JOURNAL OF AIRCRAFT
Vol. 38, No. 1, January–February 2001

Probabilistic Fault Identification Using a Committee of Neural Networks and Vibration Data

Tshilidzi Marwala*
University of Cambridge, Cambridge, England CB2 1PZ, United Kingdom

Bayesian-formulated neural network architecture is implemented using a hybrid Monte Carlo method for probabilistic fault identification in a population of ten nominally identical cylindrical shells using vibration data. Each cylinder is divided into three substructures. Holes of 12 mm in diameter are introduced in each of the substructures. Vibration data are measured by impacting the cylinders at selected positions using a modal hammer and measuring the acceleration responses at a fixed position. Modal energies, defined as the integrals of the real and imaginary components of the frequency response function over 12-Hz frequency bandwidths, are extracted and transformed into the coordinate modal energy assurance criterion. This criterion and the identity of faults are used to train the frequency response function (FRF) neural network. Modal analysis is then employed to identify modal properties. Mode shapes are transformed into the coordinate modal assurance criterion. The natural frequencies and the coordinate modal assurance criterion, as well as the identities of faults, are utilized to train the modal-property neural network. The weighted average of the modal-property network and the FRF network form a committee of two networks. The committee approach is observed to give lower mean square errors and standard deviations (thus, a higher probability of giving the correct solution) than the individual methods. This approach gives accurate identities of damage and their respective confidence intervals while requiring affordable computational resources.

Nomenclature

$[C]$	=	damping matrix
D	=	matrix containing identity of damage data
E_D, E_W	=	error functions
$F(w)$	=	integrand
$f(\cdot)$, $\{f\}$	=	activation function of $\cdot$ and force vector, respectively
$h(\cdot)$	=	mapping function of $\cdot$
j	=	$\sqrt{(-1)}$
$[K]$	=	stiffness matrix
$[M]$	=	mass matrix
N	=	identified modal energy
$p(D \mid w)$	=	likelihood function
$p(w \mid D)$	=	posterior probability distribution
$p(w \mid x)$	=	probability distribution function of the weight space
$p(y \mid x, w)$	=	distribution of the noise on y
T	=	temperature in simulated annealing
W	=	number of weights
w_{j0}, w_{k0}	=	bias parameters
$\{X\}$	=	displacement vector
$\{x\}$, $[x]$	=	vector and matrix, respectively, of modal properties or modal energies
y, y_i	=	identity of damage vector (ith vector)
α	=	coefficient of prior probability distribution
β	=	hyperparameter
η	=	data from modal energies
$\{\phi_n\}$, $\{\overline{\phi_n}\}$	=	nth natural mode shape vector, complex mode shape vector
χ	=	modal properties
$\bar{\omega}_n$, ω	=	nth complex eigenvalue and frequency, respectively

Subscripts

i, j	=	indices
m	=	measured
med	=	median

Superscripts

N	=	number of modes
M	=	number of hidden layers
$\star$	=	the complex conjugate

Received 5 August 1999; revision received 9 May 2000; accepted for publication 12 May 2000.

*Bradlow Foundation Senior Research Scholar, Department of Engineering, St. John's College, Trumpington Street.

I. Introduction

THE identification of faults in aerospace structures at the manufacturing stage offers substantial economic benefits. Vibration methods have been implemented with varying degrees of success on identifying mechanical faults.[1] These techniques can be broadly classified as being experimentally based or model based.

Experimental methods use experimental data as a basis of fault identification. These methods use changes in vibration data with little or no assumptions about the analytical behavior of the structure for fault identification. The main shortcoming with these methods is that they are usually insensitive to faults of small magnitude and experience difficulty on quantifying the severity of faults. The advantage of these methods is that they are not computationally intensive. Model-based techniques modify a numerical model (such as finite element model) to match measured vibration data as a basis for fault identification.[2] These methods are, in principle, capable of identifying faults, but they rely on the accuracy of numerical models.

In this study, the committee[3] of neural networks, which employs frequency response functions (FRFs) and modal properties simultaneously is extended to a probabilistic framework and is experimentally validated. Two Bayesian-formulated neural networks, trained using FRFs and modal properties, are weight averaged and used to identify faults in a population of cylindrical shells. The committee approach has been found to give more reliable solutions than the two individual methods.[3] The Bayesian approach is applied because it is easier to determine the confidence intervals of the identity of faults than the maximum likelihood approach.[4] It also automatically penalizes highly complex models and, therefore, is able to select an optimal model without applying independent methods such as cross validation, as is the case for the maximum likelihood approach.

The implementation of neural networks may be classified as a nondeterministic optimization problem. This is because when neural networks are implemented, only the data are required instead of deterministic mathematical relations. The optimization nature of neural networks causes a problem of not finding a global optimum solution, especially if the number of parameters that indicate the

identity of faults is high. To avoid the high incidence of finding a local optimum solution, it is often desirable to reduce the number of design variables. A method implemented to achieve this objective is the method of substructuring

Each of the 10 cylinders used in this study is divided into three substructures. Faults are located within these three substructures. The parameters corresponding to each substructure form a vector space, also known as the identity of fault, and this information is defined as the substructure space. The information from the FRFs and modal properties are transformed into substructure space using the weighted average of the two independent neural networks. This approach performs fault identification by using changes in vibration data resulting from the presence of faults, despite the presence of other changes such as those due to uncertainties in measured data because of a variation in physical properties of a population of cylinders, uncertain measurement positions, changes in support conditions, etc.

II. Theoretical Formulation

In this section, substructuring of modal and frequency equations is introduced. Any elastic structure may be expressed in terms of mass, damping, and stiffness matrices in time domain by

$$[M]\{x(t)\} + [C]\{x(t)\} + [K]\{x(t)\} = \{f(t)\} \tag{1}$$

Equation (1) can be transformed into frequency domain to give

$$\{-\omega^2[M] + i\omega[C] + [K]\}\{x(\omega)\} = \{f(\omega)\} \tag{2}$$

A. FRF Substructuring

Suppose we are interested in locating damage in either substructure 1, 2, or 3. Then Eq. (2) may be partitioned into three superelements, as follows:

$$\left(-\omega^2\begin{bmatrix}[M_{11}] & [M_{12}] & [M_{13}]\\ [M_{21}] & [M_{22}] & [M_{23}]\\ [M_{31}] & [M_{32}] & [M_{33}]\end{bmatrix} + j\omega\begin{bmatrix}[C_{11}] & [C_{12}] & [C_{13}]\\ [C_{21}] & [C_{22}] & [C_{23}]\\ [C_{31}] & [C_{32}] & [C_{33}]\end{bmatrix} + \begin{bmatrix}[K_{11}] & [K_{12}] & [K_{13}]\\ [K_{21}] & [K_{22}] & [K_{23}]\\ [K_{31}] & [K_{32}] & [K_{33}]\end{bmatrix}\right)\begin{Bmatrix}x_1\\ x_2\\ x_3\end{Bmatrix} = \begin{Bmatrix}f_1\\ f_2\\ f_3\end{Bmatrix} \tag{3}$$

Because displacements $\{x\}$ and force $\{f\}$ are not used directly, the force is assumed white, hence, $\{f\}$ has a unit force magnitude at all frequencies, and the displacement is replaced by the FRFs. If any of the substructures has a fault, this would be reflected by changes in FRFs of the three substructures. By comparing the relative changes of the FRFs of these three substructures due to faults, one may be able to identify faults.

The implicit relationship between physical properties of the structure, for example, mass and stiffness matrices, and the FRFs will be used to identify faults in structures. Sufficient data of the modal energies extracted from FRFs η and their corresponding identities of fault y_1 will be obtained from experiment, and a functional mapping between the two will be quantified using a neural network by the following equation:

$$y_1 = h(\eta) \tag{4}$$

B. Modal Property Substructuring

Equation (2) may also be transformed into modal domain to form an eigenvalue equation for the nth mode, which may be written in the substructure domain as follows:

$$\left(\bar{\omega}_n^2\begin{bmatrix}[M_{11}] & [M_{12}] & [M_{13}]\\ [M_{21}] & [M_{22}] & [M_{23}]\\ [M_{31}] & [M_{32}] & [M_{33}]\end{bmatrix} + \bar{\omega}_n\begin{bmatrix}[C_{11}] & [C_{12}] & [C_{13}]\\ [C_{21}] & [C_{22}] & [C_{23}]\\ [C_{31}] & [C_{32}] & [C_{33}]\end{bmatrix} + \begin{bmatrix}[K_{11}] & [K_{12}] & [K_{13}]\\ [K_{21}] & [K_{22}] & [K_{23}]\\ [K_{31}] & [K_{32}] & [K_{33}]\end{bmatrix}\right)\begin{Bmatrix}\overline{\phi_{n1}}\\ \overline{\phi_{n2}}\\ \overline{\phi_{n3}}\end{Bmatrix} = \begin{Bmatrix}0\\ 0\\ 0\end{Bmatrix} \tag{5}$$

If any of the substructures has a fault, this would be reflected by changes in $\{\overline{\omega_n}\}$ and $\{\overline{\phi_n}\}$ of the three substructures. By comparing the relative changes of the modal properties of these three substructures as a result of fault, one may be able to deduce the presence and the location of faults. Similarly, a functional mapping between the identity of fault y_2 and the modal properties χ may be quantified by the following equation:

$$y_2 = f(\chi) \tag{6}$$

In this section it is demonstrated how the method of substructuring can be used to reduce the order of the problem of fault identification from several thousand substructures to three (in this study three was chosen arbitrarily). Substructuring may be applied as a first step in fault diagnostics by pointing to a larger area before localized methods such as acoustics methods are applied.

C. Neural Networks

In this study, neural networks are viewed as parameterized graphs that make probabilistic assumptions about the data. Learning algorithms are viewed as methods for finding parameter values that look probable in light of the data. Learning processes occur by training the network through either supervised or unsupervised learning. Unsupervised learning is used when only the input data are available. Supervised learning is used when the input and the output are available and neural networks are used to approximate the functional mapping between the two. In this study, supervised learning is applied.

There are several types of neural network architectures, namely, multilayer perception (MLP) and radial basis function.[4] In this study, the MLP is chosen because it provides a complex nonlinear mapping between the input and the output.

A schematic illustration of the MLP is given in Fig. 1. This network architecture contains hidden units and output units. The bias parameters in the first layer are shown as weights from an extra input having a fixed value of $x_0 = 1$. The bias parameters in the second layer are shown as weights from an extra hidden unit, with activation fixed at $z_0 = 1$. The model in Fig. 1 is able to take into account the intrinsic dimensionality of the data. Models of this form can approximate any continuous function to an arbitrary accuracy if the number of hidden units M is sufficiently large. Considering several layers expands the MLP.

In this study, the output units represent the identity of damage whereas the input units represent the parameters from the FRFs or

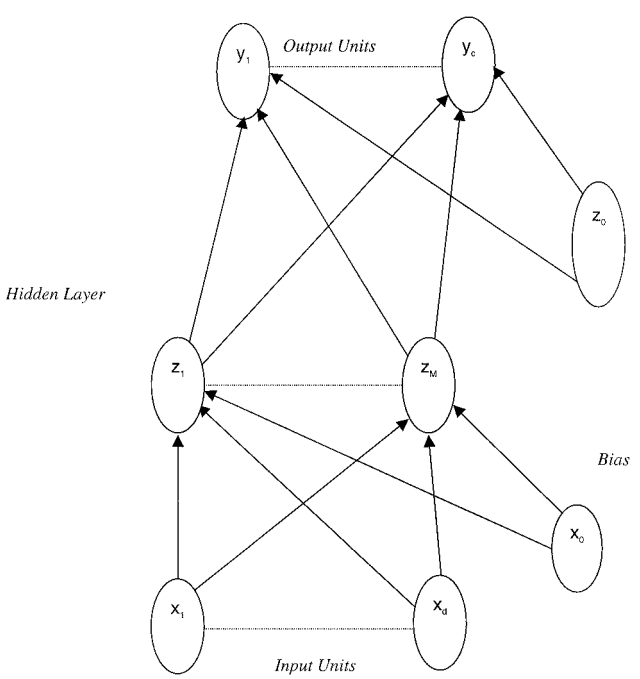

Fig. 1 Feedforward network having two layers of adaptive weights.

the modal properties. The nonlinear mathematical relation that maps the input x to the output y may be written as

$$y_k(\chi) = \sum_{j=1}^{M} w_{kj} f\left(\sum_{i=1}^{d} w_{kj} x_i + w_{j0} \right) + w_{k0} \tag{7}$$

The function $f(\cdot)$ implemented is sigmoid and is defined as

$$f(v) = 1/(1 + e^{-v}) \tag{8}$$

In the maximum likelihood approach, the weights (w_i) and biases (with subscripts 0 in Fig. 1) in the hidden layers are varied until the error between the network prediction and the output from the training data is minimized. Optimization routines, such as the scaled conjugate gradient method, are utilized for training. The ith error between the network prediction and the output from the training data is defined as

$$\text{error}_i = \frac{\| y - h(x)_i \|}{\| h(x)_i \|} \tag{9}$$

Here, $\| \cdot \|$ is the Euclidean norm of $\cdot$. The analysis employed in this section does not take into account the randomness of the input parameters. Thus, the MLP will be reformulated using the Bayesian approach.

D. Input to Neural Network

The FRFs will be transformed into modal energies. These modal energies are defined as the integrals of the real and imaginary components of the FRFs over frequency ranges that bracket the natural frequencies of the system. The identified modal energies will be transformed into the coordinate modal energy assurance criterion (COMEAC). The COMEAC is a criterion that measures the correlation between modal energies at each degree of freedom and the median over the modal energies. The COMEAC is similar to the coordinate modal assurance criterion COMAC[5] and is defined as

$$\text{COMEAC}\ (i) = \frac{\left[\sum_{j=1}^{N} |\eta_{\text{med}}(i, j) \eta_m^*(i, j)| \right]^2}{\sum_{j=1}^{N} |\eta_{\text{med}}(i, j)|^2 \sum_{j=1}^{N} |\eta_m(i, j)|^2} \tag{10}$$

Here, η_{med} is the median modal energy matrix taken over the population of undamaged structures, and η_m is the measured modal energy matrix. Similarly, when modal energy matrices are perfectly correlated, then the COMEAC for all degrees of freedom is 1. Otherwise, two modal energy matrices that are totally uncorrelated give the COMEAC for all degrees of freedom of 0. The COMEAC and the corresponding identity of fault will be used to train the FRF network.

Alternatively, modal analysis may be used to identify modal properties (mode shapes and natural frequencies). The mode shapes are transformed into the COMAC using Eq. (10) by substituting for η by mode shape vector. The natural frequency and the COMAC will be used to train the modal-property network.

To ensure that high-order input values do not dominate the training, the input parameters are normalized so that all of their values lie in the interval [0, 1]. To achieve this normalization, the following transformation is applied for the ith row of the input:

$$x_i^{\text{new}} = \frac{x_i^{\text{old}} - \min\left(x_i^{\text{old}}\right)}{\max\left(x_i^{\text{old}}\right) - \min\left(x_i^{\text{old}}\right)} \tag{11}$$

E. Committee of Neural Networks

In this paper, a method illustrated schematically in Fig. 2 is extended to a probability framework and experimentally validated. It has been shown before[3] that a committee of networks that uses both frequency and modal domain data gives results that are more reliable than when the two networks are used individually. The committee has been found to give lower mean square errors and standard deviation (thus giving solutions with a higher probability of being correct) than the individual methods.

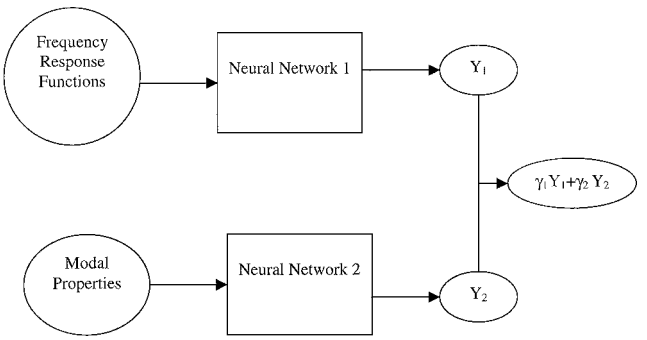

Fig. 2 Illustration of committee of networks.

F. Bayesian Approach

If the input vector $\{x\}$, is random (it has a mean and variance), because of the variation in physical properties of cylinders and variation of measurements, then the identity of damage y will be probabilistic. This will require that the weight space be assigned a probability distribution representing the relative degrees of belief in different values for the weight vector. This implies that the mapping function between input vector and output vector has a probability distribution. The weight space vector is initially assigned some earlier distribution. Once the data, in this case the FRFs or modal properties, and identities of the faults have been observed, the weight vector can be transformed into posterior distribution using Bayes' theorem. The posterior distribution can then be used to evaluate the predictions of the trained network for data not used during training.[6,7] Bayes' theorem may be written as follows:

$$p(w \mid D, x) = \frac{p(D \mid w, x) p(w \mid x)}{p(D \mid x)} \tag{12}$$

In Eq. (12), $p(w \mid x)$ is the probability distribution function of the weight space in the absence of any data (also known as prior distribution) and $D \equiv (y_1, \ldots, y_N)$ is a matrix containing the identity of damage data. The quantity $p(w \mid D, x)$ is the posterior probability distribution after the data have been seen and $p(D \mid w, x)$ is the likelihood function. The MLP network trained by supervised learning does not model the distribution of the input data. This x is a conditioning variable that always appears on the right-hand side[7] of the probabilities. For the remaining part of this study, x will be omitted to simplify the notation.

1. *Likelihood Function*

The likelihood function for a normal distribution may be written as

$$p(D \mid w) = [1/Z_D(\beta)] \exp(-\beta E_D) \tag{13}$$

The function $Z_D(\beta)$ is a normalization factor given by

$$Z_D(\beta) = \int \exp(-\beta E_D)\, \mathrm{d}D = \int \exp(-\beta E_D)\, \mathrm{d}y_i, \ldots, \mathrm{d}y_N \tag{14}$$

If the identity of damage data is a smooth function with zero-mean Gaussian noise, then the probability of observing the identity of damage data D for a given input vector may be written as follows:

$$p(D \mid w) = \frac{1}{Z_D(\beta)} \exp\left(-\frac{\beta}{2} \sum_{n=1}^{N} \{h(x^n, w) - y^n\}^2 \right) \tag{15}$$

The integral in Eq. (14) is the normalization factor that can be calculated[7] to give

$$Z_D(\beta) = (2\pi/\beta)^{N/2} \tag{16}$$

2. Prior Probability Function of Weights

The prior probability function for weights may be written as

$$p(w \mid x) = [1/Z_W(\alpha)] \exp(-\alpha E_W) \tag{17}$$

The function $Z_W(\alpha)$ is a normalization factor given by

$$Z_W(\alpha) = \int \exp(-\alpha E_W)\, \mathrm{d}w \tag{18}$$

The probability of vector w may be written as

$$p(w) = \frac{1}{Z_W(\alpha)} \exp\left(-\frac{\alpha}{2}\|w\|^2\right) = \frac{1}{Z_W(\alpha)} \exp\left(-\frac{\alpha}{2}\sum_{i=1}^{W} w_i^2\right) \tag{19}$$

The integral in Eq. (18) gives the same form of expression as Eq. (16) and is as follows:

$$Z_W(\alpha) = (2\pi/\alpha)^{W/2} \tag{20}$$

3. Posterior Distribution of Weight Vector

The distribution of the weights $p(D/w)$ after the data have been seen is calculated by substituting Eqs. (15) and (19) into Eq. (12) to give

$$p(w \mid D) = \frac{1}{Z_s} \exp\left(-\frac{\beta}{2}\sum_{n=1}^{N}\{h(x^n; w) - y^n\}^2 - \frac{\alpha}{2}\sum_{i=1}^{W} w_i^2\right) \tag{21}$$

where

$$Z_s(\alpha, \beta) = \int \exp(-\beta E_D - \alpha E_W)\, \mathrm{d}w \tag{22}$$

E_D and E_W are obtained from Eqs. (15) and (19). The optimal weight corresponds to the maximum of the posterior distribution.

4. Distribution of Network Outputs

The application of the Bayesian approach to neural networks results with the weight vector that has the mean and standard deviation. As a result, the output parameters will have a probability distribution. By the use of the rules of probability, the distribution of the outputs y for a given input vector x may be written in the following form:

$$p(y \mid x, D) = \int p(y \mid x, w) p(w \mid D)\, \mathrm{d}w \tag{23}$$

G. Monte Carlo Methods

In this section, Monte Carlo methods will be employed to find the distribution of the weight vectors and, subsequently, of the output parameters. The integral in Eq. (23) may be written as follows:

$$I = p(y \mid x, D) = \int F(w) p(w \mid D)\, \mathrm{d}w \tag{24}$$

Equation 24 may be approximated as follows:

$$I \cong \frac{1}{L}\sum_{i=1}^{L} F(w_i) \tag{25}$$

In Eq. (25), a sample weight vector $\{w_i\}$ is generated from the distribution $p(w \mid D)$. Because it is relatively difficult to generate the weight vector with a required distribution, a simpler distribution $q(w)$ will be considered, making Eq. (25)

$$I = \int F(w)\frac{p(w \mid D)}{q(w)} q(w)\, \mathrm{d}w \cong \frac{1}{L}\sum_{i=1}^{L} F(w_i)\frac{p(w_i \mid D)}{q(w_i)} \tag{26}$$

To generate the weight vectors representative of the distribution $p(w, D)$, we search through the weight space to find regions where $p(w \mid D)$ is sufficiently large. This will be achieved employing a technique called the Markov chain Monte Carlo[7] method. This technique considers sequences of vectors where each successive vector depends on the preceding vector plus a random component. This can be written as follows:

$$w_{\text{new}} = w_{\text{old}} + \varepsilon \tag{27}$$

In Eq. (27), ε is some small random vector generated from a spherical Gaussian distribution with small variance. To ensure that this technique samples at regions where $p(w \mid D)$ is high, the Metropolis et al. algorithm[8] is employed by using the following criterion:

$$\text{if} \quad p(w_{\text{new}} \mid D) > p(w_{\text{old}} \mid D) \quad \text{accept}$$

$$\text{if} \quad p(w_{\text{new}} \mid D) < p(w_{\text{old}} \mid D)$$

accept with probability

$$\frac{p(w_{\text{new}} \mid D)}{p(w_{\text{old}} \mid D)} \tag{28}$$

When steps are generated it is ensured that the probability of generating a candidate vector w_2 is the same as the probability of generating the current vector w_1. The main shortcoming with the Metropolis et al. technique[8] is that it requires that the distribution be uncorrelated. This is not the case when implementing neural networks. This problem is solved by taking into account the gradient of $p(w \mid D)$ and using it to determine the direction, which results with higher posterior probability. For neural networks, backpropagation is implemented to calculate the gradient. To ensure that the gradient information corresponds to a required distribution, hybrid Monte Carlo simulation[9] is used. To ensure that the algorithm does not spend a long time in the vicinity of poor region of local maximum probability, a simulated annealing technique[10] is used. The Metropolis et al. algorithm[8] may, thus, be modified to give

$$\text{if} \quad p(w_{\text{new}} \mid D) > p(w_{\text{old}} \mid D) \quad \text{accept}$$

$$\text{if} \quad p(w_{\text{new}} \mid D) < p(w_{\text{old}} \mid D)$$

accept with probability

$$\exp\{\ln[p(w_{\text{new}} \mid D)] - \ln[p(w_{\text{old}} \mid D)]/T\} \tag{29}$$

For temperature $T = 1$ the desired distribution is recovered. For $T \gg 1$ the system can explore weight space more freely and can escape from the minimum local error function. By using this technique, the uncertainty of the output may be assessed.

The committee procedure explained in Sec. II.D. is adapted to the Bayesian framework by generating I_1 and I_2 [in Eq. (26)], representing the distribution of the identity of faults for a given modal properties and FRFs, respectively. The overall distribution will then be evaluated by calculating the weighted summation of I_1 and I_2.

III. Experimental Example

In this section an impulse hammer test is performed on each of the 10 steel seam-welded cylindrical shells (1.75 ± 0.02-mm thickness, 101.86 ± 0.29-mm diam., and height of 101.50 ± 0.20 mm). The reason why cylinders are chosen is that many aircraft components have cylindrical shapes, for example, fuselage, casing of the engine, nacelle, etc. These cylinders are rested on a bubble wrap, to simulate a free–free environment (see Fig. 3). The cylinders are excited using a modal hammer with a sensitivity of 4 pC/N, a head of mass 6.6 g, and a cutoff frequency of 3.64 kHz. The response is measured using an accelerometer with a sensitivity of 2.6 pC/ms^{-2}, which has a mass of 19.8 g. Conventional signal processing procedures are applied to covert the time domain impulse history and response data into the frequency domain. The excitation and response data in the frequency domain are utilized to calculate the FRFs. From the FRFs, modal energies are extracted.

Each cylinder is divided into three substructures, and 12-mm-diam holes are drilled into each substructure (see Fig. 3). For example, for one cylinder, the first type of damage is a zero-fault scenario, and its identity is [0 0 0]. The second type of damage is a one-fault scenario, and if it is located in substructure 1, then its identity is

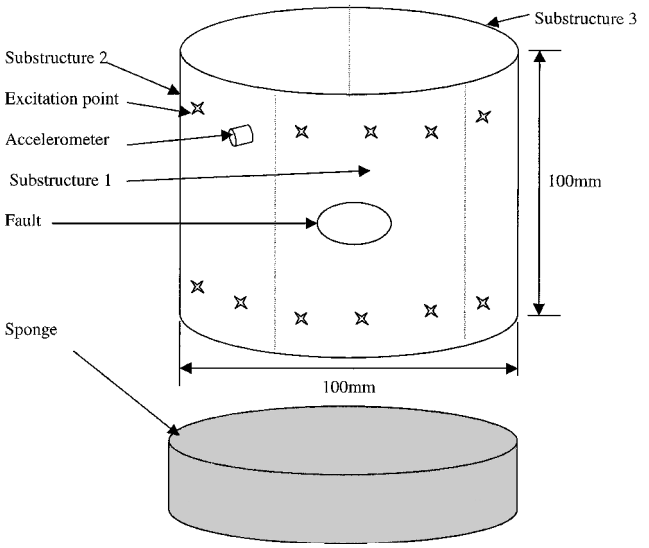

Fig. 3 Cylindrical shell of 1.75-mm thickness.

[1 0 0]. The third type of damage is a two-fault scenario, and if the faults are located in substructures 1 and 2, then the identity of this case is [1 1 0]. The final type of damage is a three-fault scenario, and the identity of this case is [1 1 1]. For each damage case, measurements are taken by measuring the acceleration at a fixed position and roving the impulse position. One cylinder gives 4 damage scenarios and 12 sets of measurements (a factor of 3 for repeatability). The total number of data collected is 120. From the 100 measured data, an additional 100 are generated by adding 5% random noise to the measured data.

The structure is vibrated at 19 different locations (see Fig. 3), 9 on the upper ring of the cylinder and 10 on the lower ring of the cylinder. Each measurement is taken three times to quantify the repeatability of the measurements. Some of the problems that are encountered during impulse testing include the difficulty in exciting the structure at an exact the position (especially for an ensemble of structures) and that the direction of the hammer cannot be accurately repeated.

Adding and subtracting 6 Hz from the resonance frequencies identifies the modal energy bandwidth. From the identified modal energies the COMEAC values are calculated. In calculating the COMEAC, natural frequencies below 1500 Hz are used. These COMEAC values are used to train the FRF neural network. This network has 19 input parameters, 11 hidden units, and 3 output units.

Furthermore, modal analysis is utilized to extract natural frequencies, damping ratios, and the mode shapes. The mode shapes are converted into COMAC. The 22 most reliable natural frequencies and COMAC values are used to train the modal-property neural network. This neural network has 22 input parameters, 13 hidden units, and 3 output units.

On training the networks, the coefficient of weight decay prior to distribution α and the coefficient of data error distribution β are initialized to be 0.001 and 100, respectively. The number of samples retained when hybrid Monte Carlo is employed is 1000. The step size for each trajectory is 0.002. From the data measured, 203 sets are chosen randomly and used to train the two networks. The remaining 17 sets of data are used to test the generalization of the networks. The FRF network and the modal-property network are combined to form a committee, which is shown in Fig. 2.

IV. Results and Discussion

The ensemble of 10 cylinders without faults is measured, and the FRFs are shown in Fig. 4. Figure 4 shows that the repeatability of the measurements is generally good at lower frequencies and, as expected, becomes poor at higher frequencies (above 1500 Hz). The presence of an accelerometer and the imperfection of the cylinders destroy the axisymmetry of the structures.

The average and the sample standard deviation of the elements of the COMEAC vector (corresponding to measured points) for both damaged and undamaged cylinders are shown in Table 1. Modal energies for higher modes (above 1500 Hz) were deliberately left out because they are generally noisy. From Table 1, the degree of repeatability of the COMEAC for the undamaged structures is highest in coordinate 3, then 7, 5, 19, 2, 14, 8, 11, 16, 18, 1, 17, 13, 9, 15, 4, 12, 6, and then 10. The repeatability of the COMEAC for data from the damaged structures is highest in coordinate 7, then 18, 16, 5, 13, 2, 11, 19, 3, 8, 15, 4, 14, 9, 17, 6, 12, 10, and then 1. From this set of data it is concluded that the COMEAC parameters at all measured positions are reliable enough to be utilized for damage identification. This conclusion is reached by comparing the standard deviations (see Table 1) of undamaged and damaged responses.

Table 2 shows the COMAC between the median mode shape for undamaged cylinders and that of all of the cylinders. The natural frequencies of this system are 413, 427, 561, 577, 1165, 1198, 1408, 1439, 1580, 2228, 2350, 2519, 2623, 3228, 3387, 3586, 3987, 4309, and 4818 Hz. The average and the sample standard deviation of the elements of the COMAC vector (corresponding to a measurement

Table 1 COMEAC between the median modal energy for undamaged cylinders and that of all of the cylinders[a]

Coordinate number	Average (COMEAC) undamaged	Standard deviation σ, undamaged	Average (COMEAC) damaged	Standard deviation σ, damaged
1	0.9370	0.0371	0.9141	0.1077
2	0.9807	0.0127	0.9849	0.0196
3	0.9824	0.0111	0.9779	0.0279
4	0.9312	0.0627	0.9645	0.0481
5	0.9828	0.0118	0.9846	0.0189
6	0.8587	0.0955	0.9003	0.0705
7	0.9856	0.0117	0.9897	0.0127
8	0.9691	0.0174	0.9644	0.0325
9	0.9557	0.0467	0.9643	0.0641
10	0.8947	0.1049	0.9427	0.0910
11	0.9695	0.0181	0.9804	0.0210
12	0.8792	0.0651	0.8968	0.0735
13	0.9592	0.0378	0.9841	0.0190
14	0.9712	0.0130	0.9667	0.0558
15	0.9277	0.0472	0.9677	0.0329
16	0.9685	0.0246	0.9804	0.0188
17	0.9055	0.0371	0.9021	0.0699
18	0.9658	0.0330	0.9848	0.0184
19	0.9710	0.0118	0.9672	0.0261

[a]Frequency bandwidth = 12 Hz.

Table 2 COMAC between the median mode shape for undamaged cylinders and that of all of the cylinders

Coordinate number	Average (COMAC) undamaged	Standard deviation σ, undamaged	Average (COMAC) damaged	Standard deviation σ, damaged
1	0.9044	0.0376	0.8546	0.1119
2	0.9191	0.1048	0.9224	0.1010
3	0.8905	0.0921	0.8938	0.1009
4	0.8662	0.0886	0.8287	0.1464
5	0.9188	0.0933	0.9304	0.0744
6	0.7488	0.1660	0.7451	0.1655
7	0.8799	0.1039	0.8751	0.1090
8	0.9419	0.0862	0.9274	0.0784
9	0.8527	0.1646	0.8829	0.0871
10	0.8560	0.1192	0.8870	0.0717
11	0.9346	0.0645	0.9464	0.0277
12	0.8655	0.0662	0.8538	0.0897
13	0.9434	0.0512	0.9335	0.0250
14	0.9591	0.0170	0.9487	0.0261
15	0.9154	0.0576	0.8836	0.0566
16	0.9610	0.0260	0.9585	0.0237
17	0.8880	0.0554	0.8139	0.0730
18	0.9471	0.0256	0.9371	0.0293
19	0.9352	0.1303	0.9484	0.0278

Table 3 Generalization of the FRF neural network[a] and confidence intervals

Fault case	Exact substructure 1	Exact substructure 2	Exact substructure 3	Identified substructure 1	Identified substructure 2	Identified substructure 3
1	0	0	0	0.00 (0.00–0.01)	0.00 (0.00–0.01)	0.00 (0.00–0.01)
2	0	0	0	0.00 (0.00–0.01)	0.38 (0.00–0.84)	0.00 (0.00–0.01)
3	0	0	0	0.01 (0.00–0.09)	0.00 (0.00–0.01)	0.00 (0.00–0.01)
4	0	0	0	0.99 (0.99–1.00)	0.00 (0.00–0.01)	0.00 (0.00–0.01)
5	0	1	0	0.00 (0.00–0.01)	0.23 (0.00–0.58)	0.89 (0.59–1.00)
6	1	0	0	0.99 (0.98–1.00)	0.00 (0.00–0.01)	0.00 (0.00–0.00)
7	1	0	0	0.57 (0.10–1.00)	0.00 (0.00–0.01)	0.00 (0.00–0.00)
8	0	1	0	0.00 (0.00–0.01)	0.00 (0.00–0.01)	0.15 (0.00–0.39)
9	0	0	1	0.00 (0.00–0.01)	0.01 (0.01–0.02)	0.99 (0.99–1.00)
10	0	1	0	0.19 (0.00–0.52)	0.99 (0.98–1.00)	0.01 (0.00–0.07)
11	0	1	1	0.00 (0.00–0.01)	0.86 (0.58–1.00)	0.68 (0.24–1.00)
12	1	1	0	0.99 (0.98–1.00)	0.97 (0.95–1.00)	0.00 (0.00–0.00)
13	1	1	0	0.99 (0.98–1.00)	0.99 (0.98–1.00)	0.04 (0.00–0.15)
14	0	1	1	0.81 (0.49–1.00)	0.99 (0.99–1.00)	0.99 (0.98–1.00)
15	1	1	1	0.97 (0.85–1.00)	0.01 (0.00–0.10)	0.99 (0.98–1.00)
16	1	1	1	0.99 (0.98–1.00)	0.56 (0.12–1.00)	0.99 (0.98–1.00)
17	1	1	1	0.99 (0.98–1.00)	0.99 (0.98–1.00)	0.99 (0.98–1.00)

[a]Bandwidth = 12 Hz.

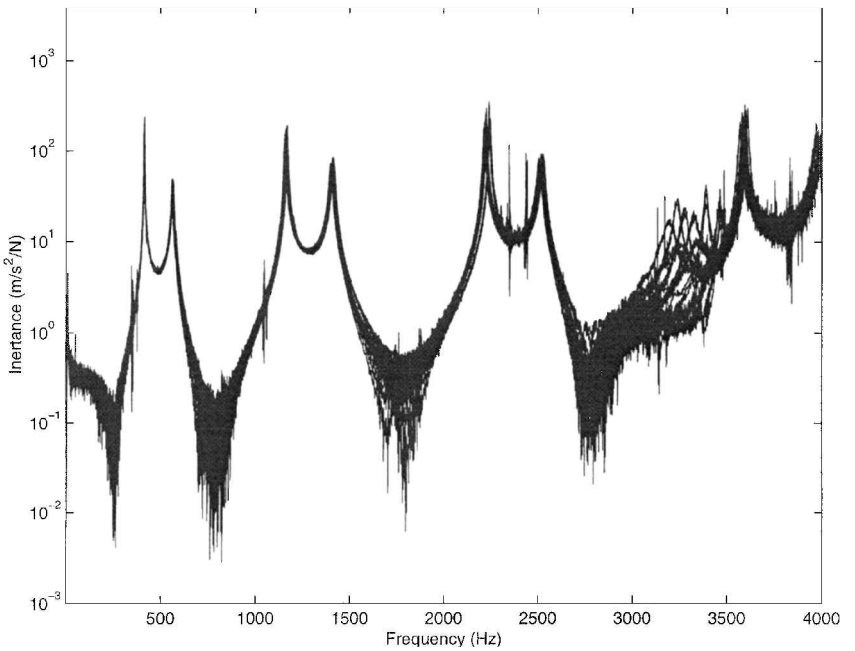

Fig. 4 Measured frequency response function of a population of undamaged cylinders.

coordinate) for both damaged and undamaged cylinders are shown in Table 2. These COMAC values are obtained by comparing the median mode shape matrices of all undamaged cases and that from each fault case. In calculating the COMAC, the mode shapes from modes 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, and 17 are used. The repeatability of the COMAC for data from the undamaged structures is highest in coordinate 14, followed by 18, 16, 1, 13, 17, 15, 11, 12, 8, 4, 3, 5, 7, 2, 10, 19, 9, and then 6. The repeatability of the COMAC for data from the damaged structures is highest in coordinate 16, then 13, 14, 11, 19, 18, 15, 10, 17, 5, 8, 9, 12, 3, 2, 7, 1, 4, and then 6. From this set of data it is concluded that the COMAC at coordinates, that is, positions of impulse, 1, 3, 4, 7, 12, 14, 17, and 18 are reliable enough to be utilized for fault identification. This conclusion is reached by comparing the standard deviations (see Table 2) of undamaged and damaged response.

The results of the generalization of the FRF neural network and modal-property network are shown in Tables 3 and 4, respectively. The CPU time taken to train the FRF and modal-property networks on a Pentium 200 MHz processor is 133.2 and 140.9 min, respectively. The results of the committee procedure are shown in Fig. 5 and Table 5. Figure 5 shows the mean square error vs the weighting function of the modal-property network. Figure 5 indicates that an optimal combination of the two networks gives a network with lower mean square errors than the individual methods. This observation has been mathematically proven before.[3] Figure 5 also shows that the optimal committee of networks (the lowest mean square error) is obtained by giving 30% weight to the modal-property network and 70% to the FRF network. Some of the reasons why the FRF network is found to be more accurate than the modal-property network are 1) the errors introduced when extracting modal properties from FRFs and 2) that the modal-property network is bigger (22 input units) than the FRF network (19 input units). Figure 6 shows that giving 50% weight to modal-property network and 50% to the FRF network offers the lowest standard deviation.

The generalization results are shown in Tables 3–5. For the zero-fault case 1, the three procedures give the correct solution. In the second case, the committee and the FRF network give the correct identity of fault, whereas the modal-property network approach incorrectly identifies the presence of a fault in substructure 1. In fault case 3, the three approaches correctly identify the faults. However,

Table 4 Generalization of the modal-property neural network and confidence intervals

Fault case	Exact substructure 1	Exact substructure 2	Exact substructure 3	Identified substructure 1	Identified substructure 2	Identified substructure 3
1	0	0	0	0.05 (0.00–0.23)	0.01 (0.00–0.05)	0.00 (0.00–0.01)
2	0	0	0	0.72 (0.32–1.00)	0.27 (0.00–0.65)	0.00 (0.00–0.01)
3	0	0	0	0.00 (0.00–0.01)	0.44 (0.00–0.91)	0.07 (0.00–0.30)
4	0	0	0	0.00 (0.00–0.01)	0.18 (0.00–0.50)	0.00 (0.00–0.01)
5	0	1	0	0.72 (0.30–1.00)	0.00 (0.00–0.01)	0.00 (0.00–0.01)
6	1	0	0	0.85 (0.53–1.00)	0.92 (0.68–1.00)	0.36 (0.08–0.79)
7	1	0	0	0.96 (0.82–1.00)	0.85 (0.53–1.00)	0.08 (0.00–0.32)
8	0	1	0	0.00 (0.00–0.01)	0.00 (0.00–0.01)	0.00 (0.00–0.01)
9	0	0	1	0.08 (0.00–0.32)	0.00 (0.00–0.01)	0.73 (0.32–1.00)
10	0	1	0	0.95 (0.78–1.00)	0.99 (0.98–1.00)	0.78 (0.40–1.00)
11	0	1	1	0.00 (0.00–0.01)	0.98 (0.87–1.00)	0.99 (0.98–1.00)
12	1	1	0	0.42 (0.00–0.89)	0.25 (0.00–0.63)	0.27 (0.00–0.67)
13	1	1	0	0.55 (0.06–1.00)	0.36 (0.00–0.81)	0.73 (0.33–1.00)
14	0	1	1	0.02 (0.00–0.15)	0.99 (0.98–1.00)	0.99 (0.98–1.00)
15	1	1	1	0.99 (0.98–1.00)	0.99 (0.99–1.00)	0.99 (0.98–1.00)
16	1	1	1	0.99 (0.98–1.00)	0.91 (0.67–1.00)	0.75 (0.37–1.00)
17	1	1	1	0.99 (0.94–1.00)	0.99 (0.97–1.00)	0.36 (0.00–0.84)

Table 5 Generalization of the committee neural network and confidence intervals

Fault case	Exact substructure 1	Exact substructure 2	Exact substructure 3	Identified substructure 1	Identified substructure 2	Identified substructure 3
1	0	0	0	0.00 (0.00–0.07)	0.00 (0.00–0.02)	0.00 (0.00–0.01)
2	0	0	0	0.22 (0.10–0.34)	0.35 (0.01–0.70)	0.00 (0.00–0.01)
3	0	0	0	0.07 (0.00–0.07)	0.13 (0.01–0.27)	0.02 (0.00–0.09)
4	0	0	0	0.70 (0.10–0.81)	0.06 (0.00–0.15)	0.00 (0.00–0.01)
5	0	1	0	0.22 (0.09–0.34)	0.16 (0.00–0.41)	0.62 (0.41–0.83)
6	1	0	0	0.96 (0.86–1.00)	0.28 (0.20–0.35)	0.11 (0.00–0.24)
7	1	0	0	0.69 (0.36–1.00)	0.26 (0.16–0.35)	0.02 (0.00–0.10)
8	0	1	0	0.02 (0.00–0.01)	0.00 (0.00–0.01)	0.10 (0.00–0.27)
9	0	0	1	0.02 (0.00–0.09)	0.01 (0.0–0.01)	0.92 (0.79–1.00)
10	0	1	0	0.42 (0.19–0.66)	0.99 (0.98–1.00)	0.24 (0.11–0.37)
11	0	1	1	0.00 (0.00–0.01)	0.90 (0.70–1.00)	0.78 (0.47–1.00)
12	1	1	0	0.83 (0.69–0.96)	0.77 (0.66–0.89)	0.08 (0.00–0.20)
13	1	1	0	0.86 (0.72–1.00)	0.81 (0.67–0.94)	0.25 (0.09–0.41)
14	0	1	1	0.57 (0.34–0.80)	0.99 (0.98–1.00)	0.99 (0.98–1.00)
15	1	1	1	0.98 (0.90–1.00)	0.31 (0.25–0.37)	0.99 (0.98–1.00)
16	1	1	1	0.99 (0.98–1.00)	0.67 (0.33–1.00)	0.92 (0.81–1.00)
17	1	1	1	0.99 (0.98–1.00)	0.99 (0.98–1.00)	0.81 (0.67–0.95)

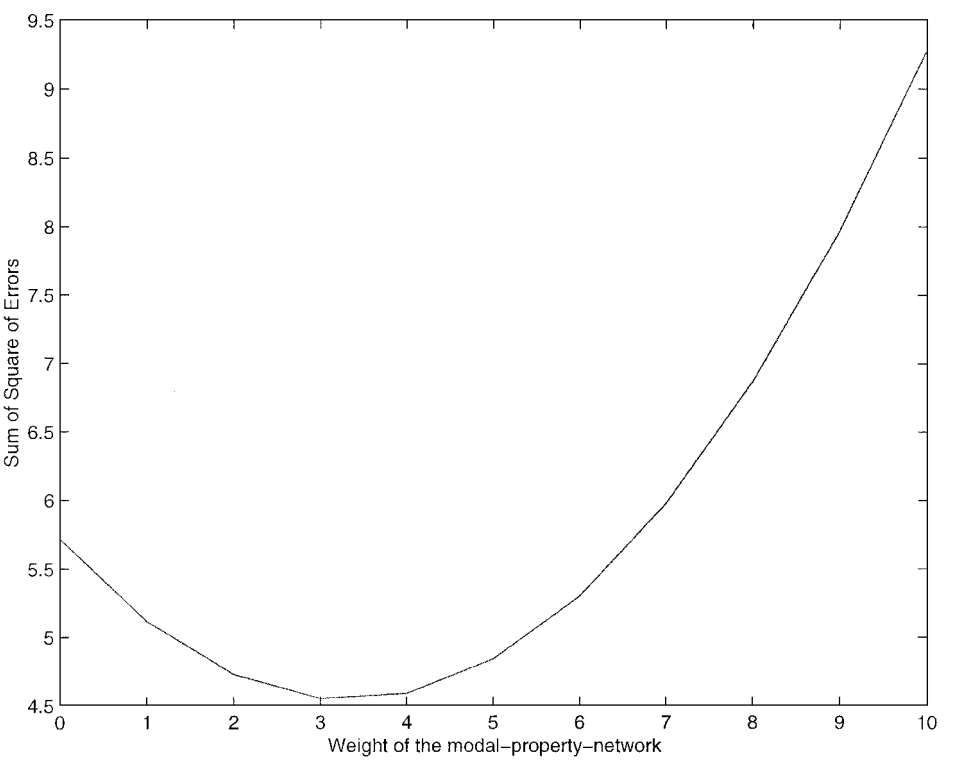

Fig. 5 Sum of square of errors vs weighting function of the modal-property network.

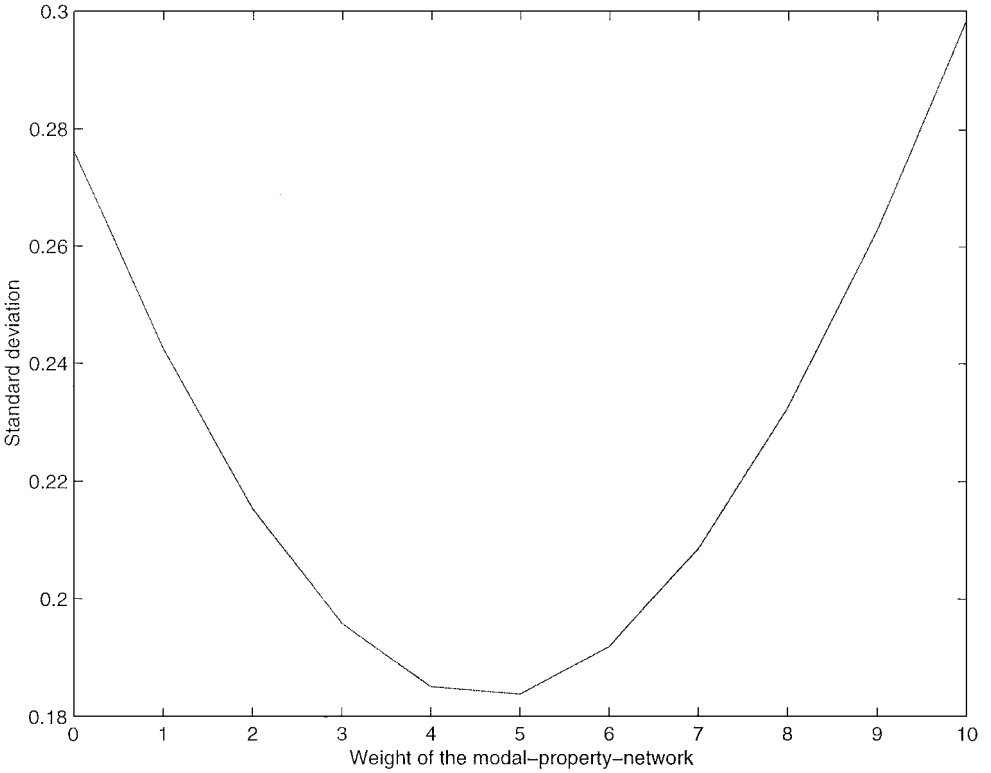

Fig. 6 Standard deviation vs weighting function.

note that the modal-property network gives wider confidence intervals than the two other approaches. In fault case 4, the modal-property network successfully identifies this fault whereas the other two approaches fail to do so. However, if confidence intervals are taken into account, the committee approach gives an indecisive solution for substructure 1.

For one-fault case 5, the three networks fail to correctly identify the presence of faults in substructure 2. For fault case 6, the committee and the FRF network correctly identify the one-fault-case, whereas the modal-property network identifies this fault case as a two-fault case. In fault case 7, the committee gives the best prediction, followed by the FRF network, whereas the modal-property network incorrectly identifies this one-fault case as a two-fault scenario. In fault case 8, the three approaches fail to identify the presence of fault in substructure 2. The confidence intervals are close to zero, implying that there are sufficient data to model this case. The reason for this failure is that the data set for this fault case gives conflicting information. In other words, the input data set (COFEAC, natural frequencies, and COMAC) for this fault case over an ensemble of cylinders has a higher variation than for other fault cases. The modal-property network identifies this one-fault case as a two-fault case. The committee gives inconclusive results, which is more informative than the two individual approaches. In fault cases 9 and 10 the three networks correctly identify the one-fault scenarios, with faults located in substructures 3 and 2, respectively.

In fault case 11, the three networks correctly identify this two-fault scenario. The modal-property network gives the best results, followed by the committee approach. In fault case 12, the FRF network correctly identifies this two-fault scenario, followed by the committee approach, whereas the modal-property network fails to identify the presence of faults in substructure 2. In fault case 13, the FRF network and the committee approach correctly identify the fault case, whereas the modal-property network fails to locate the fault in substructure 2 (it also gives a wider confidence interval). In fault case 14, the modal-property network correctly identifies the two-fault scenario. The committee approach and the FRF network incorrectly identify the presence of faults in substructure 1. However, if confidence intervals are taken into account, the committee approach is indecisive about whether a fault is present in substructure 1.

For fault case 15, the modal-property network gives the correct identity of the three-fault scenario. The committee and the FRF network fail to identify the presence of faults in substructure 2. For fault case 16, the modal-property network gives the best results, followed by the committee approach. The FRF approach experiences some difficulty in identifying the presence of faults in substructure 2. For fault case 17, the FRF network gives the best results, followed by the committee approach. The modal-property network fails to locate the presence of faults in substructure 3 (for this substructure it gives a wide confidence interval).

V. Conclusion

A committee of two Bayesian-formulated neural networks, which are trained using modal energies and modal properties, is successfully implemented to perform probabilistic fault identification in a population of cylindrical shells. The results show that the committee gives more accurate identities of faults (lower sum of squares of errors and standard deviations) and their respective confidence intervals than individual networks. Modal energies and modal properties are found to be sufficiently sensitive to be used for fault identification. The FRF network is found to be more accurate than the modal-property network. The committee is found to be a reliable alternative for fault identification if there is no prior knowledge as to which method is better. The modal-property network requires more computational time to train than the FRF network.

Acknowledgments

The author appreciates useful comments from H. E. M. Hunt, J. Woodhouse, and R. S. Langley.

References

[1] Doebling, S. W., Farrar, C. R., Prime, M. B., and Shevitz, D. W., "Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: a Literature Review," TR LA-13070-MS, Los Alamos National Lab., Albuquerque, NM, 1996.

[2]Marwala, T., and Heyns, P. S., "Multiple-Criterion Method for Determining Structural Damage," *AIAA Journal*, Vol. 36, No. 8, 1998, pp. 1494–1501.

[3]Marwala, T., and Hunt, H. E. M., "Fault Identification Using Finite Element Models and Neural Networks," *Mechanical Systems and Signal Processing*, Vol. 13, No. 3, 1999, pp. 475–490.

[4]Bishop, C. M., "Neural Networks for Pattern Recognition," Oxford Univ. Press, Oxford, England, U.K., 1995, pp. 384–433.

[5]Lieven, N. A. J., and Ewins, D. J., "Spatial Correlation of Modeshapes: the Coordinate Modal Assurance Criterion (COMAC)," *Proceedings of the 6th International Modal Analysis Conference*, Society of Experimental Mechanics, Union College, Schenectady, NY, 1988, pp. 690–695.

[6]Neal, R. M., "Bayesian Training of Backpropagation Networks by Hybrid Monte Carlo Method," TR CRG-TR-92-1, Dept. of Computer Science, Univ. of Toronto, Toronto, 1992.

[7]Neal, R. M., "Probabilistic Inference Using Markov Chain Monte Carlo Methods," TR CRG-TR-93-1, Dept. of Computer Science, Univ. of Toronto, Toronto, 1993.

[8]Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., "Equations of State Calculations by Fast Computing Machines," *Journal of Chemical Physics*, Vol. 21, No. 6, 1953, pp. 1087–1092.

[9]Duane, S., Kennedy, A. D., Pendleton, B. J., and Roweth, D., "Hybrid Monte Carlo," *Physics Letters*, Vol. 195, No. 2, 1987, pp. 216–222.

[10]Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., "Optimization by Simulated Annealing," *Science*, Vol. 220, No. 4598, 1983, pp. 671–680.