

# Reliability analysis of response surface based damage identification method

Tanmoy Mukhopadhyay, Rajib Chowdhury, Anupam Chakrabarti

**Abstract**— Non-destructive structural damage identification (SDI) and quantification of damage is an important issue for Civil, Mechanical and Aerospace engineering structures. Recently, response surface based damage identification methods have been successfully applied for this purpose. In this paper, the reliability of response surface based damage identification methodology has been studied.

**Index Terms**—  $2^k$  factorial design, D-optimal design, Genetic algorithm, Non-destructive structural damage identification, reliability, response surface methodology, sensitivity analysis.

## 1 INTRODUCTION

**D**AMAGE in structures is defined as the changes to the material and/or geometric properties of the systems, including changes to the boundary conditions and system connectivity, which adversely affect the system's performance [1]. Numerous methods have been developed so far for damage detection in structures [2]. Recently one methodology called Response surface Methodology (RSM) has been successfully applied for damage identification. SDI using RSM involves formation of response surface equation and inverse optimization to achieve some target value of the responses with the help of different statistical and mathematical techniques. This method can measure the extent of damage very explicitly along with its location. Though RSM-based damage identification techniques provide comparatively lesser accuracy, but it is very much computationally efficient and cost effective. Damage detection using RSM is a four step procedure: step-1: Identification of structure of interest, step-2: Identification of proper input and output features, step-3: Formation of response surface relating input and output features, step-4: Identification of damage.

RSM was primarily proposed by Box and Wilson [3] for application in chemical industry. After that the methodology has been modified and enriched rigorously for achieving different objectives. Comprehensive description about RSM can be found in [4]. Cundy [5] gave a preliminary idea of using RSM in damage detection of structures. Cho [6] performed an investigation using RSM to predict the accumulated damages in concrete structures. Fang and Perera [7] established a comprehensive methodology for damage identification using RSM. In this paper, a reliability analysis of such damage detection methods have been carried out. The paper is organized as, section-1: introduction, section-2: brief overview of RSM, section-3: different steps of SDI based on RSM, section-4: numerical example of damage identification using RSM, section-5: reliability of RSM based damage identification technique, section-6: conclusion.

## 2. RESPONSE SURFACE METHODOLOGY

On the basis of statistical and mathematical analysis RSM gives an approximate equation which relates the input features  $\zeta$  and output features  $y$  for a particular system.

$$y = f(\zeta_1, \zeta_2, \dots, \zeta_k) + \varepsilon \quad (1)$$

where  $f$  denotes the approximate response function and  $\varepsilon$  is the statistical error term having a normal distribution with mean zero.  $k$  is the number of input parameters. The  $\zeta$  are usually coded as dimensionless variables having mean zero and the same standard deviation of  $\zeta$ .

The metamodel is fit approximately to a set of points in the design space (which may be chosen using design of experiment approach) using a multiple regression fitting scheme. Design of experiments (DOE) is an efficient procedure for planning experiments so that the data obtained can be utilized to achieve any particular goal. In the present study two different DOE methods have been used for the purpose of sensitivity analysis and response surface formation. These two DOE methods are described below.

### 2.1. $2^k$ factorial design

One of the most common first order designs is  $2^k$  factorial design which is very useful for screening out some non-significant input parameters by determining the contribution of each parameter to the total model variance. In this design every input parameter has two coded levels ( $\pm 1$ ), that corresponds to the lower and upper value bound of the design space.

In this design, the number of experimental runs is equal to  $2^k$  provided no single design point is replicated more than once. If  $k$  is large ( $k \geq 5$ ), the  $2^k$  design requires a large number of design points. In that case, we can consider a one-half fraction design consisting of one-half the number of points of a  $2^k$  design, or a one-fourth fraction design consisting of one-fourth the number of points of a  $2^k$  design. In general, a  $2^{-m}$ th fraction of a  $2^k$  design consists of  $2^{k-m}$  points from a full  $2^k$  design.  $m$  should be chosen in such a way that  $2^{k-m} \geq$  number of unknowns in the response surface equation. Sometimes a few additional centre point samples (level=0) are added to the design to evaluate the curvature of the middle region of the design space [4].

- Tanmoy Mukhopadhyay is currently pursuing masters degree program in Structural Engineering, IIT Roorkee, India, PH-07417529653. E-mail: [tanmoy.m.89@gmail.com](mailto:tanmoy.m.89@gmail.com)
- Rajib Chowdhury is currently assistant professor in IIT Roorkee, India
- Anupam Chakrabarti is currently associate professor in IIT Roorkee, India

## 2.2. Optimal design

Optimal designs require fewer samples than the other standard design procedures and thus it is much more computationally feasible mainly in case of large number of input factors. In this design, position of design points is chosen algorithmically according to the number of factors and the desired model and the points are not at any specific positions, they are selected to meet the optimality criteria. Optimal designs can be used to create a good design for fitting a linear, quadratic, cubic or higher order models.

There are several types of optimality criteria such as D-optimality, A-optimality and E-optimality. D-optimality is achieved if the determinant of  $(X^tX)^{-1}$  is minimal. A-optimality is achieved by minimising the trace of  $(X^tX)^{-1}$ . E-optimality is achieved if the largest eigenvalue of  $(X^tX)^{-1}$  is minimal. Here, X denotes the design matrix as a set of value combinations of coded parameters and  $X^t$  is the transpose of X [4],[8],[9]. Among these, D-optimal design is the most popular one. In this study, for model construction an over-determined D-optimal design ( $n_a=k$  additional samples along with the minimum point design and  $n_l=5$  samples to estimate the lack of fit) has been used [10].

## 3. DAMAGE IDENTIFICATION BASED ON RSM

A RSM based structural damage identification procedure consisting of the aforementioned four steps has been discussed in this section.

### 3.1. Identification of structure of interest

First step of SDI is to identify the structure whose damage is needed to be identified. In this study a simply supported beam has been taken.

### 3.2. Identification of proper input and output features

Material properties, such as Young's modulus, density, Poisson's ratio and geometric properties such as section inertia may be taken as input parameters depending on the type of structure under consideration. Time domain features (peak acceleration, temporal moments, logarithmic decrement etc.) and frequency domain features (such as, modal frequencies, mode shapes etc.) are generally taken as output. For highly nonlinear structures, time domain features are more suitable than frequency domain features. Selected output features should not be highly correlated to each other. In Design Expert software, the percentage contribution of each input parameter (including the contribution of the interaction terms) to the total model variance can be obtained. The percentage contribution of each term is obtained by summing all the term sum of squares (SS) and then taking each individual SS and dividing it by the total SS and multiplying by 100 [11].

## 3.3. Formation of response surface relating input and output features

In this step the models have been formed for responses in terms of input parameters using D-optimal design. It should be mentioned that, the models have been constructed using numerical data instead of actual experiments in this study. ABAQUS and Design-Expert softwares have been employed for finite element analysis and response surface model construction respectively [11],[12].

An optimized response surface model is formed by adding or deleting input factors through backward elimination, forward addition or stepwise elimination/addition. It involves the calculation of the *P*-value (probability value, gives the risk of falsely rejecting a given hypothesis) and *Prob. > F* value (gives the proportion of time one would expect to get the stated *F*-value if no factor effects are significant).

The response surface model constructed should be checked by some criterias such as  $R^2$  (A measure of the amount of variation around the mean explained by the model),  $R^2_{adj}$  (A measure of the amount of variation around the mean explained by the model, adjusted for the number of terms in the model. The adjusted R-squared decreases as the number of terms in the model increases if those additional terms don't add value to the model) and  $R^2_{pred}$  (A measure of the prediction capability of the response surface model). The values of  $R^2$ ,  $R^2_{adj}$  and  $R^2_{pred}$  should be close to 1. A difference between  $R^2_{adj}$  and  $R^2_{pred}$  within 0.2 indicates that the model can be used for further prediction. Another check is Adequate precision, which compares the range of the predicted values at the design points to the average prediction error. A value greater than 4 indicates adequate model.

Further, some plots should also be checked such as normal plot of residuals (indicates whether the residuals follow a normal distribution, in which case the points will follow a straight line), residuals vs. predicted plot (plot of the residuals versus the ascending predicted response values), actual vs. predicted plot (A graph of the actual response values versus the predicted response values for the design points used for response surface formation. It helps to detect a value, or group of values, that are not easily predicted by the model), Box-cox plot (helps to determine the most appropriate power transformation to be applied to response data) etc [11].

If experimental data is available for a particular structure, the numerical model can be further modified using model updating technique to get predictions much closer to the actual structure [7],[13].

### 3.4. Identification of damage

Damage identification using the obtained response surface models is an inverse optimization problem i.e. knowing the measured output features, finding out the input parameters that lead to such output values. A hybrid multiobjective genetic algorithm has been used for optimization with the help of Matlab [14]. A hybrid function

(*fgoalattain* has been used in this problem as hybrid function) is an optimization function that runs after the genetic algorithm terminates in order to improve the value of the fitness function. The hybrid function uses the final point from the genetic algorithm as its initial point.

## 4. NUMERICAL EXAMPLE OF DAMAGE IDENTIFICATION

### 4.1. Identification of structure of interest

A 3m long simply supported beam having cross-section of 0.25m x 0.25m (as shown in figure 2) has been taken first for SDI. Material properties of the beam are: Young's modulus ( $E$ ) = 30 GPa, Density ( $D$ ) = 2400 kg/m<sup>3</sup>, Poisson ratio ( $P$ ) = 0.2. The beam is divided into 20 identical parts (as shown in figure 1) for the damage detection purpose.

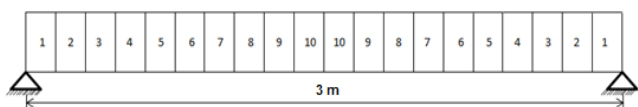


Figure 1. Dimensions of the simply supported beam

### 4.2. Identification of proper input and output features

Since the beam is having uniform cross section and material property along its length, four parameters, Young's modulus ( $E$ ), density ( $D$ ), Poisson ratio ( $P$ ) and section inertia ( $I$ ) of substructure number-4 are taken as input parameters for screening. For screening purpose, a  $2^k$  factorial design is adopted having 16 samples. The first four bending frequencies are taken as responses (output feature) in this case. The bounds ( $\pm 1$ ) of each parameter are identically set to be  $\pm 30\%$  change with respect to the initial values.

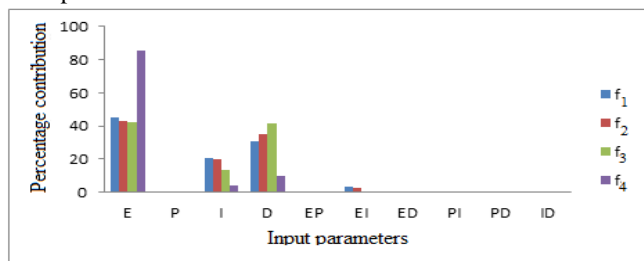


Figure 2. Parameter screening results

The percentage contribution to total model variance of each input parameter (including the two factor interaction effects) to the output features has been shown in Figure 2. From the figure it is evident that chosen output features are highly sensitive to  $E$ ,  $I$  and  $D$ , whereas Poisson ratio has almost no effect on output features. However, in most of the real applications, material property (Young's modulus) and mass (density) remain unaltered. Therefore, in the present study, the beam is assumed to be damaged only due to reduction of section inertia ( $I$ ) i.e. damage has been modelled by reducing the stiffness locally [15]. Furthermore, as modal frequency is a global quantity, damage at two symmetric locations of the symmetric beam will cause same amount of frequency change. Therefore, instead of taking 20 section inertia values, 10 section inertia values ( $I_1, I_2, \dots, I_{10}$ ) have been taken as input parameter. Here,  $I_1$  denotes the section inertia of substructure-1 consisting of two symmetric parts of the beam part-1 and part-20. Similarly,  $I_2$  denotes the section inertia of substructure-2 consisting of part-2 and part-18 and so on.

### 4.3. Formation of response surface relating input and output features

An over determined D-optimal design considering a linear model (with no interaction terms) having total 26 samples consisting of 21 model points plus 5 points to estimate lack of fit has been used. The lower and upper bounds ( $\pm 1$ ) are set to  $0.7I_0$  and  $I_0$ , where  $I_0$  represents the undamaged section inertia. In the Design Expert software, the search option used is 'Best' which tries both Point exchange and Coordinate exchange searches of the design space.

### 4.4 Identification of damage

In this section, the capability to identify damage by using the response surfaces formed by D-optimal design has been discussed. For this purpose, damage have been introduced to the structure by 30% reduction of the section inertia in sub-structure 6. The responses (first four bending frequencies) corresponding to the induced damage condition are found out first.

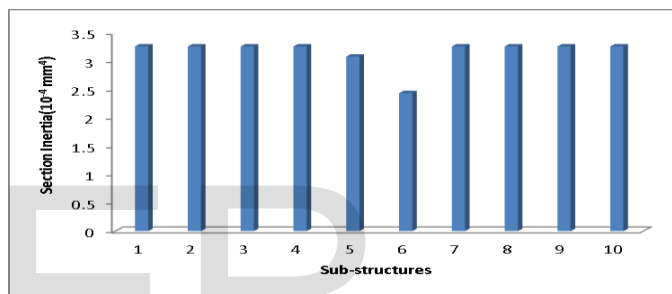


Figure 3. Damage identification result

Then, to judge how the method works for damage detection, the response surfaces formed are optimized by using multi-objective Genetic algorithm to find the value of the input parameters (section inertias of ten substructures), which can cause such responses. In Figure 3 optimization result is shown. Undamaged section inertia of the beam is  $3.2552 \times 10^{-4} \text{ m}^4$ . Figure 3 shows that damage has been correctly identified using D-optimal design.

## 5. RELIABILITY OF DAMAGE IDENTIFICATION

The estimation of the time-invariant reliability of a system or component entails the computation of multidimensional probability integrals [16], [17]

$$P_F = P(g(x) < 0) = \int_{g(x) \leq 0} p_X(x) dx \quad (2)$$

where  $x = \{x_1, x_2, x_3, \dots, x_n\}$ , represent the N-dimensional random variables of the model under consideration;  $g(x)$  is the limit state/performance function, such that  $g(x) < 0$ , represents the failure domain; and  $P_X(x)$  is the joint probability density function of the input random variables.

In the present study sources of variation are considered due to modeling error (Error due to the inaccurate modeling of the structure in finite element analysis softwares, variation in material properties, geometric configurations etc.) and the effect of noise (to simulate the actual field condition). To introduce the modeling error, four natural frequencies corresponding to each set of samples have been varied by generating some random numbers (*randn*) in the range of 0 to 1 with the help of Matlab as shown below.

$$f = f_{original} \times (1 + p_1 \times randn) \tag{3}$$

where,  $f_{original}$  and  $f$  represent the set of natural frequencies obtained from finite element software and the randomly varied natural frequencies respectively.  $p_1$  is the percentage of variation (ranging from 0 to 5 percent). The response surfaces are now formed by using the randomly varied set of natural frequencies. The effect of external noise is introduced to the set of natural frequencies of the actual structure (whose damage is to be identified) in the similar manner. In this case also, the range of percentage of variation (say  $p_2$ ) is taken as 0 to 5 percent.

For the purpose of reliability analysis, the percentage of variations  $p_1$  and  $p_2$  are randomly varied and percentage error in damage detection ( $E$ ) is computed corresponding to each set of  $p_1$  and  $p_2$ . In this study, the reliability analysis has been done by forming response surface equation for percentage error in damage detection in terms of  $p_1$  and  $p_2$ . The response surface have been constructed by using D-optimal design.

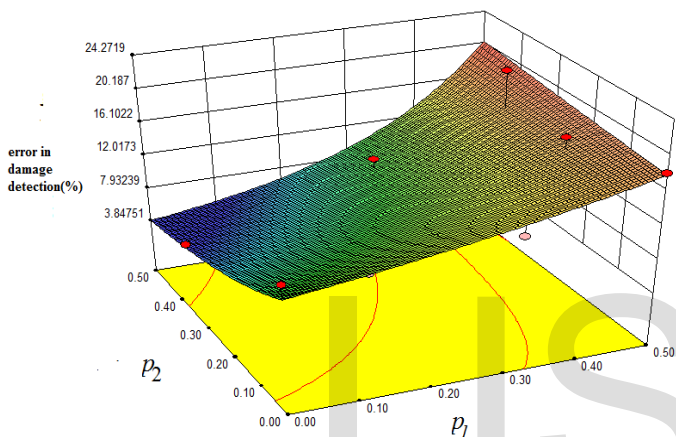


Figure 4. 3D response surface plot for percentage error in damage detection

Response ( $E$ ) is obtained by carrying out Monte Carlo simulation ( $10^5$  runs) and varying  $p_1$  and  $p_2$ . More than 5 percent deviation in the damage detection result (section inertia value) is regarded as the failure of the method to identify damage. The failure probability is finally calculated as

$$P_F = \frac{1}{N_S} \sum_{i=1}^{N_S} I[g(x^i) < 0] \tag{4}$$

where  $x^i$  is  $i^{th}$  realization of  $X$ ,  $N_S$  is the sampling size,  $I[.]$  is a deciding function of the fail or the safe state such that  $I = 1$ , if  $g(x^i) < 0$  otherwise zero.

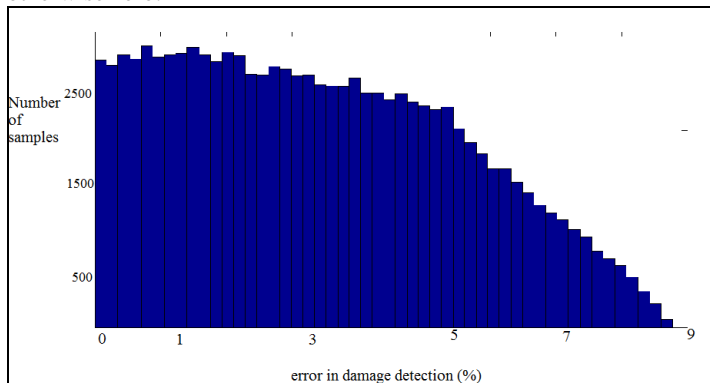


Figure 5. Histogram of error in damage detection vs. number of sam-

ples in Monte Carlo simulation

Figure 4 shows the variation of  $E$  with  $p_1$  and  $p_2$  in a 3D plot. Figure 5 shows a histogram describing the number of samples satisfying the limit state function corresponding to a particular threshold value of  $E$ . For the present study, the failure probability i.e. probability of false damage detection, comes out to be 0.26.

### 6. CONCLUSION

In this paper the damage identification method using RSM has been described in details for a simply supported beam. Then the reliability of such damage identification process based on meta-modelling approach has been explored. This methodology for reliability assessment of damage detection techniques can be extended to more complex structures. Before carrying out the actual damage detection process in any structure, this type of reliability assessment is strongly recommended for judging its probability to successfully detect the damage in that particular structure.

### REFERENCES

- [1] Farrar C R and Worden K “An introduction to structural health monitoring”, 2007, Phil. Trans. R. Soc. A 365, 303–315
- [2] Doebling S W, Farrar C R, Prime M B and Shevitz D W “ Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: A Literature Review”, 1996
- [3] Box G E P and Wilson K B “ On the experimental attainment of optimum conditions”, 1951 J. R. Stat. Soc. Ser. B 13 1–45
- [4] Myers R H and Montgomery D C “ Response Surface Methodology: Process and Product Optimization Using Designed Experiments”, 2002, 2nd edn (New York: Wiley)
- [5] Cundy A L , “Use of response surface metamodels in damage identification of dynamic structures”, Master Thesis Virginia Polytechnic Institute and State University, 2002
- [6] Cho T “ Prediction of cyclic freeze–thaw damage in concrete structures based on response surface method”, 2007 Constr. Build. Mater. 21 2031–40
- [7] Fang S E and Perera R “ A response surface based damage identification technique”, 2009 Smart Mater. Struct. 18(2009)065009
- [8] Michael J B and Norman R D “On minimum-point second-order designs”, 1974 Technometrics 16 (4) 613–616.
- [9] Unal R, Lepscht R A, McMillin M L “ Response surface model building and multidisciplinary optimization using D-optimal designs”, 1998, in: Annual AIAA/ NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, seventh, St. Louis, MO, USA pp. 10–31.
- [10] Carpenter W C “ Effect of design selection on response surface performance” 1993, NASA Contractor Report 4520
- [11] Design-Expert® Version 8.0.7.1, Stat-Ease Inc ;2011
- [12] ABAQUS CAE 6.8. Dassault Systèmes Simulia Corp.; 2008
- [13] Friswell M I and Mottershead J E “ Finite Element Model Updating in Structural Dynamics”, 1995 (Dordrecht: Kluwer Academic)
- [14] Matlab Version 7.12.0.635 (R2011a), MathWorks Inc ;2011
- [15] Friswell M I and Penny J E T “ Crack modeling for structural health monitoring”, 2002, Sage Publications Vol 1(2): 0139–148
- [16] Ditlevsen O, Madsen HO. “Structural Reliability Methods” .Chichester, Wiley; 1996.
- [17] Chowdhury R, Rao B.N. “Assessment of high dimensional model representation techniques for reliability analysis”, 2009, Probabilistic Engineering Mechanics 24 100–115