Heung Fai Lam, Ching Tai Ng
The Selection of Pattern Features for Structural Damage Detection Using an Extended Bayesian ANN Algorithm
Engineering Structures, 2008; 30(10):2762-2770

Journal of Engineering Structures (submitted Nov 2006; revised at Feb 2008)

The Selection of Pattern Features for Structural Damage Detection Using an Extended Bayesian ANN Algorithm

Heung Fai LAM[1,*] and Ching Tai NG[2]

[1] Department of Building and Construction, City University of Hong Kong, Kowloon Tong, Kowloon, Hong Kong

[2] School of Engineering, The University of Queensland, Brisbane, Australia.

[*] Send correspondence to Dr HF Lam (paullam@cityu.edu.hk).

Keywords: Artificial neural networks; Structural damage detection; Bayesian model class selection; Benchmark study.

Total no. of pages: 20 (text, including this page) + 4 (figure) + 7 (table) = 31

The Selection of Pattern Features for Structural Damage Detection Using an Extended Bayesian ANN Algorithm

Heung Fai LAM[1] and Ching Tai NG[2]

Abstract

Pattern recognition is a promising approach to the detection of structural damage using measured dynamic data. Many researches of pattern recognition have employed artificial neural networks (ANNs) as a systematic way of matching pattern features. When such methods are used, the ANN design becomes the most fundamental factor affecting the performance and effectiveness of the pattern recognition process. The Bayesian ANN design algorithm proposed in Lam *et al.* (2006) provides a mathematically rigorous way of determining the number of hidden neurons for a single-hidden-layer feedforward ANN. The first objective of this paper is to extend this Bayesian ANN design algorithm to cover the selection of activation (transfer) functions for neurons in the hidden layer. The proposed algorithm is found to be computationally efficient and is suitable for the real-time design of an ANN. As most existing ANN design techniques require the ANN model class to be known before the training process, a technique that can automatically select an "optimal" ANN model class is essential. As modal parameters and Ritz vectors are commonly used pattern features in the literature, the second objective of this paper is to compare the performance of these two pattern features in structural damage detection using pattern recognition. To make a fair judgment between the features, the IASC-ASCE benchmark structure is employed in a case study. The results show that the performance of ANNs trained by modal parameters is slightly better than that of ANNs trained by Ritz vectors.

[1] Assistant Professor, Department of Building and Construction, City University of Hong Kong.
Email: paullam@cityu.edu.hk [corresponding author].
[2] Postgraduate Student, School of Engineering, The University of Queensland, Brisbane, Australia,
Email: c.ng2@uq.edu.au.

1 Introduction

Ensuring the safety of structural systems is an important and challenging mission for engineers and researchers. The collapse of structural systems is usually caused by the continuous accumulation of damage during the service life of the structure. To prevent disasters such as collapse, therefore, remedial works to damaged structures must be carried out as early as possible. Several structural damage detection methods (Lam *et al.* 1998; Beck and Katafygiotis 1998; Papadimitriou 2004a; Papadimitriou 2004b; Ng 2007; Ng and Lam 2006; Lam *et al.* 2007) and sensor technologies have been developed to address this problem (see Sohn *et al.* 2004 for a comprehensive review), many of which are now sufficiently mature for real application in structural health monitoring projects. Indeed, instruments such as laser Doppler vibraometers and shearographs are capable of measuring the static and dynamic responses of structures to a very high degree of accuracy. However, although many methods for structural damage detection have been proposed by various researchers, none is at this stage sufficiently developed for real application. Pattern recognition is one of the most popular approaches to structural damage detection. Artificial neural networks (ANNs) are commonly adopted in pattern recognition to match pattern features, mainly because of their outstanding pattern generalisation capabilities (Wu *et al.* 1992; Elkordy *et al.* 1994; Lam *et al.* 2006). The basic idea of applying ANNs in structural damage detection is to treat the calculated pattern features from the structural model as inputs and the corresponding damage scenarios as targets in the ANN training process. The trained ANN is then able to estimate the damage scenario by fitting the measured pattern features to the inputs.

Wu *et al.* (1992) attempted to use Fourier spectra as pattern features to detect structural damage in a three-storey building model with the help of an ANN, but concluded that Fourier spectra are unsuitable features for damage detection. Elkordy *et al.* (1994) proposed an ANN-based damage detection method that used displacement and strain mode shapes as pattern features, and obtained encouraging numerical and experimental results. Lam *et al.* (2006) proposed the use of damage-induced Ritz vector changes as ANN inputs to identify damage location and severity. Yuen and Lam (2006) proposed a two-stage ANN-based damage detection methodology. In the first stage, the damage signature (Lam *et al.* 1998; Lam 1994) is employed as the pattern feature to identify damage locations, and in the second stage, the severity of the damage is estimated using another ANN that is trained by using changes in the

modal parameters as the pattern features.

The design of the ANN is the critical factor that affects the success of all ANN-based methods. The Bayesian ANN design algorithm proposed in Lam *et al*. (2006) and Yuen and Lam (2006) provides a mathematically rigorous and systematic way of determining the number of hidden neurons in a single-hidden-layer feedforward ANN. The first objective of this paper is to extend this Bayesian ANN design algorithm to cover the design of the activation (transfer) function, which is one of the main elements in the ANN structure. The proposed algorithm is computationally efficient and is suitable for the real-time design of an ANN. This technique should prove extremely useful for most 'two-step' ANN-based structural damage detection methods, in which the location of damage location is identified by a pre-trained ANN in the first step and the extent of the damage is detected by another ANN in the second step, because the number of output nodes in the second ANN depends on the results from the first step, which necessitates a real-time ANN design method.

The selection of a damage-sensitive and noise-insensitive pattern feature is important for all structural damage detection methods. Modal parameters and Ritz vectors (Cao & Zimmerman 1997; Sohn & Law 2001) are commonly used features in the literature, and thus the second objective of this paper is to compare the performance of these two pattern features in structural damage detection using pattern recognition. To make a fair judgment, the IASC-ASCE Phase I Structural Health Monitoring (SHM) benchmark is employed in a case study to demonstrate the proposed method and to compare the performance of the modal parameters and Ritz vectors in structural damage detection. The IASC-ASCE Phase I SHM benchmark study has been comprehensively documented in a special section of the Journal of Engineering Mechanics (Bernal & Beck 2004). In the case study, damage-induced changes in modal parameters and Ritz vectors are separately employed as pattern features to train two ANNs. The proposed Bayesian ANN design method is then employed to identify the "most suitable" class of ANN models so that the ad hoc assumptions and subjective decisions of ANN designers can be avoided. Very encouraging results are obtained. In addition, the damage detection results show that the performance of ANNs trained by modal parameters is slightly better than that of ANNs trained by Ritz vectors.

2 Methodology and Theoretical Background

2.1 Pattern Features

It was reported by Lam *et al.* (2006) and Yuen and Lam (2006) that if modal parameters and Ritz vectors are directly employed in structural damage detection, then the results will be very sensitive to modelling error. However, if the damage-induced changes in the measured quantities are employed, then the influence of modelling error on the damage detection results can be significantly reduced. Thus, the damage-induced changes in the modal parameters and Ritz vectors are separately employed in this paper as pattern features.

Several methods have been developed for calculating Ritz vectors, such as those of Cao and Zimmerman (1997) and Sohn and Law (2001). In this paper, the method proposed by Sohn and Law (2001) is adopted. As the limited space in this paper prevents us from going into detail about the calculation of Ritz vector based on a set of measured natural frequencies and mode shapes, interested readers are referred to (Sohn and Law 2001) for a fuller account.

Two pattern features are employed in this paper: the damage-induced changes in modal parameters (denoted as PF I) and the damage-induced changes in Ritz vectors (denoted as PF II). Based on computer simulation, the PF I for different damage scenarios can be calculated by

$$\Delta\mathbf{P}_1(k)=\left[\boldsymbol{\Phi}_1^T(k),\ldots,\boldsymbol{\Phi}_{N_F}^T(k),\omega_1(k),...,\omega_{N_F}(k)\right]^T-\left[\boldsymbol{\Phi}_1^T(0),\ldots,\boldsymbol{\Phi}_{N_F}^T(0),\omega_1(0),...,\omega_{N_F}(0)\right]^T \quad (1)$$

where $k=1,...,N$ is the index that represents a particular damage scenario and $k=0$ stands for the undamaged scenario (healthy state), N_F is the number of identified modes, and $\omega_q(k)$ and $\boldsymbol{\Phi}_q(k)$ are the natural frequency and mode shape of the *q*-th mode and the *k*-th damage scenario, respectively. The *k*-th damage scenario is represented by the damage index vector $\mathbf{E}(k)$, which is considered in the proposed method to be the ANN target in the proposed method, as follows.

$$\mathbf{E}(k)=\left\{E_1,E_2,\ldots,E_r,\ldots,E_{N_D}\right\}^T \quad (2)$$

where N_D is the total number of possible damage locations and E_r is the damage severity at the *r*-th possible damage location. In the proposed method, the value of E_r varies from 0 to 100 and represents the percentage reduction in stiffness at the corresponding damage location (e.g., $E_3=40$ refers to a 40% reduction in stiffness at the 3rd location). The ANN that uses PF I as the pattern feature is trained by a set of input-target pairs that are calculated from equations

(1) and (2). After training, the ANN can "estimate" the damage index vector $\hat{\mathbf{E}}_1$, thus showing the estimated damage scenarios of the structure, by fitting the measured damage-induced modal parameter changes $\Delta\hat{\mathbf{P}}_1$ to its input. $\Delta\hat{\mathbf{P}}_1$ is defined as

$$\Delta\hat{\mathbf{P}}_1 = \left[{}_{DS}\hat{\boldsymbol{\Phi}}_1^T, \ldots, {}_{DS}\hat{\boldsymbol{\Phi}}_{N_F}^T, {}_{DS}\hat{\omega}_1, \ldots, {}_{DS}\hat{\omega}_{N_F} \right]^T - \left[{}_{US}\hat{\boldsymbol{\Phi}}_1^T, \ldots, {}_{US}\hat{\boldsymbol{\Phi}}_{N_F}^T, {}_{US}\hat{\omega}_1, \ldots, {}_{US}\hat{\omega}_{N_F} \right]^T \quad (3)$$

where the subscripts US and DS stand for the undamaged and damaged states, respectively. The damage-induced Ritz vector change, or PF II, is defined as

$$\Delta\mathbf{P}_2(k) = \left[\mathbf{r}_1^T(k), \ldots, \mathbf{r}_{N_R}^T(k) \right]^T - \left[\mathbf{r}_1^T(0), \ldots, \mathbf{r}_{N_R}^T(0) \right]^T \quad (4)$$

where $\mathbf{r}_i$ is the i-th Ritz vector and N_R is the number of Ritz vectors to be employed in the damage detection process. This ANN, which uses PF II as the pattern feature, is trained by a set of input-target pairs that are calculated from equations (2) and (4). For PF II, the measured Ritz vector changes $\Delta\hat{\mathbf{P}}_2$ can be calculated as follows

$$\Delta\hat{\mathbf{P}}_2 = \left[{}_{DS}\hat{\mathbf{r}}_1^T, \ldots, {}_{DS}\hat{\mathbf{r}}_{N_R}^T \right]^T - \left[{}_{US}\hat{\mathbf{r}}_1^T, \ldots, {}_{US}\hat{\mathbf{r}}_{N_R}^T \right]^T \quad (5)$$

The "estimated" damage index vector $\hat{\mathbf{E}}_2$ can then be obtained as the ANN outputs by fitting the measured Ritz vector changes $\Delta\hat{\mathbf{P}}_2$ to the ANN input.

2.2 Bayesian Artificial Neural Network design algorithm

The basic mechanism of pattern recognition is first to calculate the pattern features of a selected list of possible damage scenarios by computer simulation, and then to match the measured pattern features from the possibly damaged structure with all the calculated pattern features one by one. The damage scenario that corresponds to the "best fit" calculated pattern feature is then considered to be the "true" damage scenario for the structure. Because of its pattern matching ability, ANN is adopted in the proposed method as a systematic tool for matching the measured pattern features to the calculated pattern features. As with all existing ANN-based structural damage detection methods, the success of the proposed method depends heavily on the design of the ANN, which involves the selection of an ANN model class and the training of the ANN (i.e., the selection of the "best" model in the selected model class).

The multi-layer feedforward type of ANN is adopted in the proposed Bayesian ANN design algorithm. This type of ANN is commonly used in ANN-based structural damage detection methods in the literature (Sohn *et al.* 2004), and was proved by Cybenko (1989) to

require only one single hidden layer to be able to approximate any functional relationship between inputs and outputs. However, the ANN does need more hidden neurons to simulate a more complicated functional relationship. Without any loss of generality, the proposed Bayesian ANN design algorithm focuses on the design of a single-hidden-layer feedforward ANN, as shown in Figure 1, in which N_I and N_O are the numbers of neurons in the input and output layers, respectively, and n is the number of neurons in the hidden layer. With this arrangement, the design of the ANN involves the selection of 1) the number of hidden neuron in the hidden layer and 2) the activation function for all of the neurons in the hidden layer (note that a linear function is always employed as the activation function in the output layer).

The number of hidden neurons (n) has a significant effect on the performance of the trained ANN. If n is too small, then the ANN may not be able to capture the behaviour of the training data, which will result in poor performance. It is clear that the larger the value of n, the better the performance (i.e., the smaller the discrepancy between the ANN output and the target). However, if n is too large, then the trained ANN may produce outputs that will fluctuate in the region between the training data points. If performance (the discrepancy between the ANN output and the target) is employed as the only criterion for designing n, then the largest possible value of n will always be selected.

In addition to the number of neurons, the type of activation function adopted in the hidden layer is also an important factor in ANN design, and in particular the nonlinearity of the activation function, because this affects the capability of the ANN to generalise. In the case study, the ability of the proposed ANN design algorithm to design both the number of hidden neurons and the activation function are demonstrated.

A key element of the proposed ANN design method is the Bayesian model class selection method (Beck and Yuen 2004; Lam *et al* 2007) that is employed, which is briefly reviewed as follows. Let D denote the set of input-target training pairs. The objective here is to use D to select the 'best' ANN model class from N_M model classes M_j, for $j = 1, 2, \ldots, N_M$. Let $\boldsymbol{\theta}_j \in \Re^{N_j}$ denote the vector of the ANN parameters, which includes the weights and biases of the ANN in the model class M_j. The ANN training can be treated as a process of selecting an ANN model $\boldsymbol{\theta}_j$ from the model class M_j such that the ANN can simulate the input-target relationship specified by the set of training data D and N_j is the dimension of $\boldsymbol{\theta}_j$ and is equal to the number of ANN parameters in M_j. Following Bayes' theorem, the updated probability

density function (PDF) $p(\boldsymbol{\theta}_j \mid D, M_j)$ based on the set of training data D can be expressed as (Beck & Katafygiotis 1998)

$$p\left(\boldsymbol{\theta}_j \mid D, M_j\right) = c_j\, p\left(\boldsymbol{\theta}_j \mid M_j\right) p\left(D \mid \boldsymbol{\theta}_j, M_j\right) \tag{6}$$

where c_j is a normalising constant and $p(\boldsymbol{\theta}_j \mid M_j)$ is the prior (initial) PDF of the uncertain parameters $\boldsymbol{\theta}_j$, which allows the judgment of the relative plausibility of the values of $\boldsymbol{\theta}_j$. In general, ANN designers have no knowledge of the value of the weights and biases of an ANN, and therefore a non-informative prior (a uniform distribution) is recommended. The most important term in equation (6) is $p(D \mid \boldsymbol{\theta}_j, M_j)$, which is called the likelihood and is obtained from the set of input-target training data. Under the assumption of independent Gaussian prediction errors, the likelihood can be expressed as

$$p\left(D \mid \boldsymbol{\theta}_j, M_j\right) = \left(\sqrt{2\pi}\sigma_\eta\right)^{-NN_O} e^{-\frac{NN_O}{2\sigma_\eta^2} J\left(\boldsymbol{\theta}_j \mid D, M_j\right)} \quad \text{for} \quad j = 1, \ldots, N_M \tag{7}$$

where N is the total number of input-target training pairs; N_O is the number of output neurons of the ANN, which is equal to the number of possible damage locations N_D; and $J(\boldsymbol{\theta}_j \mid D, M_j)$ is the contribution of the training data to the likelihood, which is given by

$$J\left(\boldsymbol{\theta}_j \mid D, M_j\right) = \frac{1}{NN_O} \sum_{k=1}^{N} \left\| \Psi\left(k; \boldsymbol{\theta}_j, M_j\right) - \hat{\Psi}\left(k\right) \right\|^2 \quad \text{for} \quad j = 1, \ldots, N_M \tag{8}$$

where $\Psi(k; \boldsymbol{\theta}_j, M_j)$ is the ANN output that corresponds to the k-th input for a given set of ANN parameters $\boldsymbol{\theta}_j$ of the ANN class M_j, $\hat{\Psi}(k)$ is the target corresponding to the k-th input, and $\|\cdot\|$ denotes the Euclidean norm of a vector. $J(\boldsymbol{\theta}_j \mid D, M_j)$ in equation (8) shows the performance of the ANN, and demonstrates that the smaller the value of $J(\boldsymbol{\theta}_j \mid D, M_j)$, the better the performance, as the ANN outputs are closer to the targets specified in the training data D. It is clear that the trained ANN model can be obtained by minimising $J(\boldsymbol{\theta}_j \mid D, M_j)$.

To select the 'best' ANN model class, the probability of a model class M_j conditional on the set of input-target training data D is required. This can be obtained by following Bayes' theorem again (Beck & Yuen 2004).

$$p\left(M_j \mid D, U\right) = \frac{p\left(D \mid M_j, U\right) p\left(M_j \mid U\right)}{p\left(D \mid U\right)} \quad \text{for} \quad j = 1, \ldots, N_M \tag{9}$$

where $p(D \mid M_j, U)$ is the most important part in equation (9), known as the evidence for the

ANN model class M_j, and is obtained from the set of training data D. U expresses the user's judgment of the initial plausibility of the ANN classes, which is expressed as a prior probability $p(M_j|U)$ on the ANN classes M_j, such that

$$\sum_{j=1}^{N_M} p\left(M_j|U\right)=1 \tag{10}$$

In general, there is no preference for any ANN model class, and the prior probability $p(M_j|U)$ can therefore be taken as $1/N_M$ for all model classes. The ANN class to be used is obviously that which maximises the probability $p(M_j|D,U)$ given in equation (9). As U is irrelevant in $p(D|M_j,U)$ and M_j alone specifies the PDF for the data, U can be dropped in the notation. This means that maximising $p(M_j|D,U)$ is equivalent to maximising the evidence $p\left(D|M_j\right)$ with respect to j.

Based on an asymptotic approximation (Papdimitriou *et al.* 1997), the evidence can be calculated as follows (Beck & Yuen 2004)

$$p\left(D|M_j\right)\approx\left(2\pi\right)^{\frac{N_j}{2}}p\left(\hat{\boldsymbol{\theta}}_j|M_j\right)p\left(D|\hat{\boldsymbol{\theta}}_j,M_j\right)\left|\mathbf{H}_j\left(\hat{\boldsymbol{\theta}}_j\right)\right|^{-\frac{1}{2}}\ \text{for}\ \ j=1,\ldots,N_M \tag{11}$$

where the optimal ANN parameter $\hat{\boldsymbol{\theta}}_j$ is the most probable value obtained by maximising the posterior PDF $p(\boldsymbol{\theta}_j|D,M_j)$ in equation (6), and $\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)$ is the Hessian matrix of the function $g(\boldsymbol{\theta}_j)$ evaluated at the optimal parameters $\hat{\boldsymbol{\theta}}_j$ and is given by

$$g\left(\boldsymbol{\theta}_j\right)=-\ln\left[\,p\left(\boldsymbol{\theta}_j|M_j\right)p\left(D|\boldsymbol{\theta}_j,M_j\right)\right]\ \text{for}\ \ j=1,\ldots,N_M \tag{12}$$

The factor $|\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)|^{-1/2}$ in equation (11) provides a way of quantifying the complexity of the ANN model class M_j. An ANN model class that is not sensitive to the parameter values corresponds to a large value of $|\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)|^{-1/2}$. In other words, the factor $|\mathbf{H}_j(\hat{\boldsymbol{\theta}}_j)|^{-1/2}$ serves as a natural penalty factor for the complexity of a model class, and negates the need for ad hoc assumptions.

With the help of the Bayesian model class selection method, the "optimal" model class can be selected from among a given number of model classes based on the calculated evidence of each model class. A computationally efficient numerical algorithm can then be developed for ANN design that considers the different activation functions and the number of hidden neurons.

Consider the general case in which there are N_A activation functions available. The objective of the algorithm is to identify the 'best' activation function among them and the corresponding 'optimal' (or 'best') number of hidden neurons. The algorithm (as shown in Figure 2) consists of two loops. The outer loop with counter a is used to identify the 'best' activation function, and the inner iteration loop with counter n identifies the 'best' number of hidden neurons for the a-th activation function in the list. The outer loop will be repeated N_A times ($a = 1, \ldots, N_A$). For a given value of a, the algorithm consists of a series of iterative steps. In the general n-th iteration step, the algorithm compares the evidence of the ANN with n hidden neurons, $e(a,n)$ with the evidence of the ANN with $n + 1$ hidden neurons, $e(a,n+1)$. If $e(a,n+1)$ is larger, then the iteration will continue by increasing the value of n by 1 ($n = n + 1$). Otherwise, the 'best' number of hidden neurons for the a-th activation function is obtained. If a is equal to 1 (the first activation function on the list), then the algorithm will assign the current activation function as the 'best' function (i.e., a_{best} in Figure 2), and record the 'best' number of hidden neurons (n_{best} in Figure 2) and the corresponding evidence (i.e., e_{max} in Figure 2). If the current value of a is not equal to N_A, then the loop for a will continue. At the a-th step (where a is larger than 1), the algorithm will compare the 'optimal' evidence of the current step, $e(a,n)$ with the recorded maximum evidence, e_{max}. The values of e_{max}, a_{best} and n_{best} will then be replaced by those in the current step if $e(a,n)$ is larger than e_{max}. After repeating the outer loop N_A times, the 'best' activation function, the 'best' number of neurons and the corresponding value of the evidence can be obtained.

The use of the proposed Bayesian ANN design algorithm avoids the problems of under-fitting (the ANN being insufficiently complex to fit the data), over-fitting (the ANN being overly complex) and fluctuation between the training points. As has been stated, the single-hidden-layer feedforward type of ANN is the most commonly used type among structural engineers, yet in the literature it is usually designed using ad hoc rules of thumb and the subjective judgment of engineers. The Bayesian ANN design algorithm proposed in this paper provides a systematic, mathematically rigorous and practical method for the design of this type of ANN, which is particular important when the ANN must be trained in real time.

3 IASC-ASCE SHM Benchmark Study

3.1 Benchmark study description

A brief description of the Phase I IASC-ASCE SHM Benchmark study is given here for completeness, but more detail can be found in Johnson *et al.* (2004). The benchmark structure is a four-storey, two-bay by two-bay steel frame. In this paper, only the first five cases of the benchmark problem are considered. Depending on the case, the 12-DOF or the 120-DOF model is used to simulate the dynamic responses of the benchmark structure (see Table 1). In all cases, a 12-DOF shear building model is employed as the identification model, and thus the effect of modelling error is considered for cases with the dynamic responses generated by the 120-DOF model. Both the symmetric and asymmetric mass distributions of the structure are considered. The characteristics of the five cases investigated are summarised in Table 1, the six damage patterns (DP) considered in the benchmark study are summarised in Table 2 and the corresponding percentage reductions in the horizontal stiffness are summarised in Table 3.

3.2 Identification of modal parameters and Ritz vectors from the dynamics data in the benchmark study

MODE-ID (Beck 1978; Beck *et al.* 1994) is employed to identify the modal parameters (natural frequencies and mode shapes) of the first four translation modes (two in the x-direction and two in the y-direction) from the first 40 sec of noisy floor acceleration time histories with a time step of $\Delta t = 0.001$ sec. It is assumed that the excitation is not measured for all of the cases examined. The identified natural frequencies of Cases 1 to 5 are summarised in Table 4, in which the labels "1x" and "2x" represent the first and second modes in the x-direction, respectively, and "1y" and "2y" represent the first and second modes in the y-direction, respectively. The identified natural frequencies in Table 4 are close to the values given in the benchmark study (Johnson *et al.* 2004). The identified natural frequencies and mode shapes are employed to calculate the first two Ritz vectors following the procedure in Sohn and Law (2001). The order of modes before and after damage may be different, and thus the modal assurance criteria (MAC) (Allemang & Brown 1982) is used to calculate the damage-induced changes in the modal parameters to ensure that the correct modes are matched.

3.3 Bayesian ANN design for structural damage detection

Four ANNs are designed, two of which adopt damage-induced modal parameter changes (PF I) and two of which adopt damage-induced Ritz vector changes as the ANN inputs. The ANNs are named ANNx1 (x-direction, PF I), ANNy1 (y-direction, PF I), ANNx2 (x-direction, PF II) and ANNy2 (y-direction, PF II). The last two characters of the ANN identity represent the direction of interest ("x" or "y") and the type of pattern feature (where "1" represents damage-induced modal parameter changes and "2" represents damage-induced Ritz vector changes). Only the first two modes are employed to calculate the pattern features $\Delta\mathbf{P}_1$ and $\Delta\mathbf{P}_2$ (i.e., only two modes are used for damage detection). Thus, the four ANNs are designed by following the proposed ANN design algorithm based on four sets of input-target training data (i.e., PF I and PF II for both x- and y-direction).

In the benchmark study, damage is interpreted as the reduction in the stiffness of the horizontal storey of the steel frame. A 12-DOF shear building model (as specified in the benchmark study) is employed to generate the pattern features of a series of damage scenarios for each ANN. In the benchmark structure, there are four possible damage locations in either the x- or y-direction, that is, the interstorey stiffness at the first to fourth stories. Only five damage levels are considered in this study: a 0%, 20%, 40%, 60% and 80% reduction in stiffness. In addition, only two out of the four simultaneous types of damage are considered in generating the training data, and thus the total number of damage patterns considered is 113. None of the damage patterns in the benchmark study is included in the set of the training data.

The design of ANNs that use PF I as the ANN input is discussed in detail as follows.. As the damage-induced changes in the modal parameters (PF I) of the first two modes are employed as the ANN input, the number of input neurons is equal to ten (N_I = 10). The number of output neurons is four, because there are four possible damage locations in each direction (N_O = 4). To demonstrate the ability of the proposed ANN design algorithm to deign activation functions, two activation functions, the hyperbolic tangent sigmoid (tansig) and saturating linear transfer functions (satlin), are included in the ANN design process (i.e., N_A = 2). The functional relationship between the input x_{in} and output x_{out} of tansig and satlin are shown in equations (13) and (14), respectively.

$$x_{out} = \frac{2}{1+e^{-2x_{in}}} - 1 \tag{13}$$

$$x_{out} = \begin{cases} -1 & if\ x_{in} < -1 \\ x_{in} & if\ -1 \le x_{in} \le 1 \\ 1 & if\ x_{in} > 1 \end{cases} \quad (14)$$

For ANNy1, the proposed algorithm starts by considering the tansig activation function ($a = 1$) with two hidden neurons (n = 2). When the number of hidden neurons is small, say 2, the value of the evidence is small, because the performance of the ANN is poor. Figure 3 shows the logarithm of the evidence, log(e(a,n)) for a = 1, for different numbers of hidden neurons. When the number of hidden neurons increases, the value of the evidence increases until n = 17, after which it drops. According to the proposed Bayesian ANN design algorithm then, n = 17 is the 'optimal' number of hidden neurons for an ANN with tansig as the activation function (a = 1). The proposed algorithm records the 'best' activation function (a_{best} = 1) and the corresponding number of 'best' hidden neurons (n_{best} = 17) and the logarithm of the maximum evidence (log(e_{max}) = 1575.67). Note that the logarithm of the evidence is recorded to prevent numerical difficulties, as the value of evidence is very large. The algorithm then continues by increasing the value of a by one (i.e., by considering the second considered activation function, satlin). The algorithm again calculates the optimal number of hidden neurons (n = 21) and the corresponding logarithm of evidence (log(e(2,21) = 801.59) for the satlin activation function (a = 2). The algorithm compares the value of e(2,21) with the recorded e_{max} (= 1575.67). As the e_{max} is larger, the algorithm ends with a_{best} = 1 and n_{best} = 17. Therefore, tansig is selected as the 'best' activation function, with an 'optimal' number of hidden neurons of 17.

The detailed designs of the ANNs with two activation functions in the hidden layer are given in Table 5. The table shows the values of the J function in equation (8) and the corresponding logarithm of the evidence in equation (11) of ANNs with different numbers of hidden neurons for each activation function. It is clear from the table that the value of J decreases when the number of hidden neuron increases, and that if the value of J is the sole consideration in the ANN design, then the ANN model class with the largest number of hidden neurons will always be selected. It is therefore pointless to select the number of hidden neurons by solely comparing the values of the J function. Table 5 also shows that the logarithm of the evidence of the ANN with the tansig activation function (a = 1) and n = 17 is the highest for both ANNx1 and ANNy1. This means that the corresponding 'optimal' number of hidden neurons is 17 for both the x- and y- directions. This can be explained by the fact that both the number of input and output neurons and the number of training pairs are the same for both ANNs, and that the benchmark structure is symmetric in both directions. It is clear from Table 5 that the proposed ANN design algorithm concludes that the optimal number of hidden neurons

is 17 and the ‘best’ activation function is tansig for both ANNx1 and ANNy1.

When tansig is employed as the activation function, only 17 hidden neurons are required, but when satlin is used 21 hidden neurons are needed (see Table 5). This implies that tansig allows a less complex ANN model class to be employed for the benchmark problem. Due to space limitations, only the results of the ANNs with tansig as the activation function are included in the following discussion.

One of the objectives of this study is to investigate the relative performance of modal parameters and Ritz vectors in structural damage detection using pattern recognition. To achieve this, another two ANNs (ANNx2 and ANNy2) are constructed by the proposed ANN design method using the damage-induced Ritz vector changes as the pattern features. Only the first two Ritz vectors are used to construct the calculated pattern feature $\Delta\mathbf{P}_2$. The number of input neurons is equal to eight (N_I = 8) and the number of output neurons is four (N_O = 4). The design results are shown in Table 6. According to the proposed algorithm, the ‘best’ activation is tansig and the corresponding ‘optimal’ number of hidden neurons is 16. Table 7 summarises all of the ANNs employed in the case study and the number of hidden neurons in each of them.

Note that the number of hidden neurons for the ANNs that use PF I is a little larger (more complex) than the number for those that use PF II. This can be explained by the fact that the number of input neurons for ANNs that use PF I is larger than the number for ANNs that use PF II.

Consider a commonly used rule of thumb for the design of the hidden neuron number (Kermanshahi 1999)

$$n = \frac{N_I + N_O}{2} + \alpha \tag{15}$$

where α is taken to be 1 or 2 (according to the judgment of the designer). In the case study, N_I and N_O are equal to 10 and 4 for the PF I ANN, and 8 and 4 for the PF II ANN, respectively. The number of hidden neurons determined by this rule of thumb is thus 8 or 9 for the PF I ANN and 7 or 8 for the PF II ANN (as α can take the value of 1 or 2). These values are much smaller than those calculated by the proposed Bayesian ANN design algorithm, but the results generated by the algorithm are believed to be more reliable, as the quantity of training data is not considered in Kermanshahi’s rule of thumb. Another popular rule of thumb is (Ward Systems Group, Inc. 2000)

$$n = \frac{N_I + N_O}{2} + \sqrt{N_D} \tag{16}$$

where N_D is the number of input-target training pairs, which is equal to 113 in this case study. Using this rule of thumb, the number of hidden neurons is 18 and 17 for the PF I ANN and PF ANN II, respectively. Although these values are considerably closer to those generated by the proposed ANN design method, the rule of thumb has no theoretical background and may not be generally applicable, whereas the proposed ANN design algorithm is based on the axioms of probability.

3.4 Cases 1 to 3

In Cases 1 to 3, the symmetrical mass distribution is considered. In Cases 1 and 2, the excitation is only applied in the y-direction and the mass distribution is symmetrical, and thus only stiffness changes in the y-direction are considered. ANNy1 and ANNy2 are therefore employed to identify the damage for Cases 1 to 2 in the y-direction. In Case 3, ANNx1, ANNy1, ANNx2 and ANNy2 are employed to identify damage in both the x- and y-directions. Tables 9 and 10 show the damage detection results generated by fitting the measured pattern features PF I and PF II, respectively, to the input layers of the corresponding trained ANNs. The values in brackets are the true percentage reductions in horizontal stiffness specified by the benchmark study (Johnson *et al.* 2004). In these tables, DP refers to the damage pattern.

Table 8 shows the results for the ANN that use PF I (damage-induced modal parameter changes) as the pattern feature. For Case 1, the identified percentage reduction in horizontal stiffness at the damage locations (e.g., Storey 1 in the y-direction and DP 1) is very close to the true value shown in the brackets. For locations without any damage (e.g., Storey 2 in the x-direction and DP 1), the identified percentage reduction in stiffness is either zero or a very small number. The largest is 2.11 at Storey 4 in the y-direction and DP 2, which is a very small number compared with the percentage reduction at the damaged location. It can thus be concluded that the proposed structural damage detection method can successfully identify all of the damage scenarios in Case 1 by using PF I.

Table 9 shows the results for the ANNs that use PF II (damage-induced changes in Ritz vector) as the pattern feature. In Case 1, the identified reduction in stiffness at locations with damage agrees with the true value given in the brackets. However, for undamaged locations, some of the identified stiffness reductions are not as small as those generated by the PF I ANNs (e.g., 5.36% at Storey 3 in the y-direction and DP 1). When the threshold value of 5% is employed to distinguish between the undamaged and damaged scenarios, one false alarm is

generated by the ANNs trained by PF II, whereas no such alarm arises with the ANNs trained by PF I. In this respect, ANNs trained by modal parameters have a relatively higher performance than those trained by Ritz vectors for Case 1.

Both measurement noise and modelling error are considered in Case 2, and it is thus a more difficult case than Case 1. Table 8 shows (Case 2, y-direction, DP 1 and 2) that the damage scenarios identified by the ANNs trained by PF I are very close to the true result given in the brackets. At the undamaged locations, the identified reduction in stiffness is either zero or equal to a very small value, say 1.63% at Storey 2 of Case 2 in the y-direction and DP 1. It is therefore reasonable to conclude that the proposed method can successfully identify the damage scenarios in Case 2 by using model parameters.

Table 9 (Case 2, y-direction, DP 1 and 2) shows that the results from the ANNs trained by PF II are not as good as those from the ANNs trained by PF I, as the identified reduction in stiffness at the first storey of DP 1 is 78.56%, which is not close to the true value of 71.03%. This indicates that ANNs trained by Ritz vectors are not as good as those trained by modal parameters for Case 2.

The results for Case 3 (see Table 8) show that the ANNs trained by PF I successfully identify the stiffness reductions in both the x- and y-directions for all of the damage patterns. For storeys without damage, the estimated stiffness reductions are either zero or a very small number (less than 2%). Although similar results are obtained by the ANNs trained by PF II (see Case 3 in Table 9), the performance of the ANNs trained by Ritz vectors is clearly poorer than that of the ANNs trained by modal parameters.

3.5 Cases 4 and 5

In Cases 4 and 5 the asymmetric mass distribution is considered. As the 12-DOF identification model has a symmetric mass distribution, there is modelling error in both cases. In addition to the error in mass distribution, the dynamic responses in Case 5 are generated by the 120-DOF system, which also gives rise to modelling error in the stiffness distribution. The measurement noise is also considered in both cases.

The results for the detection of structural damage with ANNx1, ANNy1, ANNx2 and ANNy2 are summarised in Tables 10 and 11 for the ANNs trained by PF I and II, respectively.

The results for Case 4 in Tables 10 (PF I) and 11 (PF II) show that the identified stiffness reductions are very close to the true values, with the performance of the ANNs trained by PF I

being slightly better than that of the ANNs trained by PF II.

DP 6 corresponds to a 1/3 loss in stiffness of a brace at the first storey. Due to the small reduction in stiffness, DP 6 is the most difficult damage pattern. The results for DP 6 in Case 4 (Tables 10 and 11) show that the identified stiffness reduction in the y-direction at the first storey is 7.43% for the ANNs trained by PF I and 8.35% for the ANNs trained by PF II. Both of these results are significantly close to the true value 5.92% shown in the brackets. The estimated percentage stiffness reductions for the undamaged locations are very small in both Tables 10 and 11. It can thus be concluded that the performance of the proposed method is reasonably good even for DP 6 when the extent of the damage is very small.

Because of the relatively large modelling error in Case 5, the results (Tables 10 and 11) are not as good as those for Case 4. However, the identified stiffness reductions generated by the ANNs trained by PF I and by PF II still have an acceptable accuracy. Although some of the identified stiffness reductions at damage locations are not very close to the true values, they are sufficiently large to indicate the existence of damage. Equally, although some of the identified stiffness reductions at the undamaged locations are not equal to zero, they are relatively small compared with those at the damaged locations. If a 5% stiffness reduction is considered as the threshold value for damage, then 1 and 4 false indicators out of 48 indicators are generated by the ANNs trained by PF I and by PF II, respectively. As the method does not "miss" any damage, the damage detection results are safe.

It must be pointed out that only five damage levels (a 0%, 20%, 40%, 60% and 80% reduction in stiffness) are considered in generating the training data. It is believed that the damage detection of the trained ANNs could be further improved by considering a finer division of damage levels.

4 Concluding Remarks

In this paper, a pattern recognition approach to structural damage detection is presented that uses an ANN as a systematic and efficient tool for matching the measured and calculated pattern features. An extended Bayesian ANN design algorithm is proposed for the real-time design of ANNs that considers both the optimal number of hidden neurons and the activation function in the hidden layer. The proposed algorithm ensures that the trained ANN is on the one hand sufficiently complex to provide good performance, and on the other hand sufficiently simple to avoid the overfitting of the data and the fluctuation in output between the training data

points. The results of the case study suggest that tansig is a better activation function in the hidden layer than satlin in terms of damage detection in the benchmark structure. The ANN design algorithm can also be applied to the design of general-purpose ANNs.

The proposed ANN design and structural damage detection methods are verified using the first five cases of the Phase I IASC-ASCE SHM benchmark study. By using the benchmark study as a fair platform for comparison, the relative performance of ANNs that adopt modal parameters and those that adopt Ritz vectors in structural damage detection using the pattern recognition approach is examined. Four ANNs are constructed to address the Phase I IASC-ASCE benchmark problem, two of which are trained by damage-induced modal parameter changes and two of which are trained by damage-induced Ritz vector changes. The results are very encouraging. Most of the stiffness reductions identified by the trained ANNs are very close to the true values given in the benchmark study. Although the identified stiffness reductions at several locations in Case 5 (which has a large modelling error and measurement noise) are not as accurate as those in other cases, they are good enough to indicate the existence of damage. The results also show that the performance of ANNs trained by modal parameters is slightly better than that of ANNs trained by Ritz vectors.

5 Acknowledgements

The work that is described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 114706).

6 References

[1] Allemang, R.J. and Brown, D.L. 1982. A correlation coefficient for modal vector analysis, Proceedings of the 1st IMAC, Orlando, Florida, pp. 110-116.

[2] Beck, J.L. 1978. Determining models of structures from earthquake records, Report No. EERL 78-01, California Institute of Technology, Earthquake Engineering Research Laboratory, Pasadena, California.

[3] Beck, J.L. and Katafygiotis, L.S. 1998. Updating models and their uncertainties I: Bayesian statistical framework, Journal of Engineering Mechanics, ASCE, 124(4), pp. 455-461.

[4] Beck, J.L. and Yuen, K.V. 2004. Model selection using response measurement: A Bayesian probabilistic approach, Journal of Engineering Mechanics, ASCE, 130(2), pp. 192-203.

[5] Beck, J.L., May, B.S., and Polidori, D.C. 1994. Determination of modal parameters from ambient vibration data for structural health monitoring, Proceedings of the 1st World Conference on Structural Control, Pasadena, California, TA3:3-TA3:12.

[6] Bernal, D. and Beck, J.L. 2004. Special issue on phase I of the IASC-ASCE structural health monitoring benchmark, Journal of Engineering Mechanics, ASCE, 130(1).

[7] Cao, T.T. and Zimmerman, D.C. 1997. A procedure to extract Ritz vectors from dynamics data, Proceedings of the 15th IMAC, Orlando, Florida, pp. 1036-1042.

[8] Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function, Mathematics of Control, Signals and Systems, 2, pp. 303-314.

[9] Elkordy, M.F., Chang, K.C. and Lee, G.C. 1994. A structural damage neural network monitoring system, Microcomputers in Civil Engineering, 9, pp. 83-96.

[10] Hagan, M.T. and Menhaj, M. 1994. Training feedforward networks with the Marquardt algorithm, IEEE Transactions on Neural Networks, 5(6), pp. 989-993.

[11] Johnson, E.A., Lam, H.F., Katafygiotis, L.S., and Beck, J.L. 2004. Phase I IASC-ASCE structural health monitoring benchmark problem using simulated data, Journal of Engineering Mechanics, ASCE, 130(1), pp. 3-15.

[12] Kermanshahi, B. 1999, Design and application of neural networks, Shokodo, Tokyo, Chapter 3.

[13] Lam, H.F. 1994. Detection of structural faults via experimental modal analysis. MPhil Thesis, Civil Engineering Department, Hong Kong Polytechnic University.

[14] Lam, H.F., Ko, J.M. and Wong, C.W. 1998. Localization of damaged structural connections based on experimental modal and sensitivity analysis, Journal of Sound and Vibration, 210(1), pp. 91-115.

[15] Lam, H.F., Ng, C.T. and Veidt M. 2007. Experimental characterization of multiple cracks in a cantilever beam utilizing transient vibration data following probabilistic approach, Journal of Sound and Vibration, 305(1-2), pp. 34-49.

[16] Lam, H.F., Yuen, K.V., and Beck, J.L. 2006. Structural health monitoring via measured Ritz vectors utilizing artificial neural networks, Computer-Aided Civil and Infrastructure Engineering, 21, pp. 232-241.

[17] Ng, C.T. 2007. Probabilistic structural health monitoring and reliability analysis, MPhil. Thesis, City University of Hong Kong.

[18] Ng, C.T. and Lam, H.F. 2006. Application of pattern recognition approach on Phase I of the IASC-ASCE structural health monitoring benchmark study, Proceedings of the 5th International Conference on Computational Stochastic Mechanics, CSM5, 20th – 23rd June 2006, Rhodes, Greece, pp. 525-534.

[19] Papadimitriou, C. 2004a. Bayesian inference applied to structural damage detection. In Proceedings of the 2nd European Workshop on Structural Health Monitoring, Boller C. and Staszewski W.J. (Eds), DEStech Publications, pp. 575-582.

[20] Papadimitriou, C. 2004b. Bayesian inference applied to structural model updating and damage detection. 9th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability, Albuquerque, New Mexico.

[21] Papadimitriou, C., Beck, J.L., and Katafygiotis, L.S. 1997. Asymptotic expansion for reliability and moments of uncertain systems, Journal of Engineering Mechanics, ASCE, 123(12), pp. 1219-1229.

[22] Sohn, H. and Law, K.H. 2001. Damage diagnosis using experimental Ritz vectors, Journal of Engineering Mechanics, 127(11), pp. 1184-1193.

[23] Sohn, H., Farrar, C.R., Hernez, F.M., Czrnecki, J.J., Shunk, D.D., Stinemates, D.W., and Nadler, B.R., 2004. A review of structural health monitoring literature: 1996 – 2001, Report No. LA-13976-MS, Los Alamos National Laboratory, Los Alamos, NM.

[24] Ward Systems Group, Inc. 2000, NeuroShell 2 Manual.

[25] Wu, X., Ghaboussi, J., and Barrett, J.H. 1992. Use of neural networks in detection of structural damage, Computers and Structures, 42(4), pp. 649-659.

[26] Yuen, K.V. and Lam, H.F. 2006. On the complexity of artificial neural networks for smart structures monitoring, Engineering Structures, 28, pp. 977-984.

List of Figures

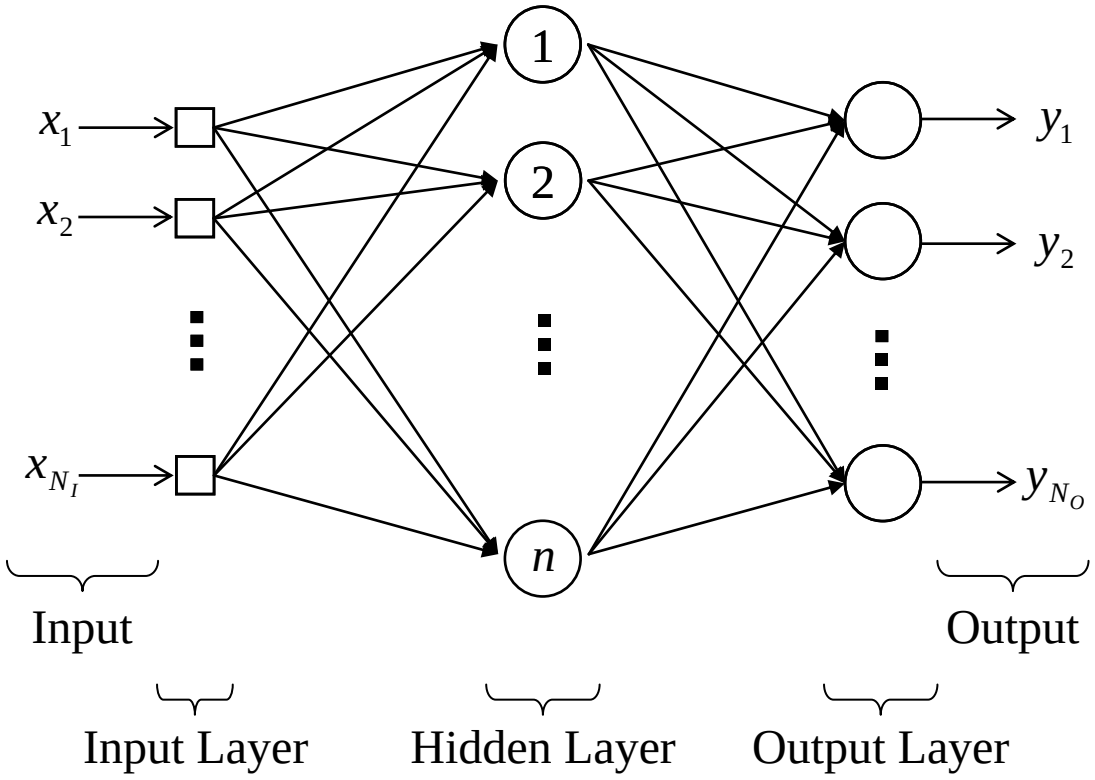

Figure 1: Structure of the single-hidden-layer feedforward ANN.

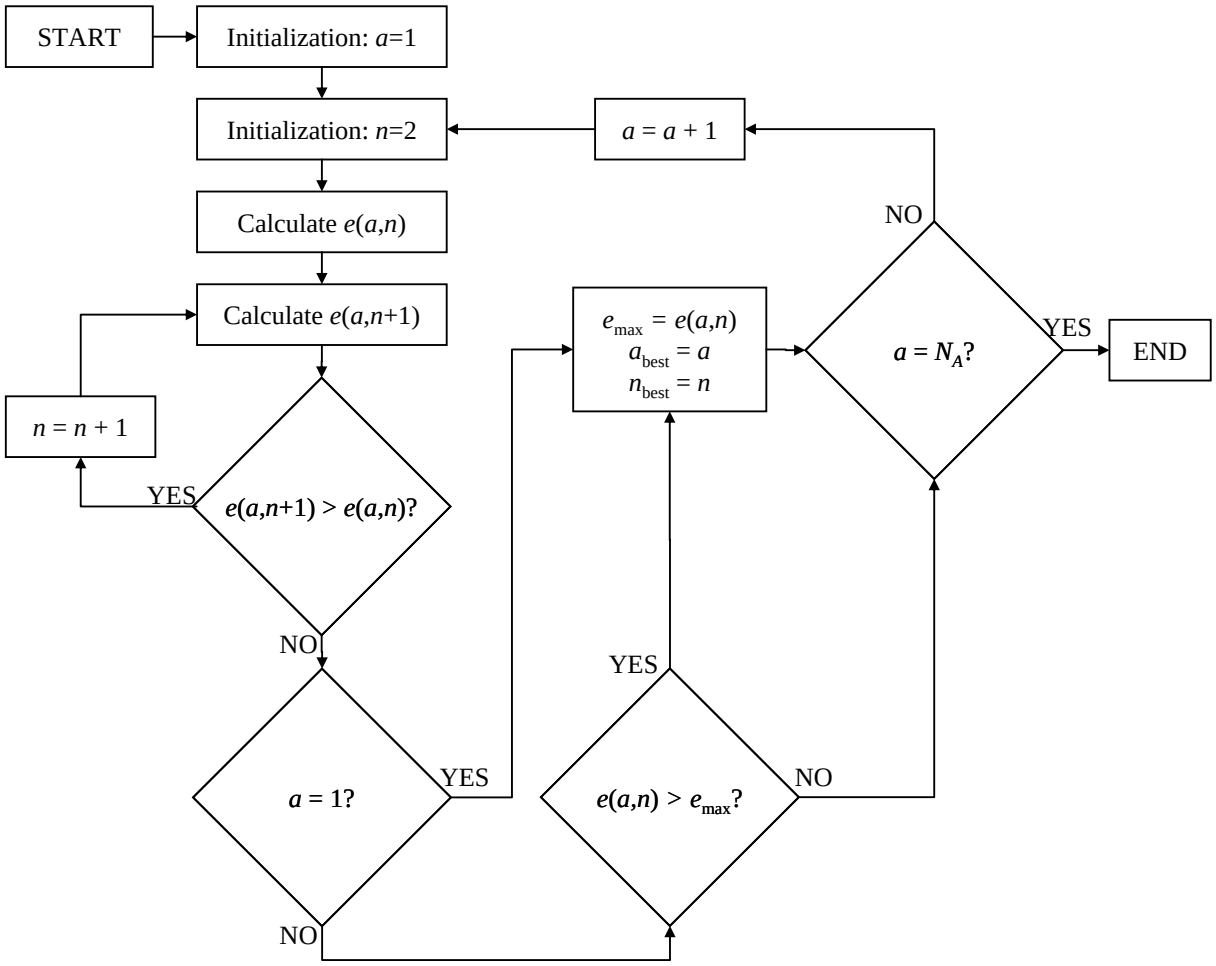

Figure 2 : Proposed Bayesian ANN design algorithm for selecting the 'optimal' number of hidden neurons and the 'best' activation function.

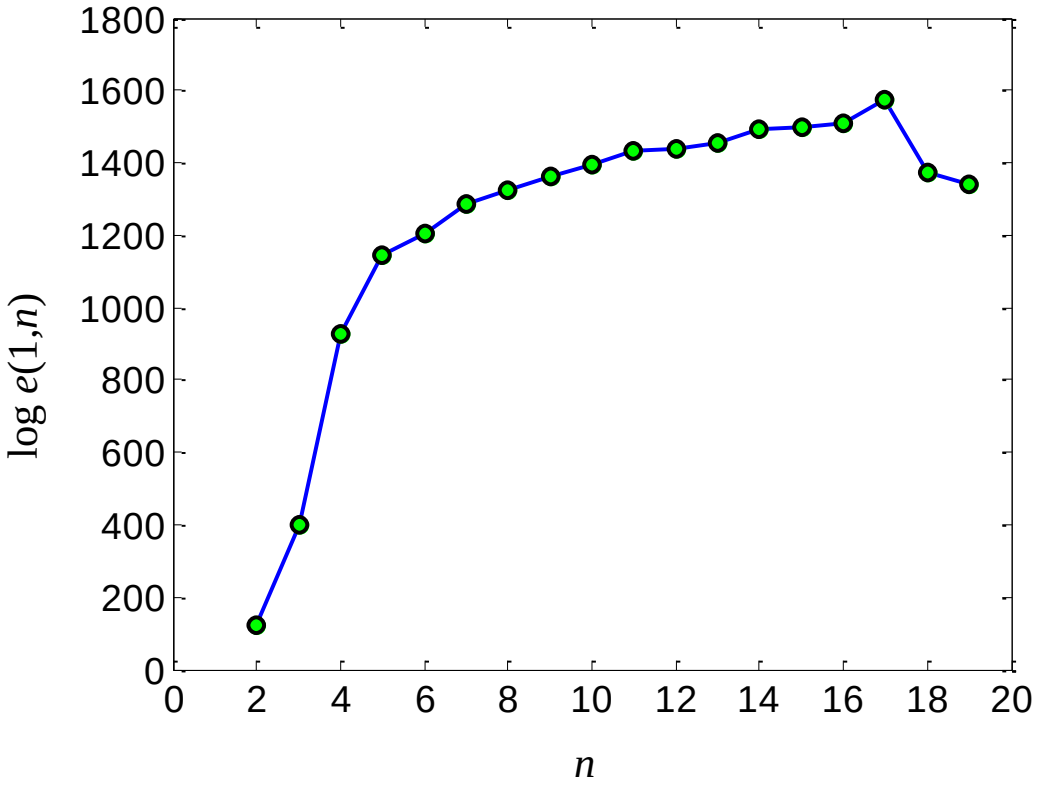

Figure 3: Iteration history of the selection of the optimal number of hidden neurons for ANNy1 with tansig as the activation function (a = 1).

List of Tables

Case	Data generation model	Mass distribution	Excitation	Damage patterns
1	12 DOF	Symmetric	Ambient	1,2
2	120 DOF	Symmetric	Ambient	1,2
3	12 DOF	Symmetric	Shaker on roof	1,2
4	12 DOF	Asymmetric	Shaker on roof	1- 4 & 6
5	120 DOF	Asymmetric	Shaker on roof	1-6

Table 1: First five cases of the IASC-ASCE SHM benchmark study.

Damage Pattern	Description
1	All braces in the first storey are removed.
2	All braces in the first and third storeys are removed.
3	One brace in the first storey is removed.
4	Two braces, one in the first and one in the third storey, are removed
5	The same as damage pattern 4 but with one additional damage that is one beam-column connection at the first floor is lessened.
6	1/3 stiffness loss in one brace, which is the same brace damaged in damage pattern 3.

Table 2: The six damage patterns considered in the SHM benchmark study.

Direction	DP	Storey			
		1	2	3	4
x-dir.	1	45.24	0.00	0.00	0.00
	2	45.24	0.00	45.24	0.00
	3	0.00	0.00	0.00	0.00
	4	0.00	0.00	11.31	0.00
	5	0.00	0.00	11.31	0.00
	6	0.00	0.00	0.00	0.00
y-dir.	1	71.03	0.00	0.00	0.00
	2	71.03	0.00	71.03	0.00
	3	17.76	0.00	0.00	0.00
	4	17.76	0.00	0.00	0.00
	5	17.76	0.00	0.00	0.00
	6	5.92	0.00	0.00	0.00

Table 3: Percentage reduction in stiffness for all damage patterns extracted from Table 4 of Johnson *et al.* (2006)

Case	DP	Mode			
		1y	2y	1x	2x
1	0	9.40	25.35	---	---
	1	6.26	21.48	---	---
	2	5.82	14.89	---	---
2	0	8.61	23.46	---	---
	1	5.48	19.37	---	---
	2	4.97	12.37	---	---
3	0	9.43	25.47	11.79	31.95
	1	6.27	21.50	9.88	28.90
	2	5.82	14.90	9.50	24.79
4	0	9.33	25.32	11.65	31.65
	1	6.16	21.26	9.81	28.47
	2	5.78	14.79	9.41	24.60
	3	8.84	24.50	11.65	31.65
	4	8.84	24.50	11.44	30.67
	6	9.15	24.95	11.65	31.65
5	0	8.47	23.10	9.04	25.56
	1	5.45	18.99	7.29	22.52
	2	4.86	12.37	6.63	17.77
	3	8.02	22.22	9.04	25.56
	4	8.01	22.20	8.85	24.62
	5	7.92	22.17	8.85	24.62
	6	8.36	22.88	9.04	25.56

Table 4: Identified natural frequencies (Hz) for Cases 1 to 5

a	n	ANNx1		ANNy1	
		J	$\log p(D \mid M_j)$	J	$\log p(D \mid M_j)$
1 (tansig)	16	5.17×10^{-7}	1467.02	4.72×10^{-7}	1508.54
	17	2.51×10^{-7}	**1588.67**	2.91×10^{-7}	**1575.67**
	18	1.32×10^{-7}	1391.60	1.92×10^{-7}	1371.61
	19	1.01×10^{-7}	1348.45	1.09×10^{-7}	1341.95
2 (satlin)	20	6.68×10^{-5}	760.46	5.09×10^{-5}	736.42
	21	2.43×10^{-5}	775.44	2.25×10^{-5}	801.59
	22	2.18×10^{-5}	699.16	1.78×10^{-5}	718.48
	23	1.17×10^{-5}	677.60	1.43×10^{-5}	638.27

Table 5: ANN model class selection for ANNx1 and ANNy1 trained by PF I

n	ANNx2		ANNy2	
	J	$\log p(D \mid M_j)$	J	$\log p(D \mid M_j)$
15	4.72×10^{-5}	999.04	5.35×10^{-5}	1017.13
16	3.49×10^{-5}	**1029.94**	3.91×10^{-5}	**1058.31**
17	2.80×10^{-5}	986.81	3.23×10^{-5}	994.75
18	2.46×10^{-5}	927.12	2.66×10^{-5}	938.43

Table 6: ANN model class selection for ANNx2 and ANNy2 trained by PF II

ANN ID	Pattern feature	Mass distribution	Applied to Cases (direction)	n
ANNx1	$\Delta\mathbf{P}_1$	Symmetric	1 to 5 (x)	17
ANNy1	$\Delta\mathbf{P}_1$	Symmetric	1 to 5 (y)	17
ANNx2	$\Delta\mathbf{P}_2$	Symmetric	1 to 5 (x)	16
ANNy2	$\Delta\mathbf{P}_2$	Symmetric	1 to 5 (y)	16

Table 7: List of ANNs applied to the IASC-ASCE SHM benchmark study

Case	Dir.	DP	Storey			
			1	2	3	4
1	y	1	**68.99** (71.03)	0.00 (0.00)	2.38 (0.00)	0.00 (0.00)
		2	**71.09** (71.03)	1.13 (0.00)	**70.30** (71.03)	2.11 (0.00)
2	y	1	**73.13** (71.03)	1.63 (0.00)	0.00 (0.00)	0.00 (0.00)
		2	**72.07** (71.03)	1.65 (0.00)	**73.79** (71.03)	0.00 (0.00)
3	x	1	**43.42** (45.24)	0.00 (0.00)	2.29 (0.00)	1.22 (0.00)
		2	**44.99** (45.24)	1.22 (0.00)	**45.59** (45.24)	1.43 (0.00)
	y	1	**70.00** (71.03)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
		2	**72.08** (71.03)	0.00 (0.00)	**70.99** (71.03)	0.00 (0.00)

Table 8: Percentage reduction in stiffness identified by the ANNs trained by PF I in Cases 1 to 3

Case	Dir.	DP	Storey			
			1	2	3	4
1	y	1	**71.08** (71.03)	0.00 (0.00)	5.36 (0.00)	3.42 (0.00)
		2	**71.60** (71.03)	2.58 (0.00)	**72.24** (71.03)	1.54 (0.00)
2	y	1	**78.56** (71.03)	0.00 (0.00)	3.01 (0.00)	0.00 (0.00)
		2	**73.55** (71.03)	8.96 (0.00)	**73.24** (71.03)	0.00 (0.00)
3	x	1	**45.70** (45.24)	5.19 (0.00)	4.76 (0.00)	3.03 (0.00)
		2	**43.59** (45.24)	0.00 (0.00)	**43.82** (45.24)	0.00 (0.00)
	y	1	**72.65** (71.03)	4.88 (0.00)	7.86 (0.00)	5.06 (0.00)
		2	**73.20** (71.03)	2.92 (0.00)	**73.96** (71.03)	0.94 (0.00)

Table 9: Percentage reduction in stiffness identified by the ANNs trained by PF II in Cases 1 to 3

Case	Dir.	DP	Storey							
			1		2		3		4	
4	x	1	**45.89**	(45.24)	0.00	(0.00)	2.42	(0.00)	0.00	(0.00)
		2	**45.29**	(45.24)	1.10	(0.00)	**44.78**	(45.24)	0.00	(0.00)
		3	0.00	(0.00)	0.04	(0.00)	0.00	(0.00)	0.20	(0.00)
		4	0.40	(0.00)	1.50	(0.00)	**11.11**	(11.31)	2.93	(0.00)
		6	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.17	(0.00)
	y	1	**72.77**	(71.03)	0.04	(0.00)	0.59	(0.00)	0.27	(0.00)
		2	**71.93**	(71.03)	0.00	(0.00)	**70.43**	(71.03)	0.00	(0.00)
		3	**18.59**	(17.76)	0.00	(0.00)	0.04	(0.00)	0.00	(0.00)
		4	**18.53**	(17.76)	0.00	(0.00)	0.11	(0.00)	0.00	(0.00)
		6	**7.43**	(5.92)	1.67	(0.00)	0.00	(0.00)	1.22	(0.00)
5	x	1	**56.94**	(45.24)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
		2	**42.74**	(45.24)	7.57	(0.00)	**59.35**	(45.24)	0.00	(0.00)
		3	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.05	(0.00)
		4	0.00	(0.00)	2.71	(0.00)	**12.19**	(11.31)	0.63	(0.00)
		5	0.00	(0.00)	2.71	(0.00)	**12.20**	(11.31)	0.64	(0.00)
		6	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.04	(0.00)
	y	1	**74.91**	(71.03)	3.59	(0.00)	0.00	(0.00)	0.00	(0.00)
		2	**74.37**	(71.03)	0.00	(0.00)	**71.88**	(71.03)	0.00	(0.00)
		3	**19.43**	(17.76)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
		4	**19.35**	(17.76)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
		5	**21.16**	(17.76)	2.62	(0.00)	0.00	(0.00)	0.00	(0.00)
		6	**5.77**	(5.92)	2.01	(0.00)	0.00	(0.00)	0.00	(0.00)

Table 10: Percentage reduction in stiffness identified by the ANNs trained by PF I in Cases 4 and 5

Case	Dir.	DP	Storey							
			1		2		3		4	
4	x	1	**45.40**	(45.24)	4.02	(0.00)	5.20	(0.00)	2.11	(0.00)
		2	**43.99**	(45.24)	0.00	(0.00)	**43.83**	(45.24)	0.00	(0.00)
		3	3.09	(0.00)	2.90	(0.00)	3.89	(0.00)	4.19	(0.00)
		4	1.73	(0.00)	4.46	(0.00)	**13.94**	(11.31)	5.71	(0.00)
		6	3.11	(0.00)	2.84	(0.00)	3.84	(0.00)	4.25	(0.00)
	y	1	**71.95**	(71.03)	0.00	(0.00)	4.98	(0.00)	3.38	(0.00)
		2	**71.92**	(71.03)	0.00	(0.00)	**72.19**	(71.03)	0.00	(0.00)
		3	**22.33**	(17.76)	0.00	(0.00)	1.79	(0.00)	0.00	(0.00)
		4	**22.45**	(17.76)	0.00	(0.00)	1.77	(0.00)	0.05	(0.00)
		6	**8.35**	(5.92)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
5	x	1	**51.90**	(45.24)	6.00	(0.00)	4.90	(0.00)	0.00	(0.00)
		2	**45.15**	(45.24)	0.36	(0.00)	**45.56**	(45.24)	0.00	(0.00)
		3	3.44	(0.00)	3.06	(0.00)	3.96	(0.00)	4.82	(0.00)
		4	2.73	(0.00)	1.70	(0.00)	**14.07**	(11.31)	5.00	(0.00)
		5	2.73	(0.00)	1.71	(0.00)	**14.08**	(11.31)	5.00	(0.00)
		6	3.46	(0.00)	2.81	(0.00)	3.80	(0.00)	4.96	(0.00)
	y	1	**81.03**	(71.03)	0.00	(0.00)	2.42	(0.00)	0.00	(0.00)
		2	**72.01**	(71.03)	8.73	(0.00)	**73.81**	(71.03)	0.00	(0.00)
		3	**22.26**	(17.76)	0.00	(0.00)	0.82	(0.00)	0.00	(0.00)
		4	**22.32**	(17.76)	0.00	(0.00)	0.74	(0.00)	0.00	(0.00)
		5	**21.62**	(17.76)	0.00	(0.00)	1.02	(0.00)	0.00	(0.00)
		6	**7.92**	(5.92)	0.66	(0.00)	0.80	(0.00)	0.00	(0.00)

Table 11: Percentage reduction in stiffness identified by the ANNs trained by PF II in Cases 4 and 5