Journal of Sound and Vibration (1997) **201**(1), 85–101

STRUCTURAL FAULT DETECTION USING A NOVELTY MEASURE

K. WORDEN

Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S13JD, England

(*Received* 25 *April* 1996, *and in final form* 27 *August* 1996)

The method of novelty detection is applied to diagnose damage in a simple simulated lumped-parameter mechanical system. It is shown that the system transmissibility provides a sensitive feature for the detection of small stiffness changes. A simple yet successful means of establishing warning levels is demonstrated.

1. INTRODUCTION

In recent years, there has been a great deal of interest in applying Artificial Neural Networks (ANNs) to the problem of damage identification. This is because the problem is essentially one of pattern recognition and ANNs have proved extremely powerful in this respect [1]. In reference [2], Rytter distinguishes four levels of damage identification; level 1. (DETECTION) the method gives a qualitative indication that damage might be present in the structure; level 2. (LOCALIZATION) the method also gives information about the probable position of the damage; level 3. (ASSESSMENT) the method gives an estimate of the extent of the damage; level 4. (CONSEQUENCE) the method offers information about the safety of the structure, e.g., estimates a residual life.

The neural network architecture of choice depends on which level of identification is required. If levels 2 or 3 are required, a *supervised learning* scheme is required which associates damage classes (location and severity) with input patterns. The input patterns of features for training must be obtained from detailed *a priori* models; Finite Element (FE) methods are often used. The most commonly used ANN structure is the feed-forward Multi-layer Perceptron (MLP), usually trained by back-propagation of errors [3–18]. There are some applications of more exotic structures: a counter-propagation network [19], a Cerebellar Model Articulation Controller (CMAC) network [20] and a Radial Basis Function (RBF) [21]. These ANNs retain the feed-forward structure but rely on an initial clustering stage as part of the training. In this respect, they are hybrids between the supervised and unsupervised paradigms.

All of the studies cited above are based on models of simplified structures: e.g., lumped mass systems with three and eight degrees of freedom [15]; models of mult-storey buildings [4, 14, 18]; trusses and frameworks [6, 7, 11, 16, 19]; metal and composite beams and plates [3, 5, 12, 17, 22]; a stiffened panel [10]; a bridge element [13]. Before any of these networks could be used with confidence on experimental data, one would require assurance that the relevant models have appropriate accuracy. There are two issues here. First, the model of structure must be correct; and, second, the model of the *damage* must be correct. In most cases, the damage is simulated by reducing the stiffness of individual elements of the FE

0022–460X/97/110085 + 17 $25.00/0/sv960747

model. Only in reference [9] is a fracture mechanical model adopted to represent a fatigue crack. The *a priori* model is clearly critical and for complex structures will represent the most intensive and expensive part of the analysis. The ANN training is usually a somewhat routine process once training data is established.

Level 4 methods present different problems as detailed questions of fatigue and fracture arise if residual life is to be predicted with any degree of confidence. The author is unaware of any work to date which addresses this problem successfully using neural networks.

The object of the present study is to eschew the model-based approach and therefore pave the way for diagnostic methods applicable to systems of arbitrary complexity. There is of course an opportunity cost; the method described here will only provide a level-one diagnostic according to Rytter's classification.

A recent thread in ANN research is associated with methods of *novelty* or *anomaly detection*. The objecting is to monitor a sequence of patterns and signal if one arises which differs significantly from the herd. The application to condition monitoring or damage detection is manifest. However, there appear to be no precedents in the engineering literature. An interesting study of anomalous masses in mammograms is presented in reference [23]. The method is also applied to the detection of epileptic episodes in EEG records in reference [24]. There are many methods of novelty detection, the approach taken here is based on the use of Auto-Associative (AA) networks as described in reference [25].

As an aside, the establishment of robust novelty detection methods has important implications for the validation of neural networks [26]. If a pattern recognition network can also return a diagnostic which indicates how far removed the input pattern was from its training experience, the user can be directed to call in expert help in cases where the network is forced to extrapolate.

The layout of the paper is as follows. In section 2 a brief description of the Multi-Layer Perceptron Network is given and in section 3 a description of how the Auto-Associative form of the network is used to provide a measure of novelty or *novelty index*. In section 4 the system under study and the features used for pattern recognition are described. Section 5 contains the results of the study. In section 6 the use of warning levels is discussed and the paper concludes with a general discussion of the method in section 7.

2. THE NEURAL NETWORK

For the sake of completeness, a brief description of the Multi-Layer Perceptron (MLP) follows; for a more detailed discussion, the reader is referred to the seminal work [27].

The MLP is simply a collection of connected processing elements called nodes or neurons, arranged together in layers (see Figure 1). Signals pass into the input layer nodes, progress forward through the network hidden layers and finally emerge from the output layer. Each node i is connected to each node j in its preceeding layer through a connection of weight w_{ij}, and similarly to nodes in the following layer. Signals pass through the node as follows: a weighted sum is performed at i of all the signals x_j from the preceding layer, giving the excitation z_i of the node; this is then passed through a non-linear *activation function* f to emerge as the output of the node x_i to the next layer: i.e.,

$$x_i = f(z_i) = f\left(\sum_j w_{ij} x_j\right). \tag{1}$$

Various choices for the function f are possible; the hyperbolic tangent function $f(x) = \tanh(x)$ is used here. One node of the network, the *bias* node, is special in that it is connected to all other nodes in the hidden and output layers, the output of the bias node

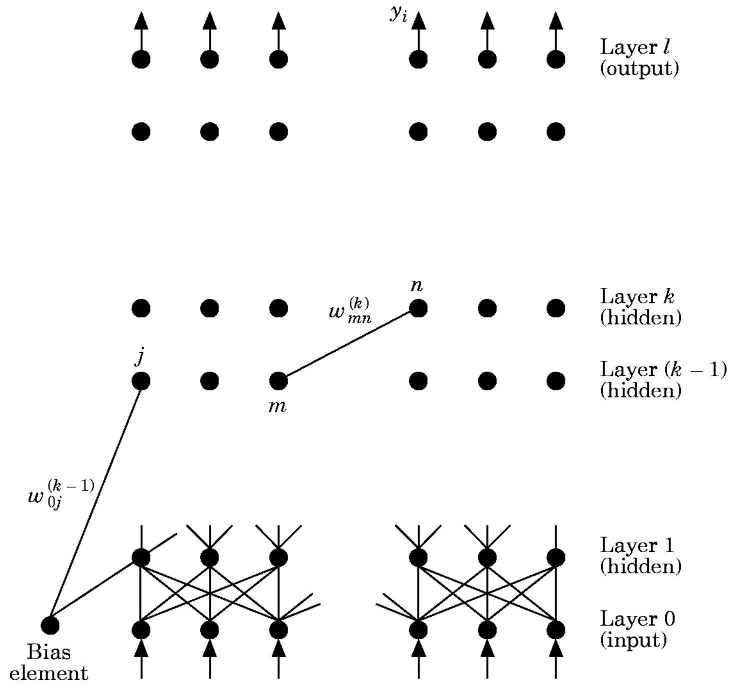

Figure 1. The Multi-Layer Perceptron (MLP) neural network.

is held constant throughout, in order to allow constant offsets in the excitations z_i of each node.

The first stage of using a network to model an input-output system is to establish the appropriate values for the connection weights w_{ij}. This is the *training* or *learning* phase. Training is accomplished by using a set of network inputs for which the desired network outputs are known. At each training step, a set of inputs is passed forward through the network yielding trial outputs which are then compared to the desired outputs. If the comparison error is considered small enough, the weights are not adjusted. If, however, a significant error is obtained, the error is passed *backwards* through the net and a *training algorithm* uses the error to adjust the connection weights. The algorithm used in this work is the *back-propagation* algorithm in which the parameter update rule is used: i.e.,

$$w_{ij}^{(m)}(t) = w_{ij}^{(m)}(t-1) + \eta\delta_i^{(m)}(t)x_j^{(m-1)}(t), \qquad (2)$$

where $\delta_i^{(m)}$ is the error in the output of the i^{th} node in layer m, and t is the index for the iteration. This error is not known *a priori* but must be constructed from the known errors $\delta_i^{(l)} = y_i - \hat{y}_i$ between the network outputs $\hat{y}_i$ and the desired outputs y_i. This is the origin of the term back-propagation. The update used here is modified by the inclusion of an additional *momentum* term which allows previous updates to persist:

$$\Delta w_{ij}^{(m)}(t) = \eta\delta_i^{(m)}(t)x_j^{(m-1)}(t) + \alpha\Delta w_{ij}^{(m)}(t-1). \qquad (3)$$

The effect of this extra term is to damp out oscillations in the weight estimates. The coefficients η and α determine the overall speed of learning; unfortunately, there are no hard and fast rules as to their optimum values for a given problem.

Once the comparison error is reduced to an acceptable level over the whole training set, the training phase ends and the network is established.

3. THE NOVELTY INDEX

As described above, the objective of novelty detection is to establish if a new pattern differs from previously obtained patterns in some significant respect. The application to on-line damage detection is clear. It is assumed that damage will alter the measured patterns, so that novelty will indicate a fault. The important point is to identify *significant* changes: i.e., those which cannot be attributed to fluctuations in the measured patterns due to noise.

The approach taken here is simply to train an Auto-Associative Network (AAN) on the patterns. This simply means a feed-forward Multi-Layer Perceptron (MLP) network which is asked to reproduce at the output layer those patterns which are presented at the input. This would be a trivial exercise except that the network structure has a "bottleneck": i.e., the patterns are passed through hidden layers which have fewer nodes than the input layer (see Figure 2). This forces the network to learn the significant features of the patterns; the activations of the smallest, central layer, correspond to a compressed representation of the input. Training proceeds by presenting the network with many versions of the pattern corresponding to normal condition corrupted by noise and requiring a copy at the output.

The novelty index $v(\mathbf{z})$ corresponding to a pattern vector $\mathbf{z} = z_i$, $i = 1, \ldots, N$ is then defined as the Euclidean distance between the pattern $\mathbf{z}$ and the result of presenting it to the network $\hat{\mathbf{z}}$,

$$v(\mathbf{z}) = \|\mathbf{z} - \hat{\mathbf{z}}\|. \tag{4}$$

It is clear how this works. If learning has been successful, then $\mathbf{z} = \hat{\mathbf{z}}$ for all data in the training set, so that $v(\mathbf{z}) \approx 0$ if $\mathbf{z}$ represents the normal condition. If $\mathbf{z}$ corresponds to damage, $v(\mathbf{z})$ is non-zero. Note that there is no guarantee that v will increase monotonically with the level of damage. This is why novelty detection only gives a level one diagnostic in Rytter's terminology [2].

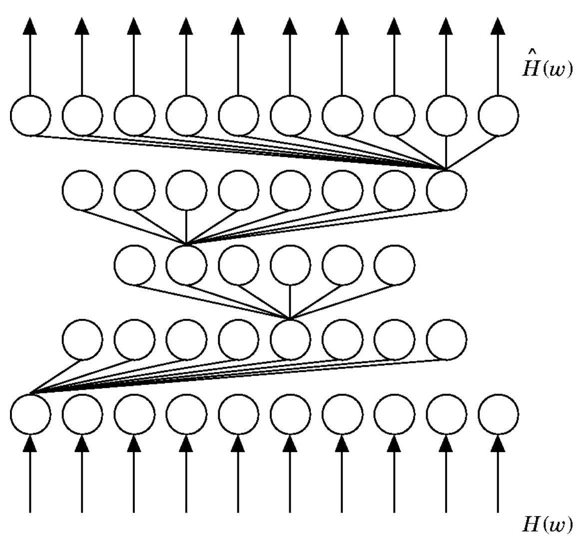

Figure 2. Auto-Associative Neural (ANN) network (most connections are suppressed).

4. THE TEST SYSTEM AND DATA

The simulated system which shall be used to demonstrate the novelty detection method is the three-degree-of-freedom (3-DOF) lumped-parameter system shown in Figure 3. The equations of motion are as follows:

$$m_1\ddot{y}_1 + c_{11}\dot{y}_1 + k_{11}y_1 + k_{12}(y_1 - y_2) = x_1(t),$$

$$m_2\ddot{y}_2 + c_{22}\dot{y}_2 + k_{21}(y_2 - y_1) + k_{23}(y_2 - y_3) = x_2(t),$$

$$m_3\ddot{y}_3 + c_{33}\dot{y}_3 + k_{32}(y_3 - y_2) + k_{33}y_3 = x_3(t) \tag{5}$$

The values $m_1 = m_2 = m_3 = 1$, $c_{11} = c_{22} = c_{33} = 20$ and $k_{11} = k_{12} = k_{21} = k_{23} = k_{31} = k_{33} = 10^4$ are used below.

The fault in this system is simulated by reducing the stiffness k_{12} by various degrees. Note that although the system is simple, and the means of simulating a fault is simple, because in the method one assumes no *a priori* model, it has much wider applicability than this simulation. The method does not seek to identify a specific damage state, but only a change. Consider, for example, the problem of detecting faults in composite plates. Even the normal condition may be difficult to model due to the complexity of the geometry or the lay-up. The fault states, e.g., fibre pullout, fibre fracture, matrix fracture and delamination, could be even more difficult to model. All the novelty index method needs is a sequence of measurements corresponding to normal condition; the simplest model can be used. Even experimental data can be used if available.

The feature for network training is chosen to be the transmissibility function between masses m_1 and m_2, denoted here by $H(\omega)$. This has previously proved useful in pattern recognition for vibration problems [28]. It is computed here by simulating the response

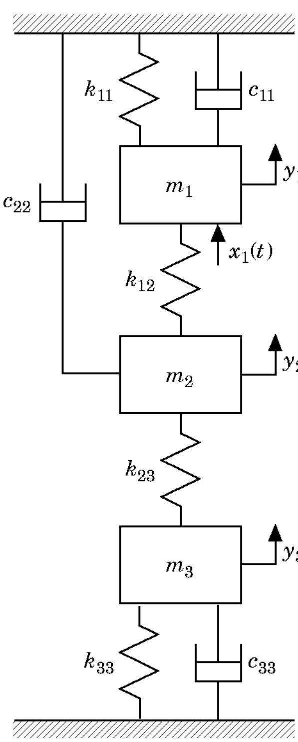

Figure 3. The three-degree-of-freedom simulated system.

to a harmonic excitation $x_1(t) = X\cos(wt)$ for a range of frequencies. The relative gain and phase between y_1 and y_2 is extracted in each case. However, only the magnitude is used below for training. A fourth order Runge–Kutta routine is used to integrate the equations of motion (5) [29]. The program used computes transmissibilities for arbitrary lumped mass non-linear systems. Note that for the linear system studied here, the functions could be obtained directly from the frequency domain versions of equations (5). The $H(w)$ is sampled at 50 regularly spaced points between 0 Hz and 50 Hz to produce the basic training vector.

The basic transmissibility function is shown in Figure 4, together with the functions corresponding to 1%, 10%, and 50% reductions in k_{12}. The 50% reduction produces a function which shows gross changes and visual inspection suffices to identify the fault in this case. The 10% reduction causes more subtle changes and the 1% reduction is of a very small order.

5. RESULTS OF NOVELTY DETECTION

The first stage of producing the novelty index is training the AA network. The training set was obtained by making 1000 copies of the transmissibility pattern corresponding to normal condition and corrupting each copy with different Gaussian noise vectors. In the absence of any prescription for the noise, the Gaussian process (with unit-proportional covariance matrix) was chosen; a minimal requirement for any pattern recognition system is that it should be transparent to normally distributed noise. In geometrical terms, the assumption here is that the normal condition set in pattern space is spherical. More complicated sets can always be contained within an appropriately large sphere: i.e., for high enough Gaussian noise. Note that such spheres will in general contain anomalous

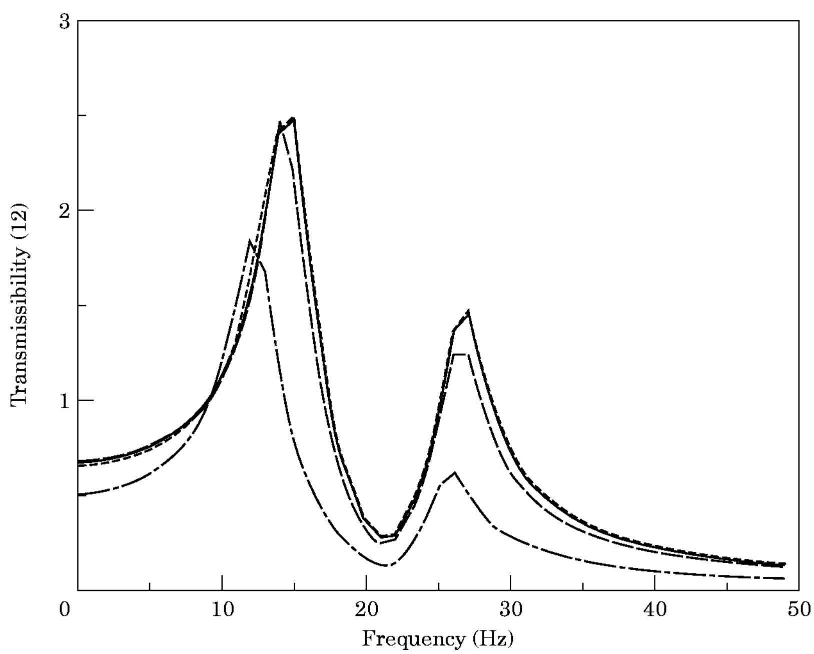

Figure 4. The transmissibility function for network training with various different stiffness reductions. ———, No fault; - - - -, 1% fault; — — —, 10% fault; — - —, 50% fault.

patterns and novelty indices based on training sets with high Gaussian noise might occasionally return false negatives.

Two simulations were carried out in order to investigate how the sensitivity of the novelty index depends on the noise level.

5.1. LOW NOISE

In this case the R.M.S. of the added noise was 0·01, compared with a peak value of 2·5 for the transmissiblity function. Three sample patterns are given in Figure 5. The degree of corruption is small but commensurate with the deviation in the 1% fault pattern (see Figure 4). This means that the 1% fault could not be detected by visual inspection at this noise level, unlike the 10% and 50% faults. The neural network structure was determined by trial and error, acceptable results were obtained for a network with five layers and node structure 50: 40: 30: 40: 50. The network was trained for 200 000 cycles, with the patterns presented in random order. An adaptive learning schedule was adopted in which the learning rate and momentum were chosen to be high in the early stages of training (0·3 and 0·4 respectively) but fell to low values (0·05 and 0·05) later. After training, the novelty index was evaluated on the training set and the results are shown in Figure 6. The mean value of v is 0·04 on the training set.

Testing sets were constructed by concatenating 1000 noise-corrupted copies of the 50%, 10% and 1% fault transmissibility on to the normal condition training data. The results of computing the novelty index for the 50% and 1% cases are given in Figures 7 and 8. The index unambiguously signals a fault in all cases, even the 1% damage case.

5.2. HIGH NOISE

The approach of the previous section was repeated on data corrupted with Gaussia noise of R.M.S. 0·05. Three sample patterns are shown in Figure 9. The level of corruption is now significant. One would usually construct transmissibilities from averaged

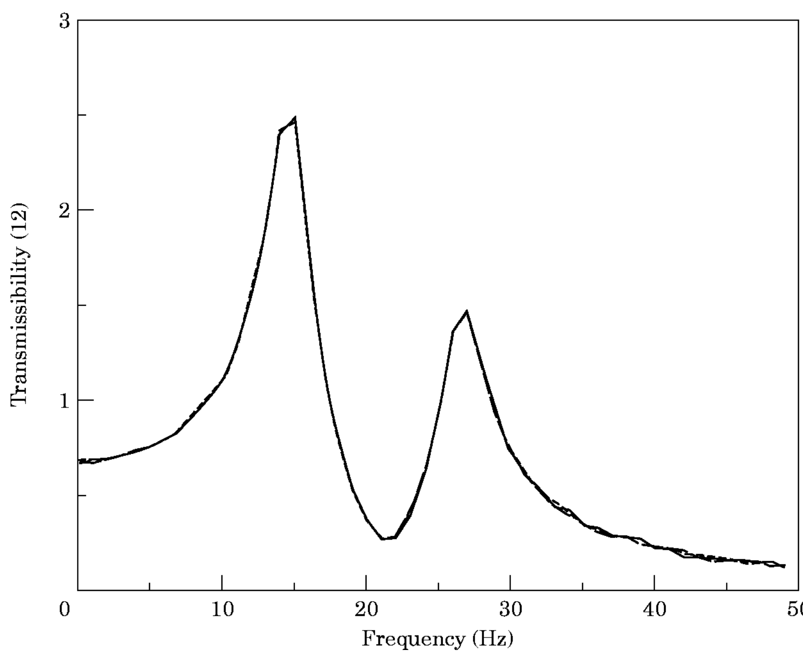

Figure 5. Sample patterns from the low noise training set. ———, Sample 1; - - - - , Sample 2; — — —, Sample 3.

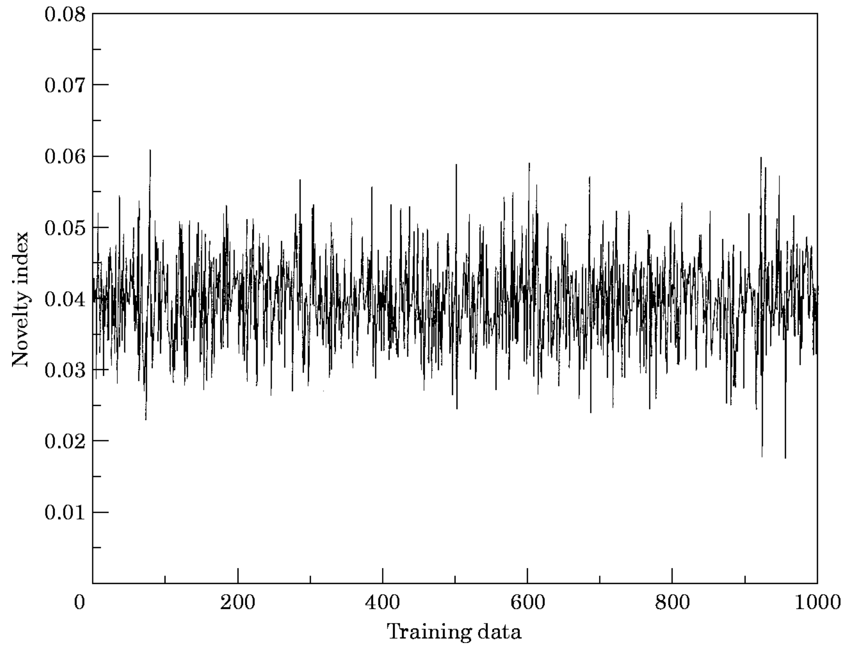

Figure 6. The novelty index evaluated on low noise 50:40:30:40:50 training set.

measurements from a frequency response analyzer. The variations between patterns would depend on the ambient noise level, the degree of "colour" in the noise and the number of averages taken. It was assumed here, on the basis of experience in the laboratory, that the higher of the levels of noise corruption used here was realistic. The deviations in the

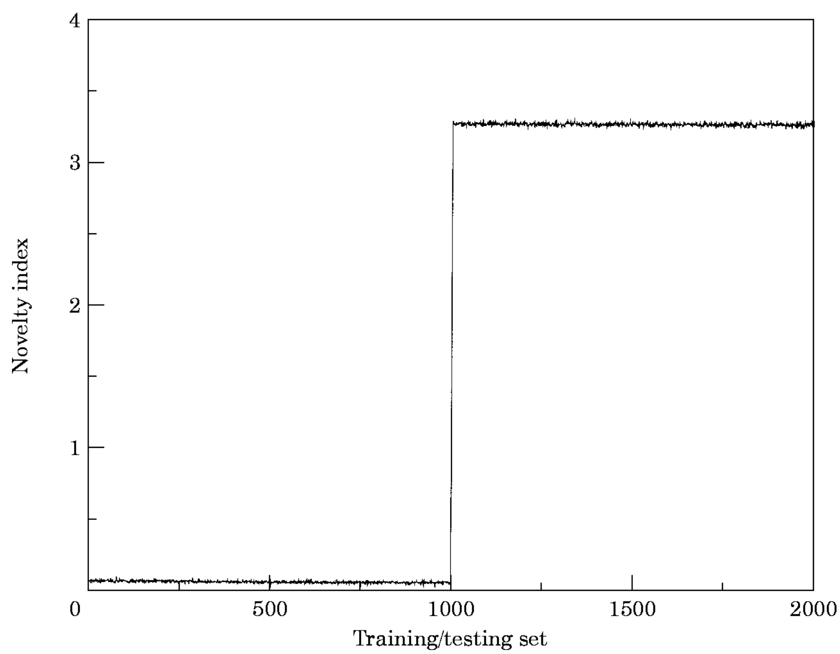

Figure 7. The novelty index evaluated on low noise data: training set and 50% fault data.

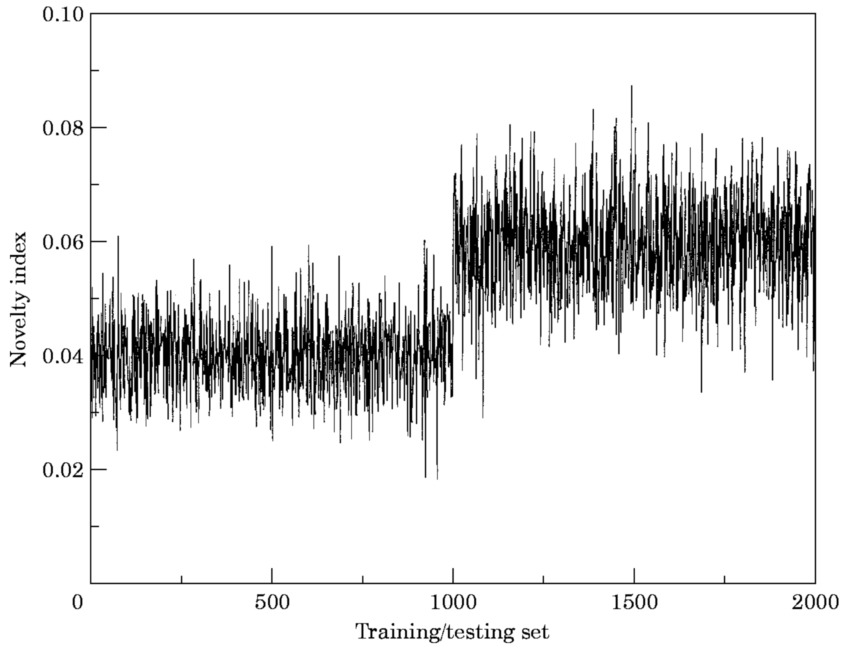

Figure 8. The novelty index evaluated on low noise data: training set and 1% fault data.

functions are large compared with the distortions induced by a 1% stiffness reduction (see Figure 4).

The AA network was trained exactly as before, the same structure and learning schedule were used. After training, the novelty index was evaluated over the training set. The mean

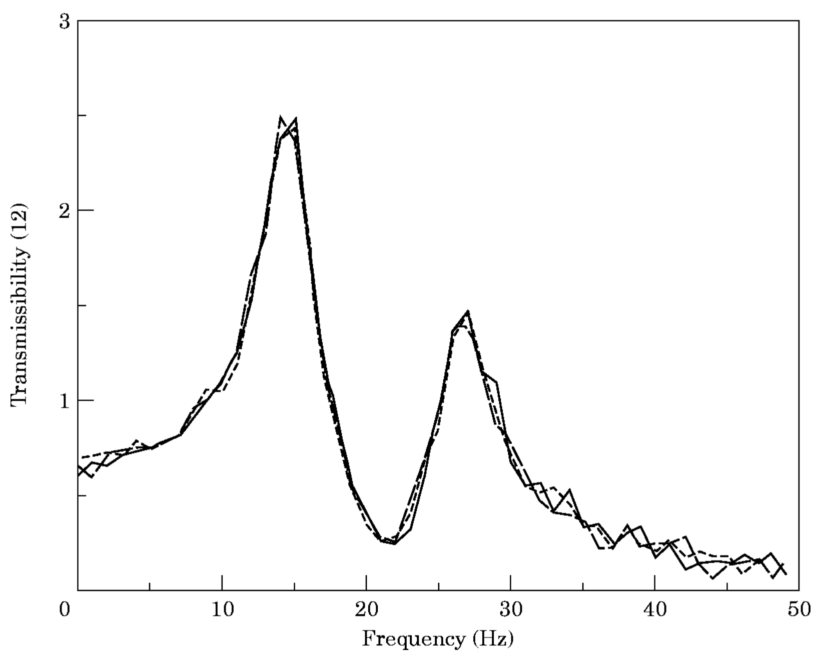

Figure 9. Sample patterns from the high noise training set. ———, Sample 1; - - - -, Sample 2; — — —, Sample 3.

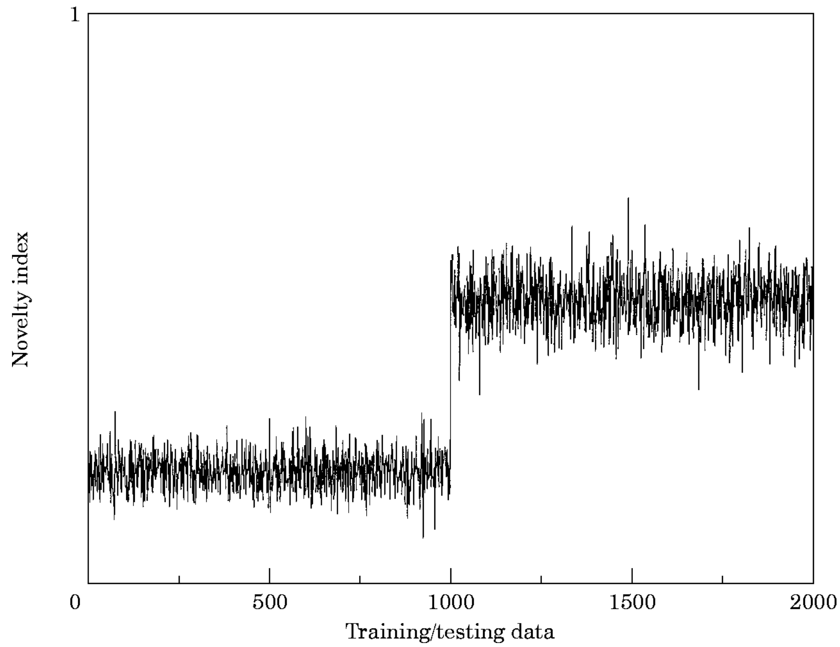

Figure 10. The novelty index evaluated on high noise data: training set and 10% fault data.

level for the index on data corresponding to normal condition was 0·2 compared with 0·04 for the low noise case.

Testing sets were constructed as before for the 50%, 10% and 1% fault cases and the results for the 10% and 1% faults are presented in Figures 10 and 11. The novelty index

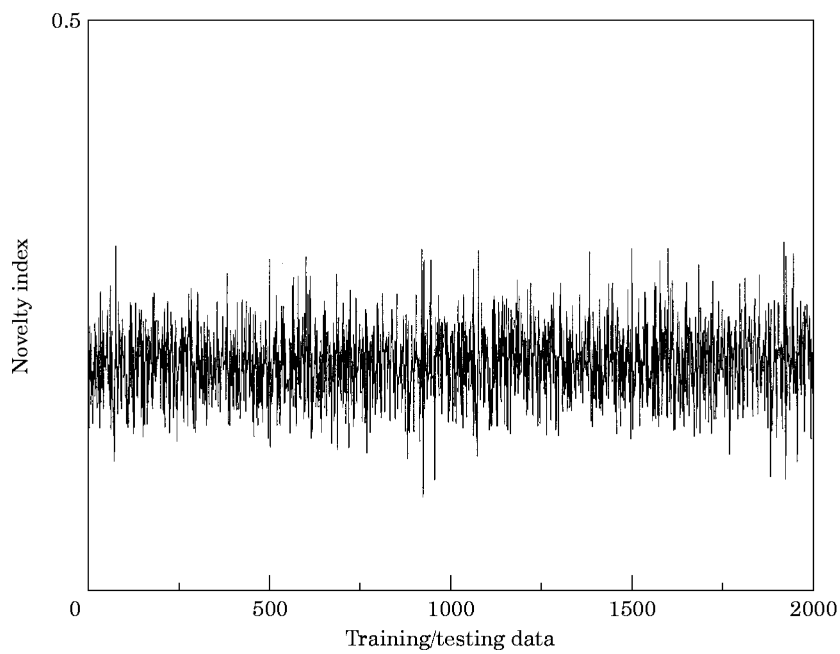

Figure 11. The novelty index evaluated on high noise data: training set and 1% fault data.

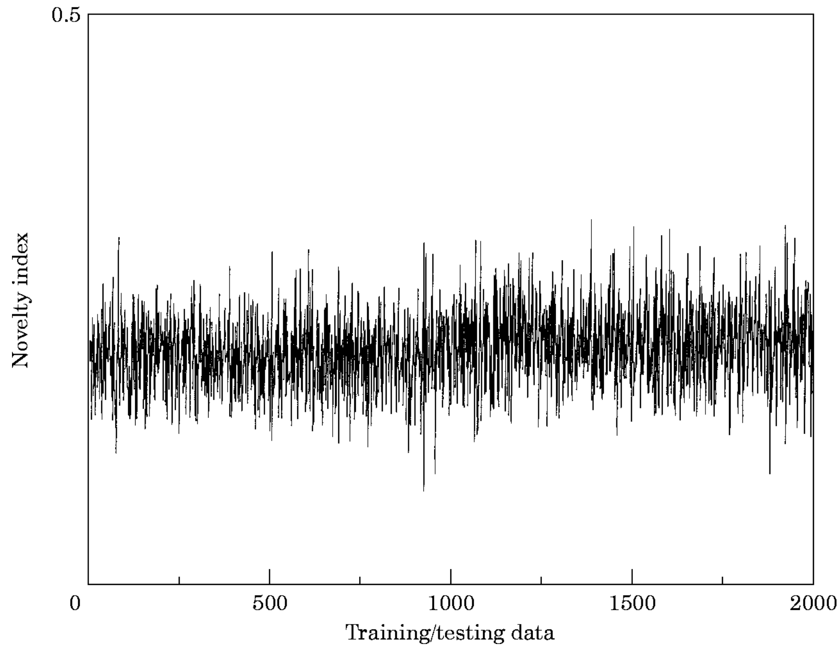

Figure 12. The novelty index evaluated on high noise data: training set and 2% fault data.

signals the 50% and 10% faults without difficulty, but is unable to flag the 1% fault. An additional simulation with a 2% reduction in stiffness produced the result given in Figure 12, the change in the novelty index is just distinguishable.

6. ALARM AND WARNING LEVELS

Given a means of tracking the time-variation of the system novelty index, it is natural to wish for some automatic means of establishing if significant change has occurred in the system condition.

The basic technique adopted here originates in the work of Cempel [30] on machine condition monitoring. To monitor the health of a structure, attention can be restricted to the time variation of one or more parameters which are associated with the condition. As an example, the temperature at a point on the structure could be recorded at regular intervals; any *significant* variation from a constant value may indicate a change in condition. The problem here is to define the meaning of *significant*. It is important to establish what range of values for a monitored parameter indicates normal condition. If monitored data is available from a sufficiently large sample of similar machines, which are in good condition, the variation in parameter values over the sample set can be used to determine confidence intervals for the expected values of parameters. However, this approach to fixing acceptable values for parameters has met with limited success; monitored parameters vary widely between nominally identical machines. These variations arise due to differences in manufacture, operating conditions, age, etc. It has been suggested that a much more reliable means of monitoring condition is to monitor *trends* in the observed parameters [30, 31].

To assist in the interpretation of condition monitoring data, and to highlight abnormal or dangerously high readings from a background of values, it is common for each monitored parameter to have associated with it two symptom limit routine values. These values are often defined as follows: the *warning level* is the value of the parameter above which it can be considered that a reading is sufficiently abnormal to require investigation; the *alarm level* is the value of the parameter above which it can be considered that the machine is no longer suitable for operation.

The situation can be simplified somewhat for novelty detection in structures. There is arguably no need to consider the alarm condition, as the novelty will not usually be a normalized quantity in the sense that the damage extent will not follow from a given value of novelty; the object is simply to signal change. The question is: How does one establish the warning level? Cempel's idea is very straightforward: the values of monitored parameters during stationary behaviour of the system are used to establish statistics for the expected variation in parameter readings. The warning level is thus a dynamical quantity, its value at a given time being dependent on the behaviour of the novelty index estimates at earlier times. If at each sampling instant i, a single estimate v_i is updated, the upper and lower warning levels for the next estimate v_{i+1} can be defined as $\bar{v} \pm \alpha\sigma_v$, where $\bar{v}$ and σ_v are, respectively, the mean and standard deviations of all novelty estimates up to and including v_i. If the distribution of v_i; $i = 1, \ldots, N$ is Gaussian, standard theory shows that a value of 1·96 for α will give estimates within the warning levels with 95% probability. The proportion of estimates which fall outside the warning interval purely due to statistical variations is termed the *percentage unnecessary repair* in reference [30], these signals are often referred to as *false positives*. α is to some extent determined by the frequency of measurement; if α is low and many measurements are taken, many false positives will result. For example, the above simulation had 1000 points in the training set for the novelty, with a value of 1·96 for α, there would occur 50 false positives in the training set if accurate statistics were estimated. The value of α taken here was 4·0; if Gaussian statistics are used, this would result in an expected number of 0·064 false positives.

To minimize data storage, the mean and standard deviation of the parameter records can be adjusted recursively by using the update formulae

$$\bar{v}_{i+1} = \left(\frac{i}{i+1}\right)\bar{v}_i + \left(\frac{i}{i+1}\right)v_{i+1}, \tag{6}$$

$$\sigma^2_{vi+1} = \left(\frac{i-1}{i}\right)\sigma^2_{vi} + \bar{v}^2_i + \frac{1}{i}\,v^2_{i+1} - \left(\frac{i+1}{i}\right)\bar{v}^2_{i+1}, \tag{7}$$

where v_{i+1} is the curent estimate and v_{i+1} is the corresponding update of the mean value.

If the warning levels are computed for the 1% fault in the low noise case described in the previous section, the results are shown in Figure 13. It is clear that the novelty estimate crosses the upper warning level almost immediately after the condition changes. It is unimportant that a number of the index values do not exceed the threshold because of statistical variations in the index: what is required is that at least one evaluation triggers the warning level within a short time of the parameter change. In the high noise case, the fault was not indicated. This is not surprising, as the fault does not produce a visible change in the novelty index. The 2% fault in the high noise case also failed to trigger the warning even though a visible change occurred in the index. This shows how critical the choice of α is; there is a trade-off between the number of false positives tolerated and the sensitivity of the warning indicator.

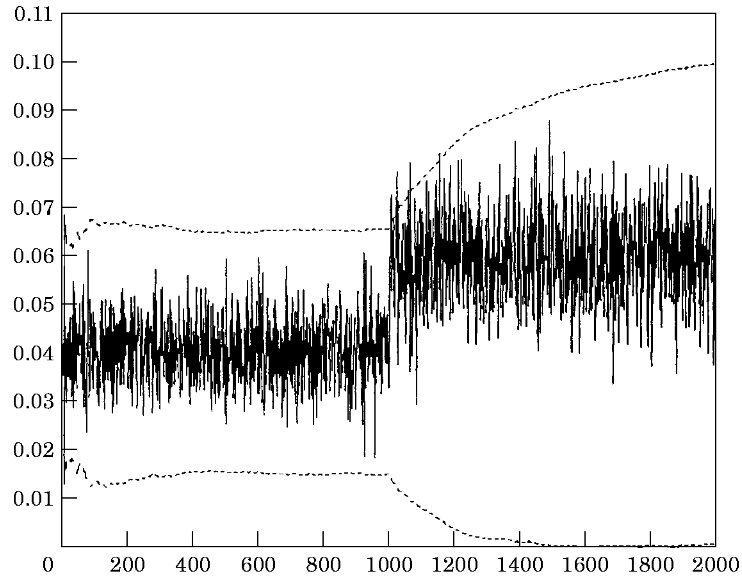

Figure 13. Warning levels for the novelty index evaluated on low noise data: training set and 1% fault data.

Note that there is no in-built forgetting factor for the mean and standard deviation calculation. As a result, the final values for the upper and lower warning levels are $\bar{v} \pm 4_{\sigma_v}$, where the statistics are calculated over the whole set of novelty values.

7. DISCUSSION

The holy grail in health monitoring of aerospace structures is an on-line method which would use the ambient vibrations of the system to detect reliably structural faults. The novelty index may allow some progress in this direction.

In the first place, no model is required, either for the structure or the fault. However, the choice of feature for the novelty index is critical. The transmissibilities used in this report are attractive because they can be measured simply and they do not depend on the type of excitation to which the structure is subjected. They therefore offer the possibility of using ambient excitation. Note that the transmissibility used throughout this paper is chosen to detect damage between masses m_1 and m_2. It will be completely insensitive to damage between mass m_1 and ground, for example. In practice, a vector of novelty indices would be required in order to ensure coverage of the structure. Some might be based on transfer function measurements in order to detect faults in connections to ground. This is a disadvantage, as the system would then require a network of sensors, although the number needed would in principle be far fewer than the number of monitored parameters. A possible advantage of using a sensor network is that information might then be available on the probable location of faults. This would lift the fault diagnostic to level two in Rytter's classification.

It is important to establish what degree of change in the index is *significant*. Suppose that a strain-based index were defined for an aircraft. There are many manoeuvres which may cause wide fluctuations in such a diagnostic, so the neural network used to define the index would have to be trained with example patterns which adequately spanned the flight envelope, so that false positives would not occur during normal flight.

An important point is that the index may change for reasons other than damage; if an aircraft drops a store, the transmissibilities will change. Some system is needed which will ignore changes in the novelty index due to planned changes in the structure. If warning levels are being used, they should be reset after planned changes. To make the warning levels sensitive to more than one change in condition, it would be necessary to adopt an adaptive estimator of the parameter statistics. The simplest approach is to take a moving window for the mean and standard deviation estimates on which the warning levels are based. If a window of 100 points is used, the warning levels for the 1% fault in the low noise case are as shown in Figure 14. The statistics clearly reset themselves after the change. An alternative to the moving rectangular window is to include a forgetting factor in the calculation of the alarm level; this would effectively induce a moving exponential window with decay constant proportional to the factor. In practice, there is always a trade-off between window length and the speed with which the alarm level responds; this is discussed in some detail in reference [32]. In flight, there will be a gradual change in the parameters of an aircraft due to burning of fuel; the novelty index and warning levels should be insensitive to this. In Figure 15 are shown the warning levels for the 1% fault data at low noise, when a linear trend has been added. The warning level is still triggered.

Finally, it was mentioned above that the novelty index need not necessarily be simply correlated with the extent of damage. For the simulations discussed here, the contrary in Figure 16 is shown. It may simply be fortuitous that the index appears to be a linear function of the damage. However, under such circumstances (even under more general conditions where the index is monotonically increasing with damage), it might be possible to calibrate the index–damage curve by using FE analysis or experimental data. This might give the diagnostic some level-three capabilities.

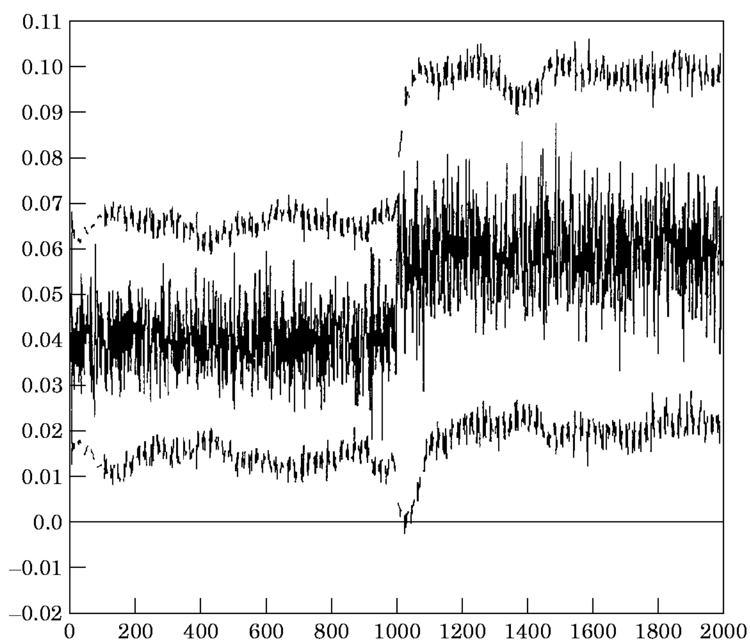

Figure 14. Warning levels for the novelty index evaluated on low noise data: training set and 1% fault data. Levels computed by using a 100-point moving window.

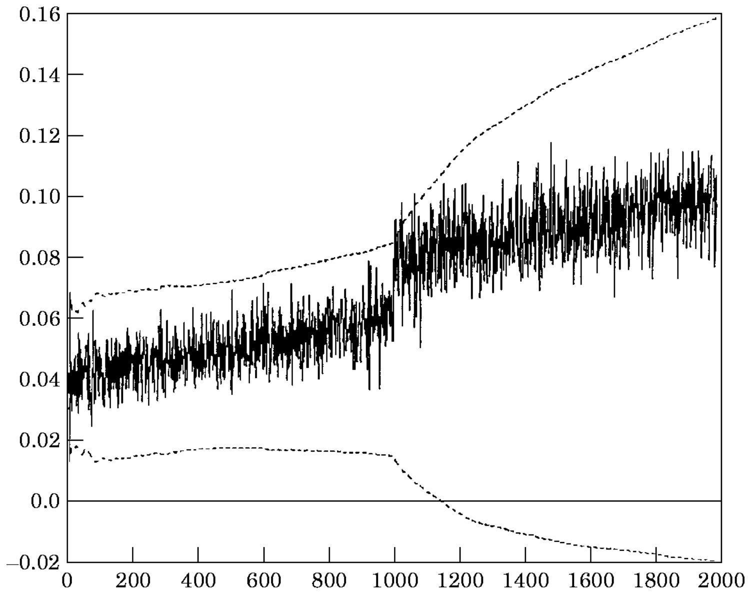

Figure 15. Warning levels for the novelty index evaluated on low noise data: training set and 1% fault data. Linear trend added to data.

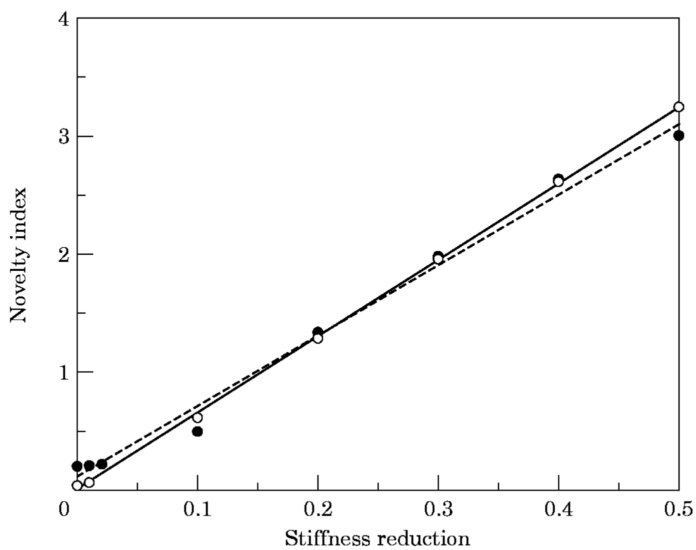

Figure 16. The variation of the novelty index with the extent of damage. ○, Novelty index, low noise; ——, linear regression, low noise; ●, novelty index, high noise; – – –, linear regression, high noise.

8. CONCLUSIONS

It has been shown how novelty detection methods can be used to diagnose damage in systems. The concept of warning and alarm levels can be adapted profitably from the field of condition monitoring to give automatic signalling of failure.

ACKNOWLEDGMENTS

The author would like to acknowledge a useful—indeed vital—discussion with Dr Peter

Cowley and Ms Rachel Pearce of Rolls-Royce and Associates, Derby. Discussions with Mr Martin Darkins of British Aerospace, Warton and Mr Andrew Ball of British Aerospace, Farnborough, proved useful, as did conversations with Dr Wieslaw Staszewski and Professor Geoff Tomlinson of the Department of Mechanical Engineering, University of Sheffield. Dr Q. Chen of the Control Technology Centre, University of Manchester, suggested transmissibility as an appropriate feature for the novelty index. Finally, Dr Lionel Tarassenko of the Department of Engineering Science, University of Oxford helped in navigating through the literature.

REFERENCES

1. C. M. Bishop 1995 *Neural networks for Pattern Recognition*. Oxford: Oxford University Press.
2. A. Rytter 1993 *Ph.D. Thesis, Department of Building Technology and Structural Engineering, University of Aalborg, Denmark*. Vibration based inspection of civil engineering structures.
3. R. Ceravolo, A. De Stefano, M. Grosjacques and D. Sabia 1995 *ASME Design Engineering Technical Conference* 1127–1133. Experimental identification of structural damage by neural networks.
4. M. F. Elkordy, K. C. Chang and G. C. Lee 1993 *Americal Society of Civil Engineers, Journal of Computing in Civil Engineering* **7**, 130–145. Neural network trained by analytically simulated damage states.
5. W. J. Leath and D. C. Zimmerman 1993 *Damage and Control of Large Structures, proceedings of the 9th VPI&SU Symposium* 583–594. Analysis of neural network supervised training with application to structural damage detection.
6. R. Manning 1994 *AIAA/ASME Adaptive Structures Forum* 160–172. Damage detection in adapative structures using neural networks.
7. P. H. Kirkegaard and A. Rytter 1994 *Proceedings of the 12th International Modal Analysis Conference, Hawaii*, 1128–1134. Use of neural networks for damage assessment in a steel mast.
8. P. H. Kirkegaard and A. Rytter 1994 *Proceedings of the 1st Workshop of the European Group for Structural Engineering Applications of Artificial Intelligence, Lausanne*. Vibration based damage assessment of civil engineering structures using neural networks.
9. P. H. Kirkegaard and A. Rytter 1994 *Proceedings of 10th International Conference on Experimental Mechanics, Lisbon*. Vibration based damage assessment of a cantilever using a neural network.
10. J. Kudva, N. Munir and P. W. Tan 1992 *Smart Materials and Structures* **1**, 108–112. Damage detection in smart structures using a neural networks and finite-element analysis.
11. C. Povich and T. Lim 1994 *AIAA/ASME Adaptive Structures Forum* 151–159. An artificial neural network approach to structural damage detection using frequency response functions.
12. J. Rhim and S. Lee 1994 *AIAA/ASME Adaptive Structures Forum* 173–180. A neural network approach for damage detection in structures.
13. W. Spillman, D. Huston, P. Fuhr and J. Lord 1993 *SPIE Vol*. 1918 *Smart Sensing, Processing and Instrumentation*, 288–295. Neural network damage detection in a bridge element.
14. J. E. Stephens and R. D. Van Luchene 1994 *Microcomputers in Civil Engineering* **9**, 119–128. Integrated assessment of damage in structures.
15. P. Tsou and M.-H. H. Shen 1994 *American Institute Aeronautics and Astronautics Journal* **32**, 176–183. Structural damage detection and identification using neural networks.
16. K. Worden, A. D. Ball and G. R. Tomlinson 1993 *Smart Materials and Structures* **2**, 189–200. Fault location in a framework structure using neural networks.
17. K. Worden and G. R. Tomlinson 1994 in *Engineering Integrity Assessment* (J. H. Edwards, J Kerr and P. Stanley, editors) 11–31. Damage location and quantification using neural networks.
18. X. Wu, J. Ghaboussi and J. H. Garrett 1992 *Computers and Structures* **42**, 649–659. Use of neural networks in detection of structural damage.
19. P. Z. Szewczwk and P. Hajela 1994 *American Society of Civil Engineers Journal of Computing in Civil Engineering* **8**, 163–178. Damage detection in structures based on feature-sensitive neural networks.
20. G. G. Yen and M. K. Kwak 1993 *American Institute Aeronautics and Astronautics* 931485- *CP*, 1549–1555. Neural network approach for the damage detection of structures.

21. G. L. Zheng and S. A. Billings 1993 *Research Report, Department of Automatic Control and Systems Engineering, University of Sheffield*. Fault diagnosis using radial basis function neural networks.
22. Y. Teboub and P. Hajela 1992 *Proceedings of the 4th AIAA/USAF/NASA/OAI symposium on Multidisciplinary Analysis and Optimisation, Cleveland, OH*. A nerual network based damage analysis of smart composite beams.
23. L. Tarassenko, P. Hayton, N. Cerneaz and M. Brady *Preprint, Department of Engineering Sciences, University of Oxford*. Novelty detection for the identification of masses in mammograms.
24. S. Roberts and L. Tarassenko 1994 *Neural Computation* **6**, 270–284. A probabilistic resource allocating network for novelty detection.
25. D. A. Pomerleau 1993 in *Advances in Neural Information Processing Systems* **5** (S. J. Hanson, J. D. Cowan and C. l Giles, editor). Morgan Kaufman. Input reconstruction reliability estimation.
26. C. M. Bishop 1994 *IEE Proceedings on Vision and Image Signal Processing* **141**, 217–222. Novelty detection and neural network validation.
27. D. E. Rumelhart and J. L. McClelland 1988 *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (two volumes). MIT Press.
28. Q. Chen, Y. W. Chan, K. Worden and G. R. Tomlinson 1994 *Proceedings of the International Conference on Vibration Engineering, Beijing, China*, 567–757. Structural fault detection using neural networks trained on transmissibility functions.
29. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling 1986 *Numerical Recipes—The Art of Scientific Computing* Cambridge: Cambridge University Press.
30. C. Cempel 1985 *Maintainence Management International* **5**, 297–304. Determination of vibration symptom limit values in diagnostics of machinery.
31. A. D. Ball 1991 *Ph. D. Thesis, Department of Mechanical Engineering, University of Manchester*. Machine condition assessment and fault diagnosis.
32. A. D. Ball, J. E. Cooper and K. Worden 1992 in *Structural Integrity Assessment*, (P. Stanley, editor) 236–245. Elsevier. On-line system identification techniques for structural analysis and monitoring.