improved performance compared to applying them separately. For example, in, the modification of the steel frame structure model is carried out using the annealing simulation algorithm hybridized with the Kalman filter. Similarly, the team of authors in used a particle swarm optimization algorithm to update the model of a continuous railway bridge. The main advantage of the proposals described above is a significant reduction in the duration of the calculation required to solve the modification problem, without losing the accuracy of the modified solution. Taking into account recent trends, this study presents a new joint machine learning and optimization algorithm for modifying a finite difference model. The proposal presented in this article combines two optimization algorithms, namely multicriteria harmony search, active set method in conjunction with ANN [4], and statistical principal component analysis. The basis of the obtained algorithm is the algorithms of multicriteria search for harmony and global optimization, while the latter turned out to be very effective for solving the problems of upgrading the finite difference model. Therefore, a set of solutions is randomly generated and then a multi-objective search algorithm is applied to minimize the multi-objective function, defined in terms of residuals between the numerical and empirical properties of the design, using the most important parameters finite-difference models as design variables. The result of this process is a sparsely populated Pareto front. As a result, the performance of this basic algorithm is enhanced by its interaction with: traditional statistical principal component analysis, which is used to transform the original Pareto front from function space to principal component space, resulting in a modified Pareto front with more pronounced variability. This allows you to get a continuous front, using a discrete number of iterations, which follows from the application of the multi-criteria search algorithm; and a local optimization algorithm, an active set algorithm that uses the properties of the resulting front. This method allows you to choose a solution among all the different optimal solutions of such a front.

The above algorithm satisfactorily handles this decision problem due to its proven efficiency in solving optimization problems. Upon completion of the process of determining the optimal solution in the space of the main components, it can be transformed into the space of functional components, which will allow us to determine a set of parameters that determine the modified model of difference elements. The proposed algorithm has the following main advantages: a significant reduction in computation time compared to the original artificial intelligence algorithm and a more accurate search for the optimal modified model.

## **2 Fundamentals of upgrading the finite difference model by the maximum likelihood method**

The goal of correcting a finite difference model is to find an accurate numerical finite difference model corresponding to the empirical dynamic behavior of the analyzed object. This goal can be achieved by evaluating the most important parameters of the model, which reduce the discrepancies between the numerical and empirical data. Therefore, the problem of upgrading a finite difference model can be reduced to the problem of identifying parameters. In some cases, the maximum likelihood method is equivalent to the least squares method.

This means that the modernization problem can be transformed into an optimization problem, its main goal is to find the minimum of the sum of relative differences between empirical and numerical properties, while the most significant parameters of the model are considered as variables. Thus, the problem of upgrading the finite difference model in the case of a multicriteria approach can be formulated as:

$$\min(f_1(\theta) \ f_2(\theta)) = \min\left(\frac{1}{2} \sum_{j=1}^{m_j} r_j^f(\theta)^2 \ \frac{1}{2} \sum_{j=1}^{m_f} r_j^m(\theta)^2\right) \quad (1)$$

$$\theta_l \leq \theta \leq \theta_u$$

where  $r_j^f(\theta)$  and  $r_j^m(\theta)$  are the residuals of the  $j$ -th natural frequency and mode of oscillation, respectively;  $m_j$  is the number of vibration modes;  $\theta$  is a vector containing the most important parameters of the finite difference model;  $\theta_l$  and  $\theta_u$  are the lower and upper boundaries of the search areas for these parameters, respectively.

The remainders in this paper are defined as follows:

$$r_j^f(\theta) = \frac{f_{num,j}(\theta) - f_{emp,j}}{f_{emp,j}} \quad j = 1, 2, \dots, m_j \quad (2)$$

$$r_j^m(\theta) = \sqrt{\left(\frac{(1 - \sqrt{MAC_j(\theta)})^2}{MAC_j(\theta)}\right)} \quad j = 1, 2, \dots, m_f \quad (3)$$

where  $f_{num,j}(\theta)$  is the numerical natural frequency  $j$ ,  $f_{emp,j}$  is the empirical natural frequency  $j$ , and  $MAC_j(\theta)$  is the value of the modal guarantee criterion [31], which is defined as the ratio used to evaluate the correlation between modal forms: numerical,  $\phi_{num,j}$ , and empirical,  $\phi_{emp,j}$ , vibration mode  $j$ , defined as:

$$MAC_j(\theta) = \frac{(\phi_{num,j}^T(\theta) \cdot \phi_{emp,j})^2}{(\phi_{num,j}^T(\theta) \cdot \phi_{num,j})(\phi_{emp,j}^T \cdot \phi_{emp,j})} \quad (4)$$

Therefore, carrying out the modernization of the finite difference model, as a problem of multicriteria optimization, involves the following steps [12]: empirical determination of the modal properties of the object under study using empirical or operational modal analysis; determination of numerical properties based on the analysis of a finite difference model; objective function evaluation; verification of convergence criteria.

The solution of the multicriteria optimization problem considered above leads to a set of parameter vectors. Each of them is a possible solution to the problem. The set of parameters can be represented graphically as a curve. Thus, each point on this curve represents an optimal upgraded model for which one of the objectives cannot be improved without compromising the other. This means that there is a need to solve the problem of decision making in order to choose the best solution on the front. To do this, various methods have been proposed in the literature based on the balance between the various members of the

front. As practice shows, these methods define the best solution as the point at which a small increase in one term of the objective function will lead to a significant decrease in another term. This point is known as the "knee" point, and each method offers a slightly different criterion for determining it.

### **3 A joint machine learning optimization algorithm for upgrading a finite difference model**

Despite the fact that the approach described above is quite popular for solving problems of upgrading finite-difference models of the objects under study, as mentioned earlier, it has two main limitations in the process of solving practical engineering problems: an intermediate decision-making problem should provide the best optimal solution for the front. To overcome these two shortcomings, the authors of the work proposed a new algorithm that takes advantage of a number of well-known methods of machine learning, statistics and optimization.

#### **3.1. Components of a machine learning collaborative optimization algorithm**

##### *3.1.1. Multipurpose Search for Harmony*

The harmony search algorithm was developed for solving single-criteria optimization problems, and over time it was modernized for a class of multi-criteria optimization problems. This metaheuristic computational algorithm. At the heart of his work lies the imitation of the process of musical inspiration: the search for harmony throughout musical creativity. The algorithm for finding the global minimum value of the objective function uses several heuristics. At the initial stage, the algorithm randomly generates a set, the elements of which are called harmonies. Each harmony contains the values of the model parameters. The harmonies are stored in a matrix called the harmony matrix,  $H$ . An objective function is evaluated for each harmony. In the next step, a new set of harmonies is iteratively generated using the following steps: memory accounting, random creation, and pitch adjustment. Each element of the new harmony vector is determined iteratively based on the information stored in the harmony matrix,  $H$ , otherwise the algorithm generates it randomly. Both mechanisms are controlled by the speed of consideration of the harmony memory, which is, in essence, an estimate of the probability of determining a new element through the previous one. The next parameter, the pitch correction rate, is calculated to determine the probability of changing a new element if it was found using the previous one. The impact of the mutation is based on the initially set bandwidth, which is added to or subtracted from the values of the harmony matrix  $H$ . The objective function is evaluated for each new harmony. To solve the problem of multicriteria optimization, it is necessary to select non-dominated solutions from the entire set of solutions that form the Pareto front. For this purpose, the authors in this paper use the non-dominated sorting method, which is described in detail in. Therefore, for each new solution, the crowding distance attribute is determined: this parameter acts as a benchmark for separating the worst solutions and restoring the original size of the harmony matrix  $H$ . The process described above is repeated until a certain stopping condition is met. As a result of the actions described above, the front is formed. Each of its points represents a possible optimal solution to the problem of upgrading the finite difference model.

### 3.1.2. *Principal component analysis*

Principal component analysis is a tool often used to identify patterns in a set of empirical data. The process of principal component analysis can be divided into the following stages: normalization of the data set; calculation of the covariance matrix of the data; finding eigenvalues and eigenvectors of the covariance matrix; and determining the evaluation of the main components.

There are two facts worth noting in this process: the eigenvector is associated with the largest eigenvalue of the first principal component, and so on; and principal component values are defined as projections of the dataset into principal component space. Therefore, one of the main advantages of Principal Component Analysis is that it allows you to transform the initial data set into a new simplified Principal Component Space. In the transformed coordinate system, the correlation between the variables that define the data set is more pronounced. This method is used in this paper to transform the front from the original function space to the space of principal components. This orthogonal transformation is defined by changing the data so that its maximum variation is progressively projected onto each principal component. In the case under consideration, this transformation makes it possible to emphasize the convex behavior of the front in the new coordinate system. Two main advantages follow from this: increasing the accuracy of any predictive model; and the local optimization algorithm is sufficiently applicable to solve the decision problem, the essence of which is to choose the best solution among the various elements of the found front. These facts indicate that principal component analysis is a key step both for increasing the performance of the proposed algorithm and, as a result, guarantees the successful completion of the process of upgrading the considered finite difference model. Thus, principal component analysis will be applied in the algorithm to the dataset. The new front obtained after applying the multi-objective search for harmony will be referred to as the transformed front. In order to find a continuous front, an artificial neural network will be tuned to the resulting transformed front.

### 3.1.3. *Artificial neural network*

An artificial neural network is a supervised machine learning. Artificial neural networks have not been used in the problems of upgrading the finite difference model and detecting damage to engineering structures.

A fairly widely used network topology is a multilayer perceptron, this method has been successfully used in various areas, for example, non-linear curve fitting. This type of networks consists of two layers: input, output and a set of hidden layers of neurons that interact with each other based on the principle of direct connection, so information flows are unidirectional: one neuron transmits information to all neurons of the next layer, but in return does not receive from them no information. Reasonable and correct setting of the number of hidden layers and neurons in each layer plays a major role in ensuring the accuracy of the network. It has been shown that one hidden layer is sufficient for a uniform approximation of any continuous function . The number of neurons in the hidden layer can be found either by trial and error, where the optimal number of neurons minimizes the mean square error, or using empirical relationships.

The output of neurons is formed using a non-linear transformation of the weighted sum of the inputs. Backpropagation of a multilayer parsetron is a learning algorithm based on adjusting the weight connection between two neurons. Thus, learning can be formulated as a non-linear minimization problem, which is solved using a gradient-based algorithm. As a result of solving the minimization problem, a set of optimal weights was found. The found

weights provide the minimum error between the outputs of the artificial neural network and the real values .

An artificial neural network will be implemented in this joint algorithm for the numerical approximation of the processed front. The front approximated by an artificial neural network will be denoted as an approximated front.

### 3.1.4. Active Set Algorithm

After finding the approximated front, it is necessary to solve the decision problem: finding the best solution among the variety of elements of the approximated front. To achieve the goal, the decision-making problem is again transformed into an optimization problem. By applying the properties of the computed approximated front, using the local optimization algorithm, taking into account its convergence rate and good accuracy in the process of solving the convex optimization problem. The solution of this optimization problem makes it possible to find the "inflection" point, i.e. the point by which the influence of various residuals can be best established. Of the many different local optimization algorithms, this paper uses the traditional active set algorithm, which is associated with its efficiency for solving optimization problems with additional conditions. The active set algorithm is an iterative method that solves the main optimization problem with additional conditions by solving a sequence of subproblems with additional equality solutions. Therefore, the main task of the algorithm is to predict the active set. The active set is defined as a set of constraints that are satisfied when solving the main optimization problem. As a rule, the active set algorithm can be conditionally divided into two phases: feasibility; and optimality. At the feasibility stage, the objective function is ignored, and the allowable point is determined for the constraints. For the optimality phase, the objective function is minimized, while satisfiability is preserved. For the sake of efficiency, the calculation of both phases must be performed by the same basic algorithm.

Thus, the decision problem is usually formulated as the following optimization problem with additional conditions:

must be minimized  $g(\theta) = y(5)$  under the condition

$$\begin{cases} \theta_{\min} \leq \theta \leq \theta_{\max} \\ -x + x_{\min} \leq 0 \\ -y + y_{\min} \leq 0 \end{cases} \quad (6)$$

where  $g(\theta)$  is the objective function,  $\theta$  is the vector of parameters,  $\theta_{\min}$  and  $\theta_{\max}$  are the minimum and maximum values of the physical parameters that form the Pareto front, respectively,  $(x, y)$  are the outputs of the integrated neural network in the spaces of the principal components,  $x_{\min}$  and  $y_{\min}$  are the minimum values of the elements of the front being approximated. The optimization problem is limited by the search area and two non-linear constraints to improve the performance of the update process.

## 3.2. Machine Learning Joint Optimization Algorithm

The considered algorithm for optimizing machine learning includes the main advantages of the four methods discussed above in order to reduce the time spent on upgrading the finite difference model, while ensuring the required accuracy.

For this, the key aspect is the adequate setting of the front. The presence of a sufficiently filled front is indispensable for accurately determining the best modernized model. Finding this front requires a significant number of iterations and a large population size, and the AI algorithm is considered to solve the upgrade problem using a maximum likelihood approach.

As mentioned earlier, the core of the algorithm will be the algorithm of multi-criteria search for harmony.

This algorithm has demonstrated a high rate of convergence and significant accuracy in relation to the problems of upgrading finite difference models. Due to the existence of a directly proportional relationship between the duration of the calculation required for this algorithm and the number of elements of the resulting front, the proposed algorithm will be used to calculate only the sparse front, in order to reduce the duration of the calculation. However, the direct use of a non-overflowing would not provide the required level of accuracy in choosing the best optimal solution. To overcome this problem, the main advantages of both statistical and machine learning methods are incorporated into the algorithm.

Thus, the original uncrowded front will be projected from function space to principal component space by analyzing the principal components of the dataset that make up the front. The calculated front will be denoted as the transformed front. This coordinate transformation is the fundamental point of the proposed algorithm, this is due to the fact that it allows you to express data in a coordinate space in which their variability is clearly traced. The information about the correlation and variation of the data that follows from the analysis of principal components makes it possible to increase the accuracy of any predictive model adapted to the data being processed. Thus, the processed front has a characteristic shape, which facilitates the subsequent solution of the decision-making problem of finding the "inflection" point.

The "inflection" point in the space of principal components coincides with the minimum of the processed front, so that the problem of decision making is reduced to the problem of function minimization. Hence, provided that any local minimum of a convex function defined on a convex set is a global minimum, such a minimization problem can be solved more efficiently with a local optimizer than with metaheuristic algorithms.

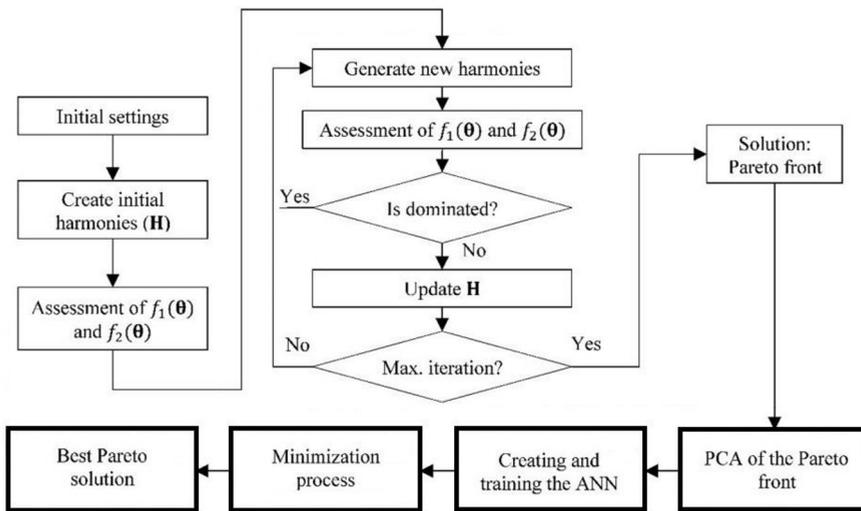
Therefore, to solve a decision-making problem using a local optimization algorithm, it is required that: the transformable front be given a convex function and a convex set is given.

To satisfy the first requirement, the properties of artificial neural networks as universal approximators were taken into account. Thus, with the help of an artificial neural network was designed, trained and tested to approximate the processed Pareto front. This artificial neural network, in fact, is an approximated Pareto front and displays the relationship between the parameters of the finite difference model and the multicriteria function that characterizes the modernization problem. Thus, the approximated Pareto front is a convex function.

Similarly, the second requirement is satisfied by specifying a convex set and the corresponding search area. The minimum and maximum values of the processed front elements are considered to define this area. Since the convex function is given in the same domain, the convex optimization problem is transformed into a convex optimization problem with additional conditions. In the end, the final convex optimization problem with additional conditions can be solved through the use of a local optimization algorithm and, therefore, determining the best upgrade of the finite difference model. Of the set of all local optimization described in the literature, the usual active set algorithm was adopted, since it was widely used to solve constrained minimization problems.

The stages of the algorithm implementation can be described as follows: a set of initial solutions is generated randomly; the multicriteria harmony search algorithm is used to find the initial sparsely populated front; the resulting front is projected onto its space of principal components, in order to find the transformable front; this transformable front is subsequently approximated continually by an artificial neural network; then the minimum value of the resulting approximated front is found; they found best solution is projected back into the original space of functional components, which will provide conditions for finding the parameters that determine the modernized finite difference model.

On fig. 1 shows a block diagram of the proposed machine learning optimization algorithm.



**Fig. 1.** Flowchart of the proposed machine learning-optimization algorithm.

## 4 Conclusions

When solving practical engineering problems, the modernization of a finite-difference model of complex engineering structures is usually carried out by applying the maximum likelihood principle. Guided by this method, the problem of upgrading a finite difference model can be transformed into a problem of multiobjective optimization. Previously, various computational algorithms were proposed to solve a complex optimization problem. However, in the process of applying these algorithms, the following complications arise: it is a long computation time until the upgrade process is completed; and the need to solve a decision-making problem, the accuracy of which depends on setting a crowded front under certain conditions. To overcome these limitations, a new algorithm has been developed and tested in this paper. This combines the advantages of two methods of search and analysis of principal components. The scheme of the joint algorithm can be summarized with the following six steps: a set of initial solutions is randomly generated; a multi-criteria search algorithm is applied to form the initial non-overflowing front; the formed front is projected into the space of its principal components; then the front is continuously approximated by an artificial neural network; the search for the minimum of the approximate front obtained using the neural network is carried out, and in the end, the optimal solution found is transformed back into the original space of functional components. As a result of the steps listed above, the authors of the work receive a set of parameters characterizing the modernized finite difference model.

## References

1. A.E. Aktan, D.N. Farhey, D.L. Brown, V. Dalal, A.J. Helmicki, V.J. Hunt, S.J. Shelley Technical paper **2,3** (1996)
2. V. Compán, P. Pachón, M. Cámara, Paulo B. Lourenço, A. Sáez Engineering Structures **140** (2017)
3. S. Fang, R. Perera, G. Roeck Journal of Sound and Vibration **313**,3–5 (2008)
4. B.N. Datta, D. Sarkission Contemporary Mathematics **280** (2001)
5. J. Carvalho, B.N. Datta, W. Lin, C. Wang Journal of Sound and Vibration **290** (2006)
6. W. D’Ambrogio, A. Fregolent Journal of Sound and Vibration **236** (2000)
7. B. Titurus, M.I. Friswell, International Journal for Numerical Methods in Engineering **75**, 4 (2008)
8. B. Weber, P. Paultre, J. Proulx, Mechanical Systems and Signal Processing **23**, 6 (2009)
9. H. Shahverdi, C. Mares, W. Wang, J.E. Mottershead, Shock and Vibration **16**, 1 (2009)
10. Z.R. Lu, S.S. Law Journal of Sound and Vibration **303** (2007)
11. M. Link, M. Weiland Mechanical Systems and Signal Processing **23** (2009)
12. Kabe, A., AIAA Journal **23**,9 (1985)
13. J. Carvalho, B.N. Datta, W. Lin, C. Wang Journal of Sound and Vibration **290** (2006).
14. S. Moradi, L. Fatahi, P. Razi Struct Multidiscip Optim. **42**,2 (2010)
15. H. Shahverdi, C. Mares, W. Wang, J.E. Shock and Vibration **16**,1 (2009)
16. M. Link, M. Weiland Mechanical Systems and Signal Processing **23** (2009)