Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Sound and Vibration 275 (2004) 931–952
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

A two-level neural network approach for dynamic FE model updating including damping

Yong Lu*, Zhenguo Tu

School of Civil and Environmental Engineering, Nanyang Technological University, Block N1, #1a-29, 50 Nanyang Avenue, Singapore 639798, Singapore

Received 7 February 2003; accepted 26 June 2003

Abstract

This paper presents a two-level neural network scheme for finite element (FE) model updating in which both the structural parameters and the damping ratios are updated. Considering the fact that in a lightly damped system the damping has only negligible influence on the resonance and antiresonance frequencies of the system, in the first-level updating the model is assumed to be free of damping and the structural parameters are updated using the natural and antiresonance frequencies as the response data. With the updated structural parameters from the above first-level updating, the second-level updating procedure deals only with the damping ratios, using the integrals of frequency response function (FRF) as reference responses. For the selection of a proper response configuration, a sensitivity analysis scheme is proposed, taking into account the carry-over error during the first-level updating in addition to the anticipated error in the measured FRF data. Through a numerical example it is shown that the approach is effective and efficient. It is also shown that by means of a noise injection learning the neural network can acquire considerable noise-resisting ability, resulting in about 50% reduction of the errors in the updated parameters as compared to the anticipated errors from the sensitivity analysis.

1. Introduction

Updating of the finite element (FE) model parameters is an essential step towards establishing a reliable FE model for an existing structure. Successfully updated FE model enables the analysis of the structural performance under a variety of user-defined loading conditions. The identification of the structural parameters from an FE model updating procedure also allows for an effective diagnosis and assessment of the structural condition.

*Corresponding author. Tel.: +65-6790-5272; fax: +65-6791-0676.
E-mail address: cylu@ntu.edu.sg (Y. Lu).

0022-460X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0022-460X(03)00796-X

The FE model updating or structural parameter identification in essence is to achieve, by updating the model parameters, a match between the predicted system response and the response measured from the actual structure. Among all applicable candidates of structural experimental data used by model updating, the modal or vibration data (and possibly its derivative information as well) are widely adopted because of the fact that the modal information, i.e., the resonance frequencies and mode shapes only depends on the structural inherent properties irrespective of the excitation applied. The vibration data based FE model updating is also called dynamic FE model updating. The basic premise of the vibration-based model updating scheme is that the FE modelling errors will produce sensible influence on the stiffness, mass or energy dissipation properties of a system, which, in turn, affect the measured dynamic response of that system.

The current common techniques for performing dynamic model updating stem primarily from the iterative-based methods and they usually employ a gradient-descent technique to directly update the structural parameters [1]. Iterative methods allow large amount of parameters to be updated simultaneously, making these schemes very attractive, and a number of applications have been reported [2–4]. Despite the successes, however, the iterative-based updating methods bear inherent limitations due to the following facts: (a) the searching procedure may be snagged at a particular local optimum instead of the global actual solution, (b) the process is vulnerable to the measurement errors, and (c) the updating usually requires a pre-processing on the response data (data reduction/expansion), which could introduce additional errors. In these respects, the recent development in artificial intelligence algorithms—the artificial neural networks (ANN)—provides a potentially powerful alternative to the traditional iterative-based methods because of the particular working mechanism with which ANN operates, i.e., the learning phenomenon. The appealing features of ANN concerning the FE model updating include: (a) it performs a function approximation of any complexity via a learning process based on the given discrete training patterns; and once trained, new updating can be carried out readily using the network (i.e., ''generalization'' in ANN terminology); and (b) it has the potential to resist the influence of noises that are contained in the input data (structural response). Furthermore, ANN in itself does not require specific form or completeness of the response data. It is noteworthy that in fact any pertinent system information can be incorporated in an ANN-based updating procedure since such a scheme works only with the response data itself and does not require other auxiliaries such as derivative information; hence, the definition of the objective function can be extremely flexible such that the procedure can be made very robust and versatile.

In recent years, some researchers have applied ANN for structural damage detection related applications [5–8]. Much experience has been accumulated through these exploratory investigations. However, there appears to be lacking a generic procedure to apply ANN in FE model updating concerning general structural parameters, especially when damping is involved. In this paper, the multi-layer feed-forward (MLP) neural network is employed to perform a general FE model-updating task including the updating of the damping parameters. To reduce the computational demand, and considering the fact that the influence of damping on the resonance and antiresonance frequencies of a lightly damped system is negligible, the updating of the FE model parameters is divided into two levels using two separate neural networks. The first-level network is trained to update the structural parameters (stiffness in particular) using the natural and antiresonance frequencies as the response data without considering the damping effect. With the updated structural parameters from the above first-level updating, the second-level network

then deals with updating the variable damping ratios. For the updating of the damping ratios, the integrals of frequency response function (FRF) are used as the reference responses because of their inherent relationship with the damping factors. In order for the identification of a proper response configuration, a sensitivity analysis scheme is proposed, taking into account the carry-over error during the first-level updating in addition to the anticipated measurement error ("noise") in the response data. A numerical example is given to illustrate the implementation of the proposed scheme and demonstrate the effectiveness of the procedure; and in particular, the noise resisting ability of the neural network trained using the noise-injection strategy is highlighted.

2. Overview of artificial neural networks (ANN)

Artificial neural networks, or simply neural networks, are essentially computational structures that mimic the operation of biological neurons of mammalian brains. Such structures encapsulate a variety of simple processing units (artificial neurons), interconnected with each other. Each neuron receives a number of inputs and produces one output as shown in Fig. 1(a); however, there exist a few different types of neurons associated with different schemes in manipulating the input information, h, to result in the output, O.

Among many different forms of network topologies (interconnection style of neurons), the MLP network (see Fig. 1(b)), which possesses layered structure and allows only connections from neurons in one layer to those in layers of its forward direction, has been applied widely due to its effectiveness and simplicity. This special topology form is mainly designed for approximating an unknown function relation, and the ability for the network to do so is realized through a learning process on the provided data patterns (training data), whereby the interconnection weights, w, are continuously adjusted until a predefined error criterion is reached.

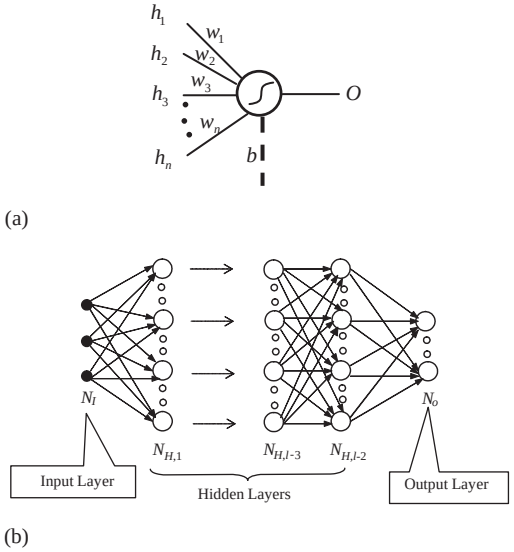

Fig. 1. Schematic illustration of the neuron (processing unit) and a typical network structure. (a) A simple neuron (processing unit): $O = \phi\left(\sum_{i=1}^{n} w_i h_i + b\right)$, ϕ: Log-sigmoid transfer function. (b) An l-layered MLP network.

In the present study, the widely used Levenburg–Marquardt (LM) back-propagation training algorithm is adopted to carry out the neural network learning towards minimizing a predefined error function, which is generally formulated as mean square error (MSE) between the network outputs and the actual values corresponding to the given set of input vectors. As a result, an optimal set of weights are obtained.

In general, the design of a MLP network needs to decide on the following: the number of hidden layers; the associated number of neurons in each hidden layer; and the interconnection patterns among the neurons. The identification of a true optimal combination can be very time consuming and some novel methods have been proposed in attempt to tackle this problem [9,10]. At the present stage, a trial procedure is still commonly practiced in the design of ANN topology for engineering applications. In fact, it has also been shown that a network with only one hidden layer suffices to approximate a large spectrum of complex functions [11,12]; therefore, the required effort in identifying a desirable network design can be considerably reduced by using just one or two hidden layers. Based on these considerations and following some trial exercises, in the present study the neural network structure is unified to contain two hidden layers and both hidden layers are to have an equal number of hidden neurons, leaving only the number of hidden neurons to be determined in the network topology design, for which a trial process is effective.

For the improvement of the network performance, a data scaling process is usually required. This is because the compiled raw training data, such as the modal frequencies of a structure used in this study, can vary significantly in their original values. When such data are directly used in the training procedure, the network could exhibit ill-conditioning and possibly does not learn at all. Besides, the application of so trained network may also bear the risk that some input components are in fact ignored; consequently, the network would no longer represent the underlying system. This problem can be avoided by a proper scaling on the raw input data patterns, such that the input data are normalized to fall within a prescribed bound, for example in the range of $[-1,1]$. For this particular value range, the transformation can be of a linear form

$$x_m^N = 2\frac{x_m - \min(x_m)}{\max(x_m) - \min(x_m)} - 1, \tag{1}$$

where x_m is a row of the input data matrix, in which each column represents one given data pattern for training, and x_m^N is the normalized quantity.

It is well recognized that the applicability of traditional updating methods is often restricted due to the measurement errors that exist in practically all measured response data. In this regard, the neural network technique is particularly appealing because of its potential noise-resisting capability. This capability can be acquired through the so-called noise-injection learning. Usually the possible margin of errors in the measured structural response information is assessable from relevant past experiences. In the implementation of the noise-injection learning algorithm for the neural network training, a similar level of random noise simulating the actual measurement errors or noises is injected into the network input data (responses) while their correct output counterparts are retained. Such noise-injection operation is straightforward and it can be expressed as

$$\tilde{R}_j = R_j(1 + v_j), \tag{2}$$

where R_j and $\tilde{R}_j$ represent the noise free (calculated) and the noise-injected response components, respectively, and v_j is a noise item simulating anticipated noise or errors in the measured response data. At this juncture, it also becomes obvious that the success of the noise-injection learning in real applications is subject to the adequacy of the noise model chosen in representing the actual noises for the particular problem under consideration.

3. Basic considerations for the two-level neural network updating scheme

3.1. Antiresonance frequencies

As damping effect is ignored at the first-level updating on structural parameters in the two-level network scheme, it is deemed appropriate not to consider mode shape data as responses because the mode shapes are more sensitive to damping. This, however, will result in a drastic reduction in the size of the available measured data set. To compensate for this, the antiresonance frequencies are considered together with the natural frequencies to enlarge the response data set.

The antiresonance frequencies are defined as the frequencies at which the magnitude of the frequency response at a measured degrees of freedom (d.o.f.) approaches zero [13], as depicted in Fig. 2(a). Lallement and Cogan [14] introduced the concept of using antiresonance frequencies to update FE models. The reason is that these antiresonance frequencies can be easily and accurately measured in a similar way as for the natural frequencies. Furthermore, a system can have much greater number of antiresonance frequencies than natural frequencies because every different

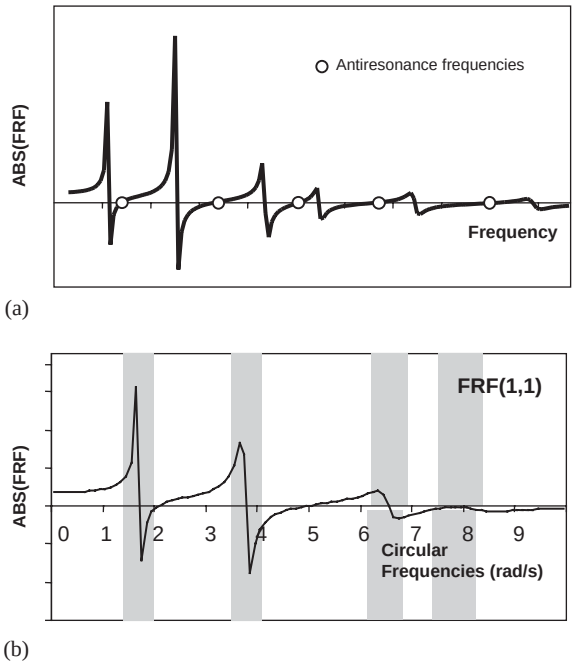

Fig. 2. Definition of antiresonance frequencies (a) and integration intervals of FRF for damping ratio updating (b).

FRF between an actuator and a sensor contains another set of antiresonance frequencies. Lallerment and Cogan referred to this increased amount of data as an "enlargement of the knowledge space" [14]. Mottershead [15] showed that the antiresonance sensitivities to structural parameters can be expressed as a linear combination of natural frequency and mode shape sensitivities, and furthermore that the dominating contributors to the antiresonance sensitivities are the sensitivities of the nearest frequencies and corresponding mode shapes. Therefore, Mottershead concluded that the antiresonance frequencies can be a preferred alternative to mode shape data.

To calculate antiresonance frequencies of a dynamic system, He and Li [16] developed an accurate and efficient method for undamped systems. Consider the eigenequation of a dynamic system given by

$$(\mathbf{K} - \omega^2\mathbf{M})\{\phi\} = 0, \tag{3}$$

where matrices $\mathbf{K}$, $\mathbf{M}$, vector $\{\phi\}$ and scalar ω denote structural stiffness matrix, mass matrix, mode shape vector and eigenfrequncy, respectively. Generally, the FRF matrix $\mathbf{H}(\omega)$ is defined as

$$\mathbf{H}(\omega) = (\mathbf{K} - \omega^2\mathbf{M})^{-1}. \tag{4}$$

The FRF for a sensor at d.o.f. p and an actuator at d.o.f. q should be the pqth element of $\mathbf{H}(\omega)$:

$$H_{pq}(\omega) = (\mathbf{K} - \omega^2\mathbf{M})^{-1}_{pq} = \frac{adj(\mathbf{K} - \omega^2\mathbf{M})_{pq}}{\mathrm{Det}(\mathbf{K} - \omega^2\mathbf{M})} = (-1)^{p+q}\frac{\mathrm{Det}(\mathbf{K}_{pq} - \omega^2\mathbf{M}_{pq})}{\mathrm{Det}(\mathbf{K} - \omega^2\mathbf{M})}, \tag{5}$$

where $\mathbf{K}_{pq}$ and $\mathbf{M}_{pq}$ indicate that row p and column q are deleted from the matrices. Therefore, the antiresonance frequencies ω_a for the FRF between p and q are the positive roots of the following equation:

$$\mathrm{Det}(\mathbf{K}_{pq} - \omega_a^2\mathbf{M}_{pq}) = 0 \tag{6}$$

for systems involving the damping effect, Eq. (3) should be modified to incorporate damping in order to calculate its antiresonance natural frequencies, which in this case are complex quantities. It is worth noting that matrices $\mathbf{K}_{pq}$ and $\mathbf{M}_{pq}$ in Eq. (6) remain symmetric in case of collocated sensor and actuator, i.e., $p = q$, and therefore the corresponding FRFs will always contain real antiresonance frequencies; otherwise, antiresonance frequencies of complex conjugates can be expected. In this research, only those antiresonance frequencies obtained from collocated FRFs are considered.

3.2. Damping model

The identification of damping in structural systems is extremely important if a model is to predict reliably the transient responses, transmissibility, decay times or other characteristics in design and analysis that are dominated by energy dissipation. However, unlike the overall stiffness and mass matrices, the damping matrix $\mathbf{C}$ cannot be constructed from the element damping matrix [17]. In the present study, it is not intended to carry out an in-depth investigation on the most appropriate damping model for a given system; but rather, the focus is placed on the identification of the damping values under a preselected damping model. For this purpose, the commonly used Rayleigh damping model in lightly damped systems is employed, which can be

expressed as

$$\mathbf{C} = \mathbf{M}\sum_{k=0}^{r-1}\alpha_k[\mathbf{M}^{-1}\mathbf{K}]^k, \tag{7}$$

where r is the number of damping ratios used to approximate the structural damping effect and coefficients α_k ($k = 0, 1, \ldots, r-1$) are obtained from the r number of simultaneous equations

$$\zeta_i = \frac{1}{2}\left(\frac{a_0}{\omega_i} + a_1\omega_i + \cdots + \alpha_{r-1}\omega_i^{2r-3}\right) = \frac{1}{2}\sum_{k=1}^{r}\alpha_{k-1}\omega_i^{2k-3} \quad (i = 1, 2\ldots, r), \tag{8}$$

where ζ_i denotes the ith modal damping ratio. Expression (7) for the damping matrix, allowing for the orthogonality of mode shapes, provides a very convenient way for incorporation into the calculation of structural dynamic responses.

The damping model represented by Eq. (7) implies that the total damping in the structure is the sum of individual damping in each mode. Thus, the ability to measure values for the damping ratio ζ_i, and hence the damping behavior of the complete structure system, is an important consideration in practice. It is noted that with $r = 2$, Eq. (7) reduces to the Rayleigh damping which is frequently called the "proportional damping". However, the Rayleigh damping model obviously damps the higher modes considerably more than the lower modes. Hence, more damping ratios would be desired to better simulate the true damping behavior of a structural system.

Although the existing modal analysis theory can identify damping parameters based on structural FRF data, the results can be highly susceptible to the measurement errors. In the present study, an ANN-based method is proposed for identifying the damping ratios under the above-mentioned damping model. The reference responses are taken from the integrals of FRFs over a specified small frequency range in the vicinity of the natural resonance frequencies, as shown in Fig. 2(b). In this way, the identified damping ratios can be of better accuracy because the noise components on the FRF curves are somehow "neutralized" through the integration process. It is noted that integrals of FRF curves can also be used as conditioned frequency domain data for structural parameter updating as proposed in a previous work [18,19]. In the present study, the integrals of FRFs are employed in its capacity for updating structural damping ratios.

3.3. Sensitivity analysis for selection of response configuration and evaluation of network performance

In the context of ANN-based FE model updating, the sensitivity analysis can be engaged for two purposes. One is to assist in the determination of a desired response configuration (i.e., the composition of the response vector). The proposed model updating herein uses the modal frequencies and FRF integrals as the response data for the updating of the FE model parameters. Since different antiresonance frequencies and FRF curves can be obtained by placing the actuator and sensor at different locations (d.o.f.s) of the structure, it is possible to have different combinations of the response components within a practical limit of the modal order. Based on the sensitivity analysis, a response configuration that yields a smaller variation of the FE parameters with a given level of perturbation to the response data is considered more desirable.

Another purpose of the sensitivity analysis is to provide a margin of anticipated errors in the FE parameters in case the response data contain a certain level of measurement errors. This estimate can be compared with the actual neural network prediction errors to evaluate the noise-resisting ability of the trained neural network.

As the proposed updating scheme involves two levels of neural networks and each performs a separate task based on different types of response data, separate sensitivity analysis is carried out for the two updating processes. Because the second network proceeds on the basis of the updated structural parameters from the first network, the errors in the structural parameters will result in errors in the computed FRFs to be used for the training of the second network. Such "carry-over" errors should be considered in conjunction with the errors that can be anticipated in the actual measured FRFs for the sensitivity analysis concerning the damping ratios.

3.3.1. Sensitivity analysis concerning structural parameters (first-level network)

In this sensitivity analysis, both natural and antiresonance frequencies are involved. Eq. (3) can be rewritten for the antiresonance eigenproblem as

$$(\mathbf{K}_{pq} - \lambda^a \mathbf{M}_{pq})\{\phi_b\} = 0. \tag{9}$$

Differentiating Eq. (9) with respect to a structure parameter p_j yields

$$(\mathbf{K}_{pq} - \lambda_b^a \mathbf{M}_{pq})\frac{\partial\{\phi_b\}}{\partial p_j} + \left[\frac{\partial \mathbf{K}_{pq}}{\partial p_j} - \lambda_b^a \frac{\partial \mathbf{M}_{pq}}{\partial p_j} - \frac{\partial \lambda_b^a}{\partial p_j}\mathbf{M}_{pq}\right]\{\phi_b\} = 0, \tag{10}$$

where λ_b^a is an antiresonance eigenvalue, p_j denotes a structural parameter, and the subscript pq denotes that row p and column q have been removed from matrices $\mathbf{M}$ and $\mathbf{K}$. Now consider a different but related eigenproblem

$$(\mathbf{K}_{pq} - \lambda_b^a \mathbf{M}_{pq})^{\mathrm{T}}\{\eta_b\} = 0. \tag{11}$$

Since the eigenvalues of a matrix are invariant with respect to the transpose operation, the eigenvalues from Eq. (11) will be the same as those from Eq. (9). However, the eigenvectors $\{\eta_b\}$ will definitely differ from $\{\phi_b\}$. Transposing Eq. (11) gives

$$\{\eta_b\}^{\mathrm{T}}(\mathbf{K}_{pq} - \lambda_b^a \mathbf{M}_{pq}) = 0. \tag{12}$$

The eigenvector $\{\eta_{\mathrm{b}}\}$ is called the *left eigenvector* because it pre-multiplies $(\mathbf{K}_{pq} - \lambda_b^a \mathbf{M}_{pq})$, whereas the standard eigenvector post-multiplies $(\mathbf{K}_{pq} - \lambda_b^a \mathbf{M}_{pq})$. Thus, pre-multiplying Eq. (10) by $\{\eta_b\}^{\mathrm{T}}$ and considering Eq. (12) give the antiresonance sensitivity

$$\frac{\partial \lambda_b^a}{\partial p_j} = \frac{\{\eta_b\}^{\mathrm{T}}((\partial \mathbf{K}_{pq}/\partial p_j) - \lambda_b^a(\partial \mathbf{M}_{pq}/\partial p_j))\{\phi_b\}}{\{\eta_b\}^{\mathrm{T}}\mathbf{M}_{pq}\{\phi_b\}}. \tag{13}$$

Notice that if $\mathbf{M}_{pq}$ and $\mathbf{K}_{pq}$ are equally symmetric matrices, then $\{\eta_b\} = \{\phi_b\}$, and hence Eq. (13) reduces to the symmetric eigenvalue sensitivity problem as in solving for the sensitivity of natural frequencies.

Denote the parameters to be updated in a vector form as $\{p\}$ and the responses used as $\{R\}$ including natural and antiresonance frequencies to a practically affordable order. The relationship between a small perturbation of responses, $\delta\{R\}$, and the corresponding variation of parameters,

$\delta\{p\}$, can be expressed as

$$\delta\{R\} = \mathbf{S}\delta\{p\}, \tag{14}$$

where sensitivity matrix $\mathbf{S}$ consists of two types of sensitivity components, namely natural frequency sensitivity and antiresonance sensitivity as

$$\mathbf{S} = \frac{1}{2\pi}\begin{bmatrix} \frac{\partial\lambda_1}{\partial p_1}(2\omega_1)^{-1} & \frac{\partial\lambda_1}{\partial p_2}(2\omega_1)^{-1} & \cdots & \frac{\partial\lambda_1}{\partial p_m}(2\omega_1)^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\lambda_n}{\partial p_1}(2\omega_n)^{-1} & \frac{\partial\lambda_n}{\partial p_2}(2\omega_n)^{-1} & \ddots & \frac{\partial\lambda_n}{\partial p_m}(2\omega_n)^{-1} \\ & & & \\ \frac{\partial(\lambda_b^a)_1}{\partial p_1}[2(\omega_b^a)_1]^{-1} & \frac{\partial(\lambda_b^a)_1}{\partial p_2}[2(\omega_b^a)_1]^{-1} & & \frac{\partial(\lambda_b^a)_1}{\partial p_m}[2(\omega_b^a)_1]^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial(\lambda_b^a)_l}{\partial p_1}[2(\omega_b^a)_l]^{-1} & \frac{\partial(\lambda_b^a)_l}{\partial p_2}[2(\omega_b^a)_l]^{-1} & & \frac{\partial(\lambda_b^a)_l}{\partial p_m}[2(\omega_b^a)_l]^{-1} \end{bmatrix}, \tag{15}$$

where $(\omega_b^a)_i = \sqrt{(\lambda_b^a)_i}$ is the antiresonance eigen-frequency. Inverting Eq. (14) gives the operational formula

$$\delta\{p\} = \mathbf{S}^{-1}\delta\{R\} = \mathbf{G}\delta\{R\}, \tag{16}$$

which gives the variation of the physical parameters as a result of a small perturbation of the responses. Such sensitivity results can be used in judging the adequacy of a particular response configuration. On the other hand, when the response perturbation is set to represent an anticipated margin of errors in the measured response, the parameter variation as expressed in Eq. (16) will indicate the errors that can be expected in the updated parameters. This error estimation can then be used as a norm to evaluate the performance of the trained neural network.

3.3.2. Sensitivity analysis concerning damping parameters (second-level network)

The second neural network is used to update structural damping ratios using the integrals of FRFs as reference responses. The sensitivity thus refers to the variation of damping ratios to small perturbation of FRF integrals. Rewriting Eq. (8) in matrix form,

$$\{\zeta\}_{r\times 1} = \mathbf{A}_{r\times r}\{\alpha\}_{r\times 1}, \tag{17}$$

where $A_{ij} = \frac{1}{2}\omega_i^{2j-3}$, with ω_i $(i = 1, 2\ldots, r)$ being the circular frequencies. Thus,

$$\{\alpha\}_{r\times 1} = \mathbf{A}_{r\times r}^{-1}\{\zeta\}_{r\times 1}. \tag{18}$$

The individual response components, i.e., the integral of FRFs, I_B, is obtained from

$$\mathbf{I}_B = \int_{\omega_1(\zeta_i)}^{\omega_2(\zeta_i)} \mathbf{B}(\omega, \{\zeta\}, \{p\})\,\mathrm{d}\omega, \tag{19}$$

where $\mathbf{B}$ denotes the dynamic stiffness matrix, $\mathbf{B}(\omega, \{\zeta\}, \{p\}) = (\mathbf{K} - \lambda\mathbf{M} + \mathrm{j}\omega\mathbf{C})^{-1}$, $\{\zeta\}$ is the damping ratio vector, $\{\zeta\} = [\zeta_1, \zeta_2, \ldots, \zeta_r]^{\mathrm{T}}$, and $\{p\}$ is the structural parameter vector, $\{p\} = [p_1, p_2, \ldots, p_r]^{\mathrm{T}}$. The integration bounds ω_1 and ω_2 are determined by a proportion (say 0.9 and 1.1, respectively) of the individual frequency ω_i. Since the influence of small damping on ω_i is negligible, the sensitivity of $\mathbf{I}_B$ with respect to ζ_i can be obtained by

$$\frac{\partial \mathbf{I}_B}{\partial \zeta_i} = \int_{\omega_1}^{\omega_2} \frac{\partial \mathbf{B}(\omega, \{\zeta\}, \{p\})}{\partial \zeta_i} \mathrm{d}\omega. \tag{20}$$

In the above expression, the parameter vector $\{p\}$ has known values from the output of the first network, hence

$$\frac{\partial \mathbf{B}}{\partial \zeta_i} = -\mathbf{B}\left[\mathrm{j}\omega \frac{\partial \mathbf{C}}{\partial \zeta_i}\right]\mathbf{C}, \quad \text{where} \quad \frac{\partial \mathbf{C}}{\partial \zeta_i} = \mathbf{M}\sum_{k=0}^{r-1} \frac{\partial \alpha_k}{\partial \zeta_i}[\mathbf{M}^{-1}\mathbf{K}]^k. \tag{21}$$

Substituting Eq. (21) into Eq. (20),

$$\frac{\partial \mathbf{I}_B}{\partial \zeta_i} = \int_{\omega_1}^{\omega_2} -\mathbf{B}\left(\mathrm{j}\omega\mathbf{M}\sum_{k=0}^{r-1} \frac{\partial \alpha_k}{\partial \zeta_i}[\mathbf{M}^{-1}\mathbf{K}]^k\right)\mathbf{B}\,\mathrm{d}\omega. \tag{22}$$

Hence, the relationship between a small variation of the damping ratios and the corresponding variation of integrals of FRFs can be established via the sensitivity matrix $\mathbf{S}_d$ with entries calculated from Eq. (22)

$$\delta\{I\} = \mathbf{S}_d \delta\{\zeta\}. \tag{23}$$

Inverting Eq. (23)

$$\delta\{\zeta\} = \mathbf{G}_d \delta\{I\}, \tag{24}$$

where the gain matrix $\mathbf{G}_d = \mathbf{S}_d^{-1}$, and the vector $\{I\}$ consists of entries from matrix $\mathbf{I}_B$.

It has to be pointed out the fact that the structural parameters ($\{p\}$) used to compute the FRFs and thereby the integral $\mathbf{I}_B$ for training of the second neural network are the output from the first neural network, and these parameters contain errors in themselves already. Hence, the error vector $\delta\{I\}$ in Eq. (24) should consist of two parts; one due to the above carry-over error ($\delta\{I\}^C$) which takes effect during the neural network training, and another due to the (random) errors in the actual measured FRFs ($\delta\{I\}^M$) when subsequently apply the trained ANN for actual damping updating. Thus,

$$\delta\{I\} = \delta\{I\}^M + \delta\{I\}^C. \tag{25}$$

$\delta\{I\}^C$ can be evaluated based on the parameter error vector $\delta\{p\}$ from Network-1. From Eq. (19),

$$\frac{\partial \mathbf{I}_B}{\partial p_i} = \int_{\omega_1(\{p\})}^{\omega_2(\{p\})} \frac{\partial \mathbf{B}(\omega, \{\zeta\}, \{p\})}{\partial p_i} \mathrm{d}\omega, \tag{26}$$

where

$$\frac{\partial \mathbf{B}(\omega, \{\zeta\}, \{p\})}{\partial p_i} = -\mathbf{B}\frac{\partial[\mathbf{K} - \omega^2\mathbf{M} + \mathrm{j}\omega\mathbf{C}]}{\partial p_i}\mathbf{B}.$$

In actual application, Eq. (26) can be approximated by a finite difference solution

$$\left.\frac{\partial \mathbf{I_B}}{\partial p_i}\right|_{\{p^*\}} \approx \left.\frac{\Delta \mathbf{I}_B}{\Delta p_i}\right|_{\{p^*\}} = \left(\int_{\omega_1(p_i+\Delta p_i)}^{\omega_2(p_i+\Delta p_i)} \mathbf{B}(p_i+\Delta p_i)\,\mathrm{d}\omega - \int_{\omega_1(p_i)}^{\omega_2(p_i)} \mathbf{B}(p_i)\,\mathrm{d}\omega\right)/\Delta p_i, \tag{27}$$

where vector $\{p^*\}$ represents the nominal state of structure parameters at which sensitivity analysis is conducted and Δp_i denotes a small incremental quantity corresponding to a particular structure parameter p_i. Thus, Eq. (27) forms the entries of the sensitivity matrix $\mathbf{S}$. Subsequently, the errors in the FRF integrals induced by the inaccuracy in the parameters $\{p\}$ can be computed in a forward manner

$$\delta\{I\}^C = \mathbf{S}\delta\{p\}. \tag{28}$$

4. Design of neural networks and training

An ANN-based model updating scheme generally consists of the following stages: (1) generating the training data, (2) training the neural network, (3) testing the neural network. The trained neural network can subsequently be used to update the FE model parameters when fed with measured response data from the actual structure.

The preparation of proper training data plays a crucial role for a successful updating since their quality directly affects the "expressing power" (the generalization capacity) of the network. The training data actually comprises a number of paired vectors (called "training pairs"), and each pair includes a parameter vector (network output) and the corresponding response vector (network input). Before the generation of the training pairs, a proper response configuration must be determined. This can be carried out using the aforementioned sensitivity analysis scheme. It has to be pointed out the fact that the results of a sensitivity analysis are dependent upon the chosen nominal state around which the perturbation takes place. Therefore, in cases where the intended coverage of the structure states is wide as presumed in the present study, the sensitivity analysis concerning the response configuration selection should be conducted on a bunch of randomly selected structural states, and the final judgment should be made based on the collective trend from the sensitivity results.

Once the response configuration is decided, the generation of the training pairs becomes straightforward: For n number of randomly sampled parameter vectors within the targeted variation range, $\{p\}_j$ $(j = 1, 2, \ldots, n)$, n number of response vectors, $\{R\}_j$ $(j = 1, 2, \ldots, n)$, are calculated by a forward FE analysis. Thus, n number of training pairs (or "training patterns") are readily available for the training of the neural network.

The next decision is on the network topology, i.e., the number of hidden layers and the number of neurons (processing units) in each hidden layer. Note that the number of neurons in the input and output layers are already determined upon the selection of the response and structural parameter configuration. Since in this study the neural network is chosen to have two hidden layers with equal number of neurons in each hidden layer, only a desired number of neurons needs to be decided. This can be achieved by performing a trial procedure. After the number of neurons is determined, the Levenberg–Marquardt training algorithm is employed to train the network. The termination of the training is decided based on the cross-validation output.

For the testing of the generalization ability during the network training, a group of data patterns (called "test patterns") not included in the training patterns is utilized for the purpose to

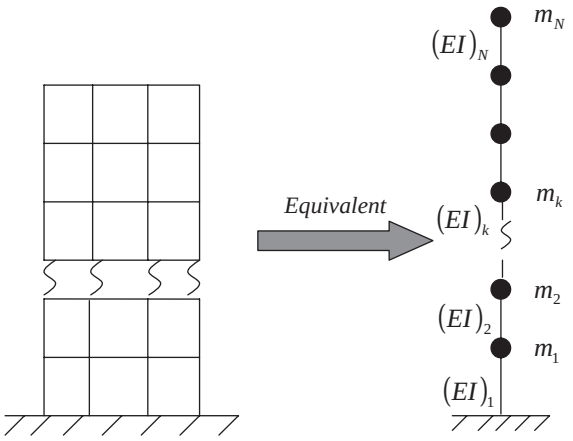

Fig. 3. Numerical example of a multi-storey building frame.

assess the network performance throughout the training process. A network is regarded as a good candidate if it could result in small errors not only on the training data but also on the test data.

5. Numerical example

A numerical example is given to illustrate the implementation of the above-described procedure. For demonstrative purpose, a simple multi-storey building frame is chosen. Assuming a pure-sway deformation mode (rigid floor), the frame can be simplified as a lumped-mass stick FE model shown in Fig. 3, allowing only sway d.o.f.s. All the effective masses are lumped to the nodal points at the floor levels. The inter-storey sway stiffness of the frame at all individual storeys is considered as the structural parameters to be updated, along with the damping ratios for the first four natural modes.

5.1. Updating of the structural parameters (first-level network)

For the purpose of better network conditioning, in the first network the stiffness variables are represented by stiffness modification factors (SMFs) defined as the absolute FE model stiffness normalized with respect to a reference-state stiffness (e.g., the uncracked stiffness for reinforced concrete members), such that the range of the SMFs falls within (0,1). A SMF close to zero represents a total-damage state, while a SMF equal to one indicates no damage. In the present example, the target SMF variation range is set to (0.35, 1.0) to cover from the undamaged to a heavily damaged state. For a six-storey frame considered in this example, six SMFs are subject to updating. Table 1 lists the basic properties of the frame.

5.1.1. Selection of response configuration

Several combinations of natural frequencies and antiresonance frequencies are configured for selection. For practicality concerns, the response components are restricted within the lowest few modes. Table 2 lists eight such response configurations, where a higher designation number roughly represents a higher demand on the measured data.

Table 1
Physical properties of a six-storey RC frame used in the numerical example

Storey	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6
Height (m)	5	3	3	3	3	3
Column cross-section (m^2)	0.6×0.6	0.5×0.5	0.45×0.45	0.45×0.45	0.35×0.35	0.35×0.35
Lumped mass (kg)	47,500	42,000	40,500	39,500	37,000	35,500
Initial modulus	$E = 26$ GPa (for a typical class of concrete)					

Table 2
Eight candidate response configurations for Network-1

Configuration no.	1	2	3	4	5	6	7	8
Number of natural frequencies	1	2	3	4	4	4	4	4
Number of antiresonance frequencies[a]	5	4	3	2	4	2,2	3,3	4,4

[a] Single numbers refer to antiresonance frequencies from FRF(1,1), while double numbers separated by a comma refer to antiresonance frequencies, respectively, obtained from FRF(1,1) and FRF(2,2). FRF(i,j) denotes the frequency response at d.o.f. i due to a unit impulse applied at d.o.f. j.

As the targeted variation range of the structural states is broad (SMFs within 0.35–1.0), a total of 20 nominal states of the structure are chosen in a random manner within the above range. Sensitivity analysis is then conducted with respect to each of these nominal models for all the candidate response configurations using an automated procedure. The basic routine is as follows. For any particular nominal state defined by a set of SMFs, an FE analysis is first performed to compute the required modal response, from which the response data for the ith response configuration are extracted. A perturbation is imposed on the extracted response data with a perturbation factor assumed to follow a Gaussian distribution with zero mean and 1% variance (note that other distribution patterns, if deemed appropriate, can be applied in a similar way). Thus, by random sampling a sufficient number (say 1000 as in this example) of the response perturbation vectors, $\delta\{R\}$, are generated. Subsequently, the corresponding parameter variation (error) vectors, $\delta\{p\}$, are calculated according to Eq. (16). These parameter error vectors are then examined on a statistical basis and the MSE is used to represent the overall sensitivity of the parameters for the particular response configuration. Besides the MSE, the detailed distribution of the error for individual parameters can also be examined to further assess the adequacy of the response configuration. Fig. 4 summarizes the sensitivity analysis results in the form of scatter plot of the MSEs of the structural parameters against all the response configurations being considered.

It can be observed that four response configurations, namely $\{R\}^{(5)}$, $\{R\}^{(6)}$, $\{R\}^{(7)}$ and $\{R\}^{(8)}$, exhibit a quite stable error margin in terms of the overall MSE. Further inspection of the error distributions of the individual SMFs, as shown in Fig. 5 for one particular nominal state, reveals that the error band of SMFs reduces from above $\pm 10\%$ to within $\pm 6\%$ when $\{R\}^{(7)}$ is used instead of $\{R\}^{(5)}$. Slight further improvement of the results can be achieved when using $\{R\}^{(8)}$, which however includes two more response components. Therefore, the response configuration $\{R\}^{(7)}$ is chosen for the subsequent preparation of the training pairs and the neural network training.

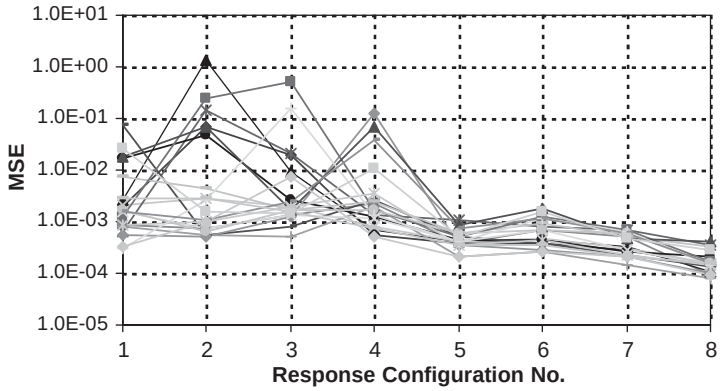

Fig. 4. MSE of SMFs at various nominal states versus different response configurations from sensitivity analysis.

5.1.2. Training of the neural network

After selected the response configuration, standard FE calculations can proceed to generate the parameter-response data pairs for the network training. The required number of the training data pairs is determined based on two factors; the dimension of the input layer (response vector) and the value ranges of the physical parameters (the "SMFs" herein) within which the network is expected to generalize. Now that the response vector has been selected as $\{R\}^{(7)}$ which contains 10 components, the dimension of the network input layer is thus 10. On the other hand, the variation range of the structural parameters, i.e., SMFs, is targeted to cover 0.35–1.0. Thus, according to Vapnik–Chervonenkis dimension theory [20], it is found that 1500 training pairs are appropriate for the network training. The numbers of testing and cross-validation data patterns are subsequently chosen to be 800 and 600, respectively. All these required data patterns are then generated from the FE analysis, and they are used as "noiseless" data patterns for the network training.

Fig. 6(a) illustrates the network performance in terms of the overall MSE on the training and testing data for 10 different networks with different number of neurons in their hidden layers. It can be seen that the MSE generally decreases with increasing number of neurons in the hidden layers (i.e., larger networks). As expected, all networks perform better on the training patterns than on the testing patterns. In conjunction with an analysis of the percentage error distributions of the individual SMFs from the network output for the test patterns, it is found that the neural network with 12 neurons in each of the two hidden layers can already achieve an error band within $\pm 3\%$ for all individual SMFs, given noiseless input data. Therefore, this network topology is selected.

During the network training process, the development of the network performance is continuously monitored. Fig. 6(b) depicts the progressive network performance in terms of the MSE during the training iterations. As can be seen, the training procedure stops at epoch 35 since there is no further improvement of the network performance on the cross-validation data patterns in nearly 10 epochs (from epoch 25 to 35).

The above neural network training is based on the noiseless data patterns directly obtained from the FE analysis. The so-trained neural network is designated as Network-1A. To produce a noise-resisting neural network, a noise-injection learning strategy is implemented to repeat the

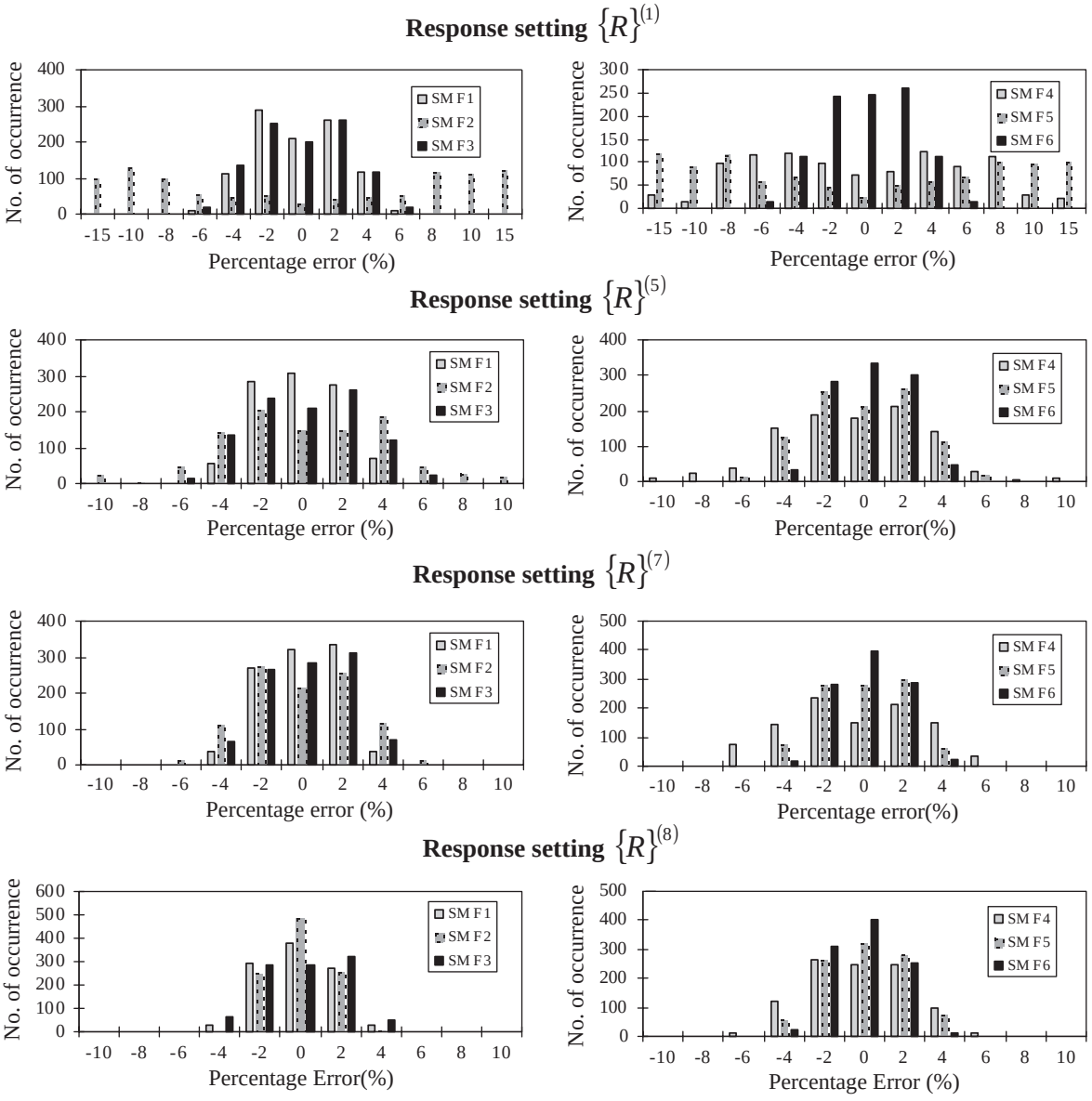

Fig. 5. Typical parameter error distributions corresponding to different response configurations based on sensitivity analysis.

training procedure with the same network topology setting. In this process, however, the 1500 training data patterns are treated by injecting noise (representing the measurement errors) to the response side (network input) according to Eq. (2). The "noise" component, v_j, is obtained by random sampling from a prescribed probability distribution. Since only frequency data are used here, it is assumed that the error in the measured frequency data follows a Gaussian distribution with a zero mean and 1% variance. The output side of the training data, i.e., the SMFs, remains

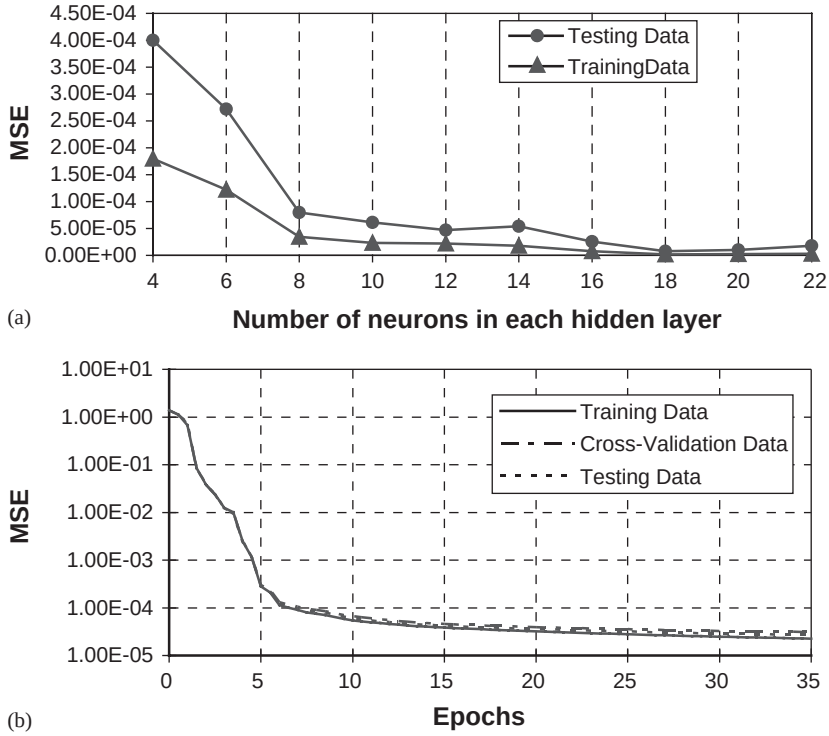

Fig. 6. Performance of neural networks with different topology settings. (a) Variation of MSE of updated parameters with different number of hidden units. (b) Typical network performance curves with ongoing training process.

Table 3
Four model scenarios used for performance comparison between Network-1A (noiseless data trained) and Network-1B (noise-injection trained)

Case no.	SMF					
	Storey 1	Storey 2	Storey 3	Storey 4	Storey 5	Storey 6
1	0.734	0.867	0.762	0.699	0.786	0.767
2	0.821	0.812	0.856	0.733	0.902	0.881
3	0.781	0.947	0.780	0.718	0.798	0.948
4	0.761	0.897	0.962	0.863	0.866	0.751

unchanged and the original pairing is also retained. The network trained using the above noise-injected training patterns is designated as Network-1B.

To evaluate the performance of the above two networks (Network-1A and Network-1B) under a noisy measurement data environment, four arbitrary structure states are subjected to updating using these two networks. Table 3 lists the values of the SMFs for the four structure states. For each state, the exact response data are first calculated using the FE analysis, and they are then "polluted" by adding random error to generate 1000 sets of noisy measurement responses. These noisy response data are then fed into the two networks one by one to perform the updating. Fig. 7

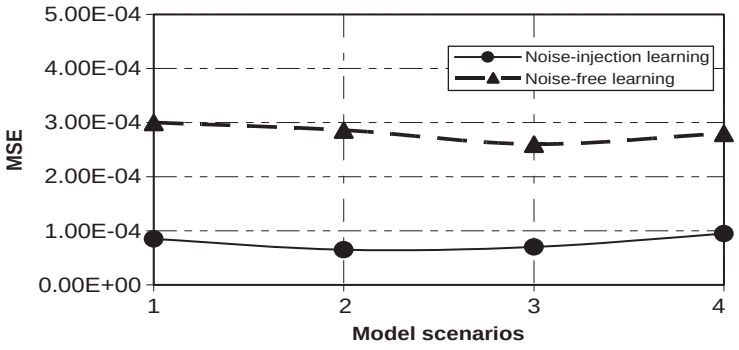

Fig. 7. Comparison of MSE of updated parameters from Network-1A (noiseless data trained) and Network-1B (noise-injection trained) when fed with noisy measurement data.

compares the overall MSE of the output parameters from the two networks. The distributions of the percentage error for the individual parameters (SMFs) from the two networks are compared in Fig. 8. As can be clearly observed, Network-1B exhibits a significantly improved performance under the simulated noisy measurement data. The maximum percentage error for the individual parameters is reduced to within $\pm 4\%$ as compared to $\pm 8\%$ from Network-1A. The latter is comparable with the prediction of the parameter errors from the sensitivity analysis.

5.2. Updating of structural damping ratios (second-level network)

In this example, four damping ratios, ζ_i ($i = 1, 2, \ldots, 4$), corresponding respectively to the lowest four natural modes, are considered for updating.

5.2.1. Selection of response configuration

The responses used for the damping updating are the integrals of FRFs around the natural frequencies (see Fig. 2(b)). For a multi-d.o.f. system, it is possible to measure a number of FRFs, so there could be different response configurations to consider for the sake of a better network performance. The identification of a desired configuration beforehand also helps the planning of the test programme in real applications. For an illustrative purpose, eight candidate response configurations, namely $\{I\}^{(1)}, \{I\}^{(2)}, \ldots, \{I\}^{(8)}$ as shown in Table 4, are subjected to the sensitivity analysis in a way similar to that described in Network-1. Note that in Table 4 the number in parentheses, (i,j), denotes the FRF with force (actuator) applied along the d.o.f. j and the response (sensor) measured at the d.o.f. i. The four lowest natural modes are considered, thus with three arbitrary FRFs each response vector will consist of 12 components (4 modes $\times$ 3 FRFs). Each integral is obtained by integrating the FRF over a frequency interval from $0.9f_k$ to $1.1f_k$, with f_k being the kth natural frequency up to the fourth mode.

A sensitivity analysis is conducted to calculate the variation of the damping parameters, $\delta\{\zeta\}^{(i)}$, corresponding to a perturbation on the FRF integrals included in a particular configuration $\delta\{I\}^{(i)}$. As mentioned in Section 3.3.2, the perturbation on FRFs is considered to represent two parts of errors, one is the so-called carry-over error that is associated with the errors from the first-level neural network, $\delta\{I\}^{(i)C}$, and the other represents the errors in the actual measured FRF

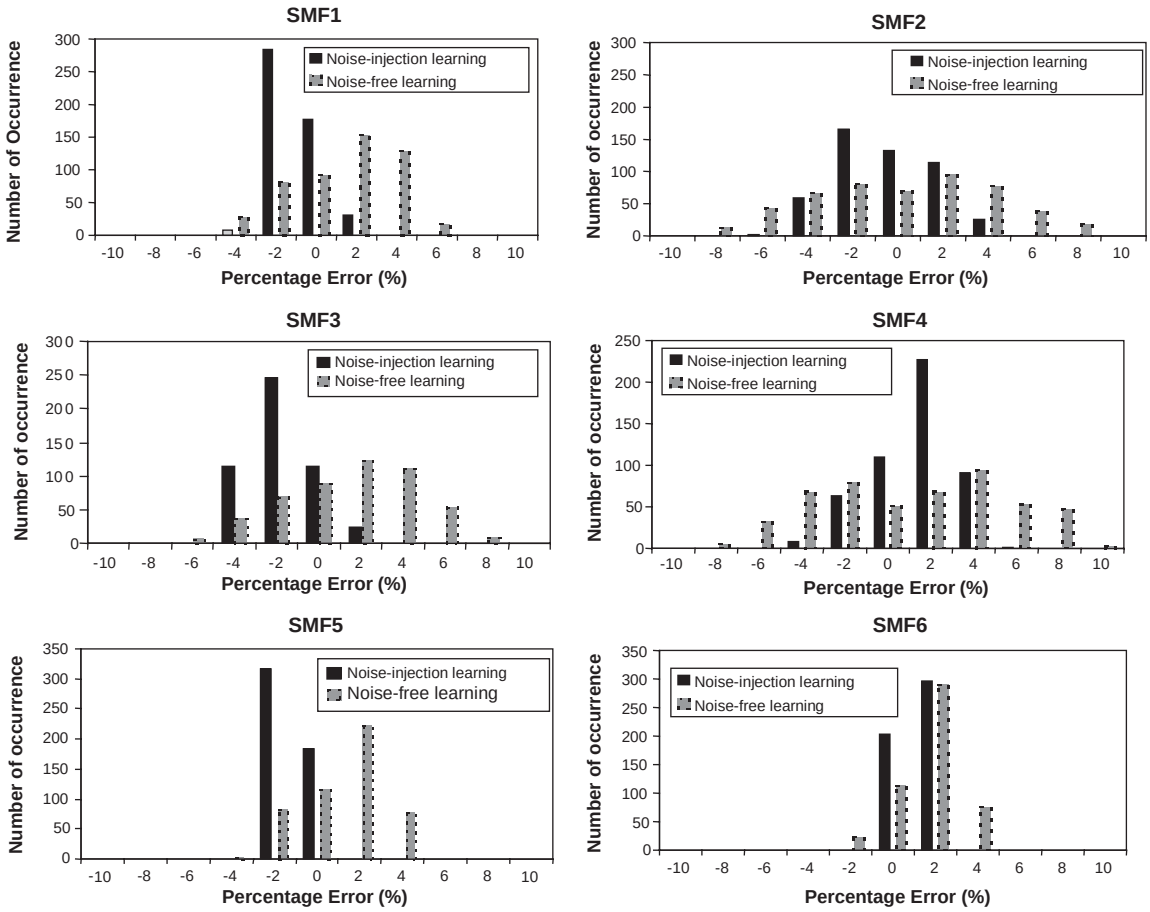

Fig. 8. Comparison of percentage errors in the updated parameters from Network-1A and Network-1B.

Table 4
Eight candidate response configurations for Network-2

Configuration no.	1	2	3	4	5	6	7	8
FRF components	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)	(1,1)
	(2,1)	(2,1)	(2,1)	(3,1)	(3,1)	(4,1)	(4,1)	(5,1)
	(3,1)	(5,1)	(6,1)	(5,1)	(6,1)	(5,1)	(6,1)	(6,1)

integrals, $\delta\{I\}^{(i)M}$, which is assumed to also follow a Gaussan noise with zero mean and 1% variance. The overall error vector, $\delta\{I\}^{(i)}$, is obtained as the sum of $\delta\{I\}^{(i)M}$ and $\delta\{I\}^{(i)C}$. Subsequently, the error in the damping ratios, $\delta\{\zeta\}^{(i)}$, can be calculated according to Eq. (24). Fig. 9(a) depicts the overall MSE in $\delta\{\zeta\}^{(i)}$ from the above sensitivity analysis for 20 arbitrary chosen nominal states of the structure, for which both the SMFs and the damping ratios are

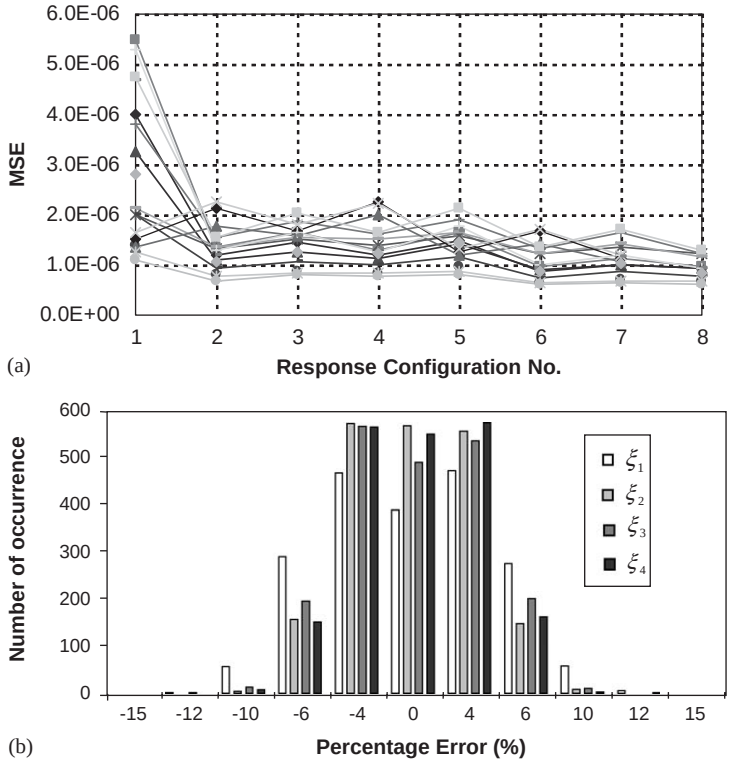

Fig. 9. Predicted errors of damping ratios for different response configurations based on sensitivity analysis. (a) MSE of damping ratios with different response configurations. (b) Error distributions of damping ratios from response setting $\{I\}^{(8)}$.

chosen in an arbitrary manner within their respective range. It can be seen that the 8th response configuration exhibits the best outcome and so it is selected for the Network 2 training. The distribution of the errors corresponding to one particular state of the structure is illustrated in Fig. 9(b), which indicates that the maximum anticipated damping updating error is about $\pm 10\%$ with the assumed margin of error in the measured FRF integrals.

5.2.2. *Training of the neural network*

The number of training pairs for Network-2 is chosen to be 900 according to the Vapnik–Chervonenkis dimensional analysis. The numbers of testing and cross-validation patterns are chosen to be 600 and 500, respectively. After a trial procedure, it is found that a neural network with 9 neurons in each of the two hidden layers is adequate.

The training data as well as the testing and cross-validation data are generated so that a damping variation range from 0.5% to 10% is covered. The neural network, designated as Network-2B, is then trained following the noise-injection learning strategy in a similar way as for the previous Network-1B.

To examine the generalization and noise-resisting ability of the above trained network, an arbitrary model with assumed parameters shown in Table 5 are subjected to updating. For 1000

Table 5
An example case and corresponding updating results from the two-level networks

	Stiffness parameters (SMFs) (p_i)						Damping ratios (ζ_i)			
	p_1	p_2	p_3	p_4	p_5	p_6	ζ_1	ζ_2	ζ_3	ζ_4
Actual parameters	0.65	0.78	0.80	0.59	0.72	0.70	0.01	0.014	0.032	0.056
Updating results	0.65	0.772	0.813	0.578	0.714	0.689	0.0098	0.0142	0.0330	0.0545
Updating errors (%)	2.3	−1.0	1.7	−2.0	−0.8	−1.6	−2.5	1.6	3.3	−2.6

sets of error-polluted FRF integral data ($\{I\}_j^8$, $j = 1, \ldots, 1000$), considering an Gaussian error distribution with zero mean and 1% variance, Fig. 10 shows the error distributions in the updated 4 damping ratios as compared with the theoretically predicted error from the sensitivity analysis. Once again, a remarkable reduction of the error margin is achieved by implementing the noise-injection learning strategy in the neural network training. A summary of the updated damping ratios for ζ_1 to ζ_4, along with the updated structural stiffness parameters using Network-1B, is given in Table 5. Of course, it should be mentioned again that the realization of the noise resisting capacity of a neural network in real applications is subject to the adequacy of the noise model used in the neural network training in representing the actual noise or error in the measurement data. Inconsistency in the noise description could otherwise contribute to erroneous updating results.

6. Conclusions

In this paper, a comprehensive procedure is presented to use ANN for FE model updating including both structural and damping parameters. The methodology involves a series of sensitivity analyses for the selection of a desired response configuration as well as for the evaluation of the post-trained network performance; it also covers the selection of the network topology through trial training and the incorporation of a noise-injection learning strategy. A two-level neural network FE model updating scheme is developed so that the structural parameters (SMFs) and the damping ratios (ζ_i) can be updated using two separate networks. A numerical example is given to demonstrate the effectiveness and efficiency of the proposed method.

The two-neural network scheme proves to work out successfully in updating the structural parameters and the damping ratios. Apart from the satisfactory accuracy that can be achieved with noiseless response data, the numerical example also shows that the noise-injection learning strategy can result in a neural network with substantially enhanced noise-resisting ability. In the example shown, the error margin in the updated structural parameters and damping ratios from the noise-injection learning is reduced by more than 50% as compared to the results from the noiseless-data trained neural work when fed with noisy data. For a given level of noise with zero mean and 1% variance in the frequency response data, the neural network is able to identify the structural parameters as well as the damping ratios within an error of $\pm 4\%$.

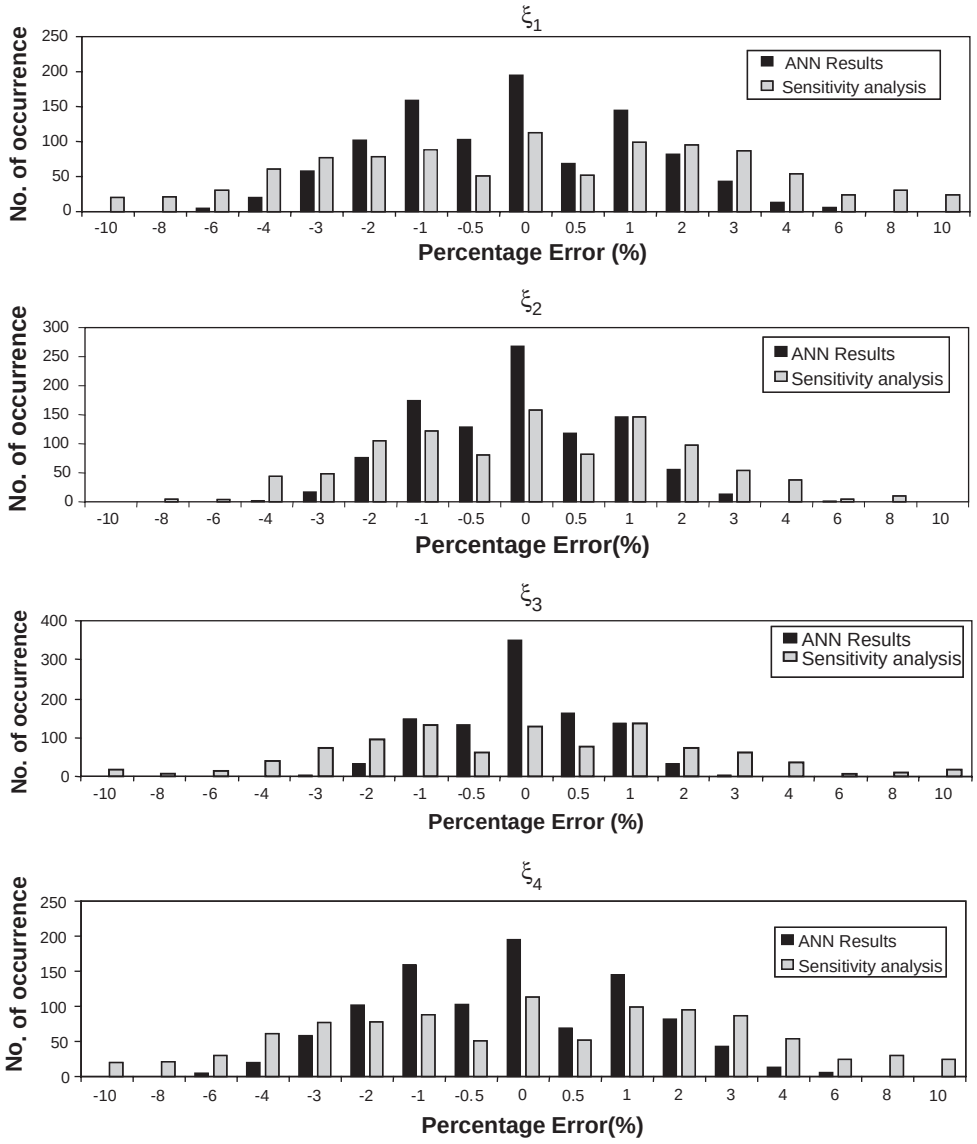

Fig. 10. Comparison of percentage errors in the updated damping ratios from Network-2B with the predicted errors from sensitivity analysis.

It should be pointed out that the realization of the noise-resisting ability in real applications is subject to the adequacy of the noise model used in the network training in describing the actual noise in the measured data. The applicability of the updated damping ratios using the proposed damping model also depends on the adequacy of such damping model in representing the actual damping mechanism of the system under consideration. Should other noise models or damping models be deemed more adequate for a particular structure, they can be implemented in the neural network training in a similar way as described in this paper.

References

[1] M.I. Friswell, J.E. Mottershead, *Finite Element Model Updating in Structural Dynamics*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.
[2] J.E. Mottershead, M.I. Friswell, Model updating in structural dynamics: a survey, *Journal of Sound and Vibration* 167 (2) (1993) 347–375.
[3] J.E. Mottershead, C. Mares, Selection and updating of parameters for an aluminium space-frame model, *Mechanical Systems and Signal Processing* 14 (6) (2000) 923–944.
[4] K.W. Jones, Finite Element Model Updating using Antiresonance Frequencies, MS Thesis, AFIT/GA/ENY/00-M08, Air Force Institute of Technology (AU), Wright-Patterson AFB, Dayton, OH, 2000.
[5] S. Wu, J. Ghaboussi, J.H. Garrett Jr., Use of neural networks in detection of structural damage, *Computers & Structures* 42 (4) (1992) 649–659.
[6] P. Szewezyk, P. Hajela, Damage detection in structures based on feature-sensitive neural networks, *Journal of Computers in Civil Engineering, Transactions of American Society of Civil Engineering* 8 (2) (1994) 163–178.
[7] S.F. Masri, M. Nakamura, A.G. Chassiakos, T.K. Caughey, Neural network approach to detection of changes in structural parameter, *Journal of Engineering Mechanics* 122 (4) (1996) 350–360.
[8] C.B. Yun, J.H. Yi, E.Y. Bahng, Joint damage assessment of framed structures using a neural networks technique, *Engineering Structures* 23 (2001) 425–435.
[9] S.A. Harp, T. Samad, A. Guha, Designing application-specific neural networks using the genetic algorithm, *Proceedings of the IEEE Conference on Neural Information Processing Systems*, Denver, CO, 1989, Vol. 2, Morgan Kaufmann, San Mateo, CA, 1990, pp. 447–454.
[10] N. Funabiki, J. Kitamichi, S. Nishikawa, Evolutionary neural network algorithm for max cut problems, *Proceedings of the 1007 IEEE International Conference on Neural Networks*, Part 2, IEEE Press, Piscataway, NJ, USA, 1997, pp. 1260–1265.
[11] K.-I. Funahashi, On the approximate realization of continuous mappings by neural networks, *Neural Networks* 2 (3) (1989) 183–192.
[12] E.J. Hartman, J.D. Keeler, J.M. Kowalski, Layered neural networks with Gaussian hidden units as universal approximations, *Neural Computation* 2 (2) (1990) 210–215.
[13] M. Peyt, *Introduction to Finite Element Analysis*, Cambridge University Press, New York, 1990.
[14] G. Lallement, S. Cogan, Reconciliation between measured and calculated dynamic behaviors: enlargement of knowledge space, *Proceedings of the 10th IMAC Conference*, San Diego, CA, 1992, pp. 487–493.
[15] J.E. Mottershead, On the zeros of structural frequency response functions and their application to model assessment and updating, *Proceedings of the 16th International Modal Analysis Conference*, Santa Barbara, CA, 1998, pp. 500–503.
[16] J. He, Y.-Q. Li, Relocation of anti-resonances of a vibratory system by local structural changes, *The International Journal of Analytical and Experimental Modal Analysis* 10 (4) (1995) 224–235.
[17] K-J. Bathe, *Finite Element Procedures*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1996.
[18] M.J. Atalla, Model Updating Using Neural Networks, Ph.D. Dissertation, Virginia Polytechnic Institute and State University, 1996.
[19] M.J. Atalla, D.J. Inman, On model updating using neural networks, *Mechanical Systems and Signal Processing* 12 (1) (1998) 135–161.
[20] V.N. Vapnik, A.Y.A. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory of Probability and its Applications* 16 (2) (1971) 264–280.