Genetic Algorithm-based Approach for Bayesian Damage Identification Using Spectral Density Analysis in Beam-like Structures

M. Varmazyar[1] , N. Haritos[2] and E. Gad[3]

1. Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
   Email: maryamv@pgrad.unimelb.edu.au

2. Principal Fellow, Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
   Email: nharitos@unimelb.edu.au

3. Professor, Faculty of Engineering and Industrial Sciences, Swinburne University of Technology, Hawthorn, VIC 3122, Australia
   Email: egad@swin.edu.au

ABSTRACT

This paper, which is a part of an ongoing research study, describes a Genetic Algorithm-based approach towards a global damage identification framework for the continuous/periodic monitoring of civil structures. In order to localise and estimate the severity of damage regions, a one stage model-based Bayesian probabilistic damage detection approach is proposed. The method is based on response power spectral density of the structure which enjoys the advantage of broadband frequency information and can be used for input-output as well as output-only damage identification studies. The suitability of the proposed method is investigated for an error-free numerical model with noise-free response data sets in a beam-like structure with different severity of damage. The results obtained indicate that the Genetic Algorithm-based damage identification approach is suitable for damage detection and can be considered for further implementation using more realistic noisy response data that would be associated with the monitoring of real civil structures.

Keywords: Damage identification, Genetic Algorithm, damage severity, mean square value, power spectral density, Bayesian probabilistic approach

1. INTRODUCTION

Structural health monitoring, (SHM), is the process of assembling general information of the current condition of a structure whose aim is to indicate the existence, location, and severity of damage in that structure if damage occurs. Damage might occur in a structure after long-term deterioration under service loads such as fatigue and corrosion or due to extreme incidents such as earthquakes and impact loads. Accordingly, damage will affect the structural physical properties, namely stiffness, mass and damping, which will change structural response behaviour and alter current and future performances. Civil structures such as buildings, bridges, dams, tunnels, etc. are the most important assets of modern society which need to be functional for a very long time under complex conditions. Thus, continuous/periodic monitoring of civil structures is essential to ensure their safety and acceptable performance during their life span and to prevent a catastrophe.

In the last few decades, structural vibration response monitoring based damage identification methods have gained substantial attention, particularly due to their potential applications in the areas of aerospace, civil and mechanical engineering. The theoretical basis for vibration response measurement based damage identification methods is that structural damage causes changes in structural dynamic properties, which in turn causes changes in the global dynamic characteristics of the structure (such as its associated natural frequencies, mode shapes, modal damping, frequency response functions, etc.). Therefore, estimation of the variations in dynamic response characteristics provides useful information regarding existence, location and severity of structural damage without the need for prior knowledge of any damage condition states. However, earlier research reports and simulation studies indicate significant differences in the levels of sensitivity of the modal and frequency response parameters to different types of damage such as local, global, crack or fatigue damage. Different active damage mechanisms tend to display different response characteristics and changes in the modal and frequency response parameters are found to show substantial variations across the modes and across damage condition states. Neither modal nor frequency response parameters are known to be consistent in providing reliable integrity assessment of a structure under investigation.

In previous research (Bayissa and Haritos, 2007), a two-stage vibration-based damage identification technique that can be used effectively for structural damage identification and condition assessment in one dimensional and two-dimensional plate-like structures was proposed. Firstly, damage-sensitive vibration response parameters that utilise the broadband frequency information (as opposed to resonance frequency based traditional counterparts) that have strong physical relationships to structural dynamic properties (Bayissa and Haritos, 2011) were identified using time-domain, frequency domain, spectral-domain and wavelet-domain analysis methods and implemented in non model based damage identification approaches. Some other salient features are: sensitivity to both local and global damage; low sensitivity to noise and modal truncation errors; identification of linear as well as nonlinear damage conditions. Secondly, a model-based damage identification method was presented in order to formulate a Bayesian probabilistic damage identification algorithm as well as to conduct damage severity predictions using optimisation techniques. This method is more flexible in its application and computationally easier to implement for practical problems and can be used for input-output as well as output-only damage identification studies.

In the previous research (Bayissa, 2007), a distinct solution technique for the Bayesian damage identification algorithm had not been proposed. Accordingly, taking regard of the salient features of the proposed method, identification of the best solution technique and implementation of the Bayesian method on real response data from structures under investigation would be considered a significant contribution to the research field of damage identification and the health monitoring of structures.

2. DERIVATION OF BAYESIAN PROBABILISTIC FRAMEWORK FOR STRUCTURAL DAMAGE IDENTIFICATION

In the former study, a Bayesian probabilistic framework was presented to estimate the severity of damage by updating the probability density function which can be used for structural health monitoring (Bayissa, 2007). This section describes the theoretical background of the Bayesian framework for damage identification.

The mean square value (MSV) of the spectrum which has been presented as the most sensitive parameter for damage detection (Bayissa and Haritos, 2005), is described as the overall energy content of the signal that can be computed either from continuous or discrete time domain signals and is identified as follows:

$$\mu_0(f) = \int_{-\infty}^{\infty} S_{yy}(f)df = \sum_{n=0}^{N/2-1} \left| F[f_n] \right|^2 \tag{1}$$

where $F[f_n]$ is the respective Fourier transform of the N point sequence data series and $S_{yy}(f)$ is the power spectral density.

In order to employ the Bayesian framework, the analytical model is parameterised in terms of structural stiffness K as an assembly of element stiffness matrices assuming that damage affects only the stiffness properties of the structure. The overall stiffness matrix $K(\beta)$ in terms of N_β number of elements is given as follows:

$$K(\beta) = \sum_{i=1}^{N_\beta} (1-\beta_i)K_i \tag{2}$$

where K_i is the stiffness matrix for the *ith* element (or substructure) and β_i ($0 \le \beta_i \le 1$) is a set of non-dimensional model parameters that represents the contribution of the *ith* element stiffness to the global stiffness matrix. In the case that no stiffness loss has occurred, the value of β_i is 0 and in situations of damage for elements or substructures, it would be determined to be greater than 0. Therefore, the value of β_i is an indicator of the location as well as the amount of stiffness loss if any damage has taken place. Therefore, in order to employ Bayes' theorem, all the uncertain quantities were represented as probability distributions and then by creating the *posterior* conditional probabilities for the different variables of interest, inferences can then be determined. Accordingly, by multiplying the *prior* distributions and *likelihood* functions, the result of the statistical inverse problem is provided by the posterior probability distribution.

A joint posterior distribution for the set of model parameters conditioned on the observations can be obtained from Bayes' theorem (Gilks et al., 1996) which can be used as a damage indicator and is defined as follows (Bayissa, 2007 and Sohn and Law, 1997):

$$p(\beta|D,M) = \frac{p(D,\beta|M)}{p(D|M)} = \frac{L(D|\beta,M)p(\beta|M)}{\int_B p(D|\beta,M)p(\beta|M)d\beta} \tag{3}$$

where D denotes vectors of measured data sets from the undamaged and damaged structural condition states, $\beta=[\beta_1,....,\beta_{N\beta}]^T$ indicates the non-dimensional model parameters included in the parameter space, $p(D,\beta|M)$ is the joint probability distribution over all random quantities; $p(\beta|\mathrm{M})$ is the prior probability distribution function (PDF) of the initial model parameters β for a structural model class M,

$L(D|\beta,\mathrm{M})$ is the likelihood density, also known as the conditional probability of observing the data D, $p(\beta|D,\mathrm{M})$ is the posterior density or the updated PDF of the unknown parameters after observing the data; $p(D|\mathrm{M})$ is the normalising factor for the posterior PDF. In that situation in which the main sources of uncertainties are from modelling error and measurement noise, the measured response value $D(s)$ after considering measurement noise $\varepsilon_N(s)$, modelling error $\varepsilon_M(\beta)$ and computed response value $D(\beta)$, is defined as follows:

$$D(s) = D(\beta) + \varepsilon_M(\beta) + \varepsilon_N(s) \tag{4}$$

The normal distribution can be used for defining a mathematical explanation for the numerical approximation error and measurement noise, as follows (Bayissa, 2007):

$$\varepsilon_M \sim \mathcal{N}(\overline{\mu_\mu}, \textstyle\sum_\mu^2), \quad \varepsilon_N \sim \mathcal{N}(0, \textstyle\sum_\mu^2) \tag{5}$$

$\overline{\mu_\mu}$ is the conditional expectation; $\sum_\mu^2$ is the positive definite covariance matrix of the approximation error which can be determined using the inverse-gamma prior distribution as: $\sum_\mu^2 \sim I\mathcal{G}(a,b)$, a and b are parameters of the inverse-gamma prior distribution.

For an independent and distributed zero mean Gaussian noise, the likelihood probability functions for the response measurements is the discrepancy between the theoretical parameters computed from the analytical model and those obtained from measured response data which can be defined as follows:

$$L(D|\beta, M) \propto \exp\left\{-\frac{1}{2}[D(s) - D(\beta) - \varepsilon_M]^T \textstyle\sum_\mu^{-2}[D(s) - D(\beta) - \varepsilon_M]\right\} \tag{6}$$

The conditional PDF of the response MSV parameters for a single data set can be expressed as follows:

$$\begin{aligned} L(\mu_n^0|\beta, M) &= \prod_{\delta r=1}^{N_{\delta r}} L(\mu_{\delta r}^0|\beta, M) \\ &= \frac{1}{f_\mu(\sum)} \exp\left\{-\frac{1}{2}\sum_{\delta r=1}^{N_{\delta r}}[\mu_{\delta r}^0(s) - \Gamma\mu_{\delta r}^0(\beta) - \overline{\mu}_{\varepsilon_M}]^T \textstyle\sum_\mu^{-2}[\mu_{\delta r}^0(s) - \Gamma\mu_{\delta r}^0(\beta) - \overline{\mu}_{\varepsilon_M}]\right\} \end{aligned} \tag{7}$$

in which, $L(\mu_n^0|\beta, M)$ is the conditional PDF for the MSV determined from the r^{th} frequency bandwidth, $\mu_{\delta r}^0(s)$ and $\mu_{\delta r}^0(\beta)$ indicate the vectors of the MSV determined from the measured and computed response data, respectively. s indicates the observed data set number, $s = 1, ..., N_s$. δr is the frequency bandwidth including the r^{th} mode, $\delta_r = 1, ..., N_{\delta r}$. $\overline{\mu}_{\varepsilon_M}$ is the expected value of the modelling error; $f_\mu(.)$ is the normalising factor for the conditional PDFs, given by $f_\mu(\sum) = [2\pi]^{N_{\delta r}/2}\left\|\sum_\mu^2\right\|^{1/2}$, Γ is a matrix that transforms the MSVs computed at full model degrees of freedom to the measurement grid points.

The prior PDF is assumed on the model parameters as white noise and the model parameters, β , can be described as uncorrelated Gaussian random variables of equal covariance centred around $\overline{\beta}$, $\beta \sim \mathcal{N}(\overline{\beta}, \sum_\beta^2)$. $\overline{\beta}$ is the best initial estimate of the model parameter distribution before any data is obtained and $\sum_\beta^2$ is the covariance of the prior PDF, which represents the initial level of uncertainty in the analytical model.

Therefore, the prior PDF on the model parameters can be described using a multi-variate Normal distribution, as follows:

$$p(\beta|M) = \frac{1}{f_\beta(\sum_\beta)} \exp\left\{-\frac{1}{2}\sum_{i=1}^{N_\beta}[\beta_i - \overline{\beta}]^T \sum\nolimits_\beta^{-2} [\beta_i - \overline{\beta}]\right\} \tag{8}$$

where $p(\beta|M)$ is the prior PDF; $f_\beta(.)$ is the normalising factor; $\sum_\beta$ expressed the level of confidence in the initial model parameters.

The joint posterior PDF of the model parameters can be computed by substituting the likelihood and prior PDFs given in Equation (7) and (8), respectively, into Equation (3), as follows:

$$p(\beta|D,M) = \overline{\chi}_1 \frac{1}{f_\mu(\sum)} \exp\left\{-\frac{1}{2}\sum_{\delta r=1}^{N_{\delta r}}\left[\mu^0_{\delta r}(s) - \Gamma\mu^0_{\delta r}(\beta) - \overline{\mu}_{\varepsilon_M}\right]^T \sum\nolimits_\mu^{-2} \left[\mu^0_{\delta r}(s) - \Gamma\mu^0_{\delta r}(\beta) - \overline{\mu}_{\varepsilon_M}\right]\right\}$$

$$\times \frac{1}{f_\beta(\sum_\beta)} \exp\left\{-\frac{1}{2}\sum_{i=1}^{N_\beta}[\beta_i - \overline{\beta}]^T \sum\nolimits_\beta^{-2} [\beta_i - \overline{\beta}]\right\} \tag{9}$$

in which, $\overline{\chi}_1$ is the normalising factor for the posterior PDF of the model parameters. Finally, the posterior PDF can be described in the form:

$$p(\beta|D,M) = \tilde{f}(\sum\nolimits_\mu^2, \sum\nolimits_\beta^2)\exp[-Q(\beta)] \tag{10}$$

where $Q(\beta)$ is the objective (or cost-function) and states the final objective of the problem. The objective function for Equation (10) is described as:

$$Q(\beta) = \frac{1}{2}\sum_{s=1}^{N_s}\sum_{\delta r=1}^{N_{\delta r}}\left[\mu^0_{\delta r}(s) - \Gamma\mu^0_{\delta r}(\beta) - \overline{\mu}_{\varepsilon_M}\right]^T \sum\nolimits_\mu^{-2} \left[\mu^0_{\delta r}(s) - \Gamma\mu^0_{\delta r}(\beta) - \overline{\mu}_{\varepsilon_M}\right]$$

$$+ \frac{1}{2}\sum_{i=1}^{N_\beta}[\beta_i - \overline{\beta}]^T \sum\nolimits_\beta^{-2} [\beta_i - \overline{\beta}] \tag{11}$$

In order to determine the most probable values of the model parameters, some kind of optimisation algorithm should be employed. By maximising the posterior PDF, the maximum *a posteriori* estimate of the parameter of interest can be computed, as follows (Bayissa, 2007):

$$\hat{\beta}_i = arg\ max\ \ p(\beta_i|D,M) \tag{12}$$

in which, $\hat{\beta}_i$ is an optimal model parameter that represents all the information required for assessment of structural damage.

3. GLOBAL OPTIMISATION USING THE GENETIC ALGORITHM METHOD

The Genetic Algorithm (GA) which was developed by John Holland (1975) is a stochastic optimisation technique inspired by natural evolution principles that can be used for discrete as well as continuous optimisation problems dealing with a large number of variables (Haupt et al., 2004). This method produces new points in the search space by applying operators to current points and statistically moving toward more optimal regions in the search space. In order to minimise the cost function, by exploring a broad space of values, the fittest individuals which have been applied into the cost function are selected. (Chambers L. 2001, Reeves et al., 2003).

GA starts with a group of chromosomes known as *population* and every *chromosome* implies the solution for the optimisation problem. The number of unknown parameters in optimisation problems are related to the genes where every *gene* is a basic component of the chromosome, for instance, $c_i=[c_{i1},c_{i2},\ldots,c_{im}]$ where c_{i1} is the first gene of the chromosome c_i etc. The initial values of the chromosomes are mostly generated randomly and then *fitness* $f(x)$ of each chromosome c in the population set is evaluated. In order to create the next *generation*, the Darwinian principles of reproduction are employed through the crossover and mutation function. In the process of reproduction, the chromosomes with a better fitness have a higher chance to survive and be considered for a new offspring population.

Crossover is another operator of GA for the reproduction process in which different parts of the two parents' chromosomes involved in the process, are randomly selected from the crossover points to create two new offspring. In order to distract the algorithm from converging on a popular solution and increase the freedom of the algorithm to search in a broad space of solution, a *mutation* operator is employed where a random gene in a chromosome is replaced by one chosen randomly from the solution space. A detailed description of types of selection of different operators in GA is out of the scope of this paper. Further details can be found in Goldberg (1989).

4. STRUCTURAL DAMAGE SEVERITY ESTIMATION USING MODEL-BASED STUDY

GA optimisation technique for damage identification has been trialled on a simply supported reinforced concrete beam. In order to develop the parameterised FEM of the beam, a MATLAB toolbox known as CALFEM has been employed and meshed using two-dimensional beam element with 10 elements in the definition. The material properties for the undamaged beam are: mass density of 2400 kg/m^3, Young's modulus of 30GPa, and Poisson's ratio of 0.25. Localised damage was simulated by percent reduction in the Young's modulus of: 1%, 5%, 10%, 15% and 20%, at the damage locations of interest. The locations of damage for both the single and multiple damage condition are illustrated in Figure 1. A broadband impact hammer excitation simulated using an impulsive load with a single integration time step, has been implemented.

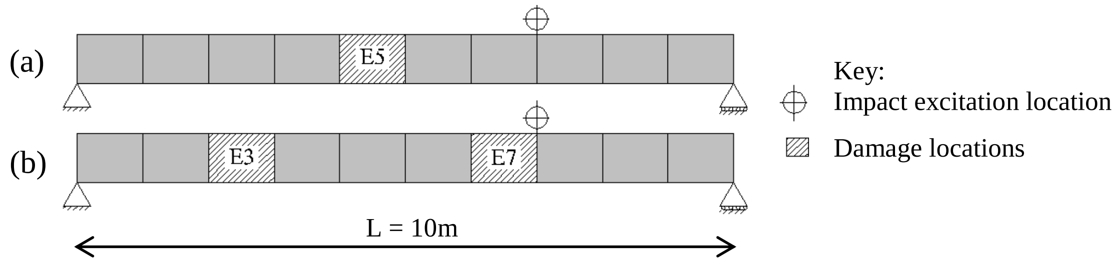

Figure 1 FE model of the beam with damage locations and simulated measurement grid points: (a) single damage condition (at model element 5); (b) multiple damage condition (at model element 3 and 7)

In order to investigate the suitability of the proposed method, damage with different severity for an error-free numerical model with noise-free response data sets was simulated for a "proof of concept" study. A constant modal damping ratio of 0.01 was applied and the first 6 flexural modes determined from both the undamaged and damaged states were implemented for computation of the undamaged and damaged response MSVs (Mean square values) at each nodal point subjected to stationary white noise. The number of model degrees of freedom where MSVs were computed were the same as the simulated measurements points (Figure 1). A single frequency bandwidth, $N_{\delta r}=1$, that included the first 6 flexural modes was used. $\sum_{\mu}^{2}$ is the positive definite covariance matrix of the approximation error which was considered as a constant value as all unknown parameters are totally uncorrelated. Finally, GA optimisation technique was employed to maximize the *a posteriori* (objective function) which would inversely predict the severity of damage that would have been induced for the different damage conditions. To provide information regarding detecting and estimating the degree of damage, the objective function resultant from the Bayesian framework was updated after generating the new points operated by GA. In order to maximise the Bayes' theorem, different values for the unknown parameters (percent stiffness reduction) moving toward the optimal values were generated and evaluated.

A population size of 40 was considered with a randomly selected initial population in the range of 0 to 1 for damage parameters. The stochastic uniform method was chosen for the type of parents' selection for creating the next generation. The heuristic crossover was set to the crossover function with a crossover fraction of 0.5 for the reproduction. Furthermore, uniform mutation was used for the mutation function with a probability rate of 0.1 and in order to scale the raw fitness scores to values in the range, the rank fitness scaling function was employed. The general optimisation process adopted for predicting the location and the estimation of structural damage severity and the results obtained are demonstrated in Table 1 and Figures 2-4. Furthermore, a typical plot of the objective function in terms of the number of generations, depicted in Figure 5, indicates the convergence characteristics of GA.

Table 1 Damage quantification results of 10 runs by Genetic Algorithm optimisation.

No.	Damaged Element No.	Degree of Damage (%)	Generation No.	Runtime (m)
1	5	1	12441	528
2	5	5	10609	446
3	5	10	8229	347
4	5	15	7227	316
5	5	20	6152	269
6	3 , 7	5 , 5	14576	623
7	3 , 7	10 , 10	11613	504
8	3 , 7	15 , 15	7735	345
9	3 , 7	20 , 20	6863	300
10	3 , 7	1 , 15	7856	346

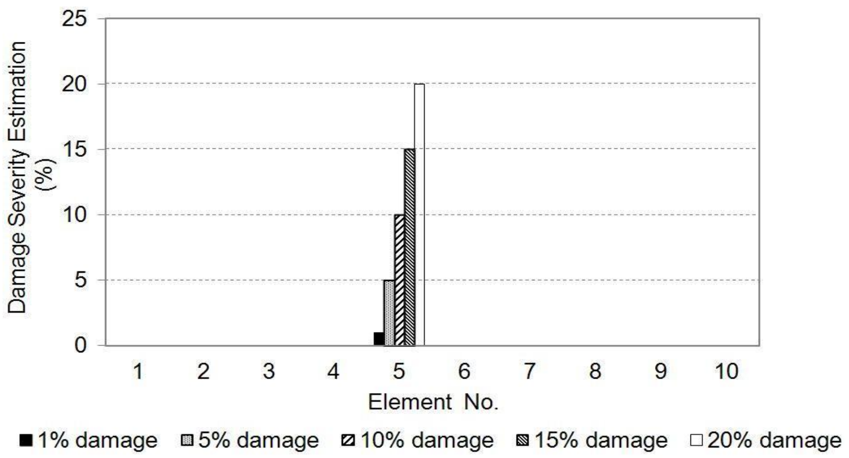

Figure 2 Single damage severity estimation using Bayesian approach of MSV response PSD and GA optimisation technique.

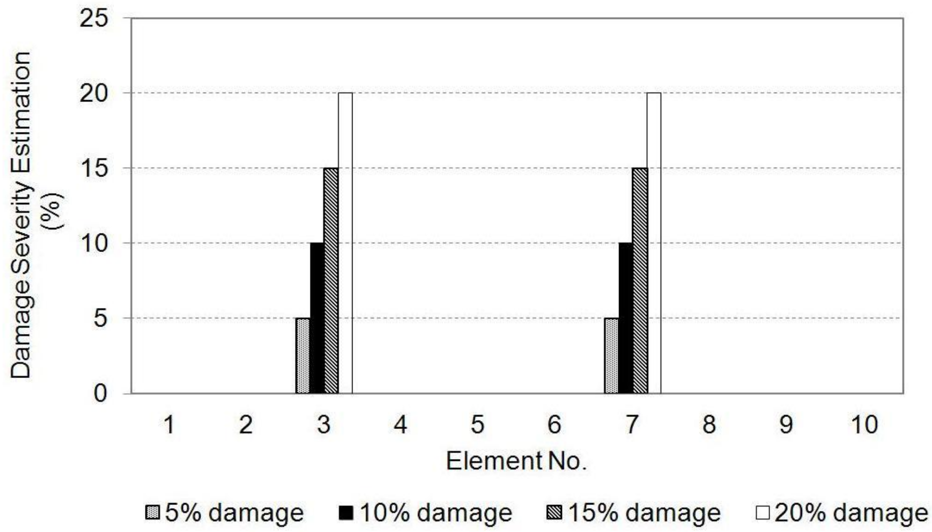

Figure 3 Multiple damage severity estimation using Bayesian approach of MSV response PSD and GA optimisation technique.

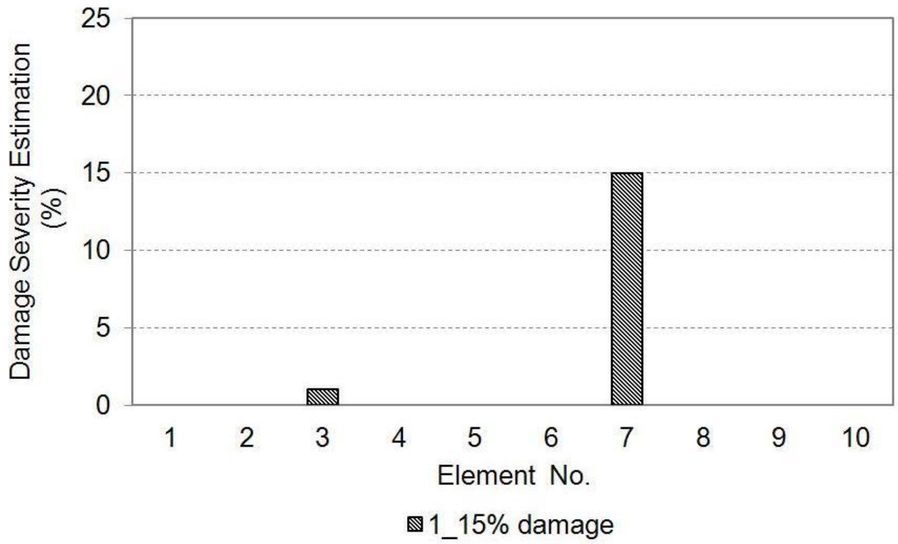

Figure 4 Multiple (irregular) damage severity estimation using Bayesian approach of MSV response PSD and GA optimisation technique.

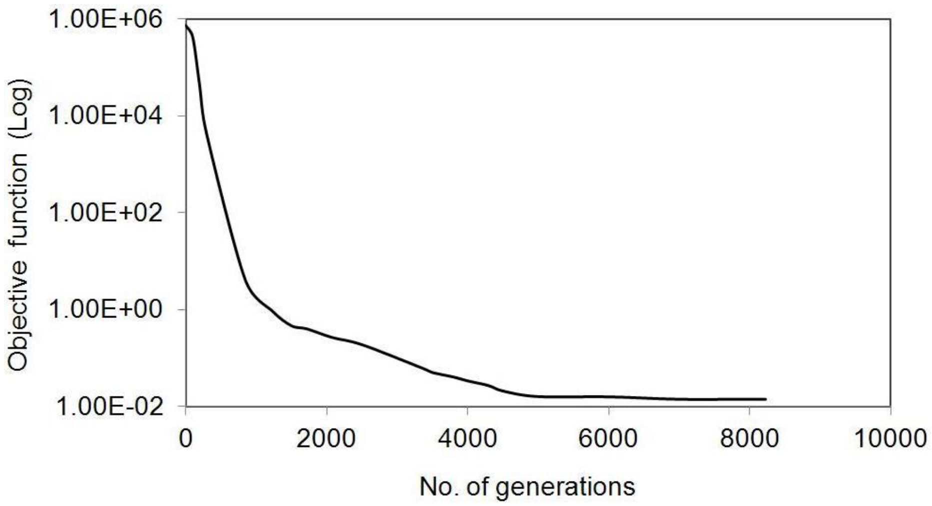

Figure 5 Typical plot of the objective function values versus the number of generations.

5. DISCUSSION AND CONCLUDING REMARKS

A Genetic Algorithm-based approach for Bayesian probabilistic damage identification using power spectral density of the response of the structure has been demonstrated in a "proof of concept" study. The results obtained from GA optimisation, that can be used for detection of the location and degree of damage, are illustrated in Figures 2, 3 and 4. It can be clearly seen from the results that the GA method is able to accurately detect the severity as well as the location of damage through a one-stage model-based damage identification process - even for the 1% damage severity level and from just a single set of response data, for this example study using noise-free response measurement data. The results also indicate that the MSV of the spectrum is quite sensitive to structural damage existence, location and damage severity.

Therefore, the proposed GA technique using the Bayesian framework is considered suitable for further implementation using more realistic noisy response data that would be associated with the continuous/periodic monitoring of real civil structures.

6. REFERENCES

Bayissa, W. L. (2007) Damage identification and condition assessment of civil engineering structures through response measurements, Ph.D Thesis, Department of Civil & Environmental Engineering, The University of Melbourne, August 20, Australia.

Bayissa, W. L. and Haritos, N. (2005) Structural damage identification using spectral moments of the vibration response signal, 1st International Conference on Structural Condition Assessment, Monitoring and Improvement, Dec 12-14, Perth, W. Australia.

Bayissa, W. L. and Haritos, N. (2011) Comparative study of broadband damage localization methods applied to test data, Ocean Engineering, 38, 329–340.

CALFEM Version 3.4 User's Manual, A Finite element toolbox, Department of Structural Mechanics, Lund University, Sweden.

Chambers L. (2001) The Practical Handbook of Genetic Algorithms, Applications, Chapman & Hall/CRC, US.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (1996) Markov Chain Monte Carlo in practice, Chapman & Hall, London, UK.

Goldberg, D. E. (1989) Genetic Algorithms in Search, Optimisation, and Machine Learning, Reading, MA: Addison-Wesley.

Haupt, R. L. and Haupt, S. E. (2004) Practical Genetic Algorithms, John Wiley & Sons, Inc., Hoboken, New Jersey, US.

MATLAB (2002) User's Manual Release 13, The Mathworks, Inc., Natick, MA, USA.
Mattson, S.G. and Pandit, S.M. (2006) Statistical moments of autoregressive model

residuals for damage localisation, Mechanical Systems and Signal Processing, 20, 627-645.

Reeves, C. R. and Rowe, J. E. (2003) Genetic Algorithms-Principles and Perspectives, Kluwer Academic Publishers, London, UK.

Sohn, H. and Law, K. H. (1997) Bayesian probabilistic damage detection of a reinforced-concrete bridge column, Earthquake Engineering & Structural Dynamics, 29**,** 1131-1152.

Technical Committee on Optimal Structural Design (1997) Guide to Structural Optimisation, American society of civil engineers (ASCE), New York, US.