Genetic optimisation of a neural network damage diagnostic

G. Manson, E. Papatheou and K. Worden
Dynamics Research Group, Department of Mechanical Engineering
University of Sheffield
Sheffield, UK

ABSTRACT

This paper presents an automated optimisation procedure for the feature selection stage of a previously proposed structural health monitoring methodology using a genetic algorithm. The same diagnostic is used in the attempt to progress up the levels of damage detection to location and severity. It was validated experimentally on a Gnat aircraft wing. An artificial neural network is used as a classifier and the work is compared with the previous selection strategy based on engineering judgement.

NOMENCLATURE

$Y_1(\omega)$ Acceleration frequency spectrum recorded at the reference accelerometer

$Y_2(\omega)$ Acceleration frequency spectrum recorded at the non-reference accelerometer

1.0 INTRODUCTION

The safety and economic benefits resulting from a robust damage identification strategy are numerous. Nowadays, it is commonplace for many civil and aerospace structures to continue to be used beyond their intended operational life which will often result in lengthy inspection procedures. These operations may account for a large proportion of the cost of ownership of a structure. Any approach which may increase the time between inspections or may decrease the length of an inspection will assuage this financial burden. At present the inspection methods most commonly used by industry fall within the area of nondestructive evaluation (NDE)[1] which includes such techniques as ultrasonic inspection, radiography, dye penetration and acoustic emission, among others. These methods usually require *a priori* specification of an area of interest, meaning that they work on a local scale. A large degree of expertise is also usually required in the identification of the areas of interest and in the interpretation of the resulting data. It is also generally the case that it is required that the structural areas of interest be accessible. All in all, applying these inspection techniques on structures of any great size generally results in very lengthy inspections which may be directly translated into economic terms.

It is against the backdrop discussed in the previous paragraph that the highly active research field of structural health monitoring (SHM) should be viewed. It would be very desirable to owners of high-cost structures if reliable, robust, global and intelligent health monitoring systems could be developed to reduce inspection times without jeopardising safety. In general the SHM approaches to identifying damage have concentrated upon changes to the dynamic properties of the system. Doebling *et al*[2] have published an extensive literature review of such vibration-based methods. According to Pai and Jin[3] they can be categorised into four main domains; namely spatial, modal, time and frequency-domain. Spatial domain methods use changes in mass, damping and stiffness matrices while the modal domain methods use changes in natural

Paper No. 3198. Manuscript received 5 April 2007, revised version received 10 December 2007, accepted 13 December 2007.

https://doi.org/10.1017/S0001924000002219 Published online by Cambridge University Press

frequencies and mode shapes. Time domain methods examine changes in input and output signals and finally the frequency domain methods use changes in frequency responses functions (FRFs) or transmissibility functions (TFs).

Although damage detection can be a very difficult problem, it is generally considered to have a hierarchical structure, starting from the simple detection of damage through to an estimation of the residual life of the structure. Perhaps the most well-known such framework is that of Rytter[4], with four levels of damage identification; namely, Level 1 – Detection, Level 2 – Localisation, Level 3 – Assessment and Level 4 – Prognosis. A series of papers from two of the authors has previously proposed a method for a health monitoring system, based on novelty detection techniques, where an attempt was made to progress up the levels of the hierarchy[5-8]. The whole idea was validated experimentally on a Gnat aircraft wing in a body of research funded by DERA/QinetiQ.

By viewing the problem of damage identification as a one of pattern recognition, an organising principle to progress from the acquisition of raw data to decision making is required. It is arguable that the feature extraction stage is the most crucial part of such an attempt. More specifically, in the context of novelty detection, a feature would be some set of values derived or calculated from measured (or pre-processed) data which will distinguish between different data classes. Again, a previous attempt to optimise this phase achieved excellent results[9] although the optimisation phase was preceded by the engineering judgement feature selection phase outlined in[7]. The current work aims to fully automate the feature selection process through a genetic algorithm (GA) optimisation framework.

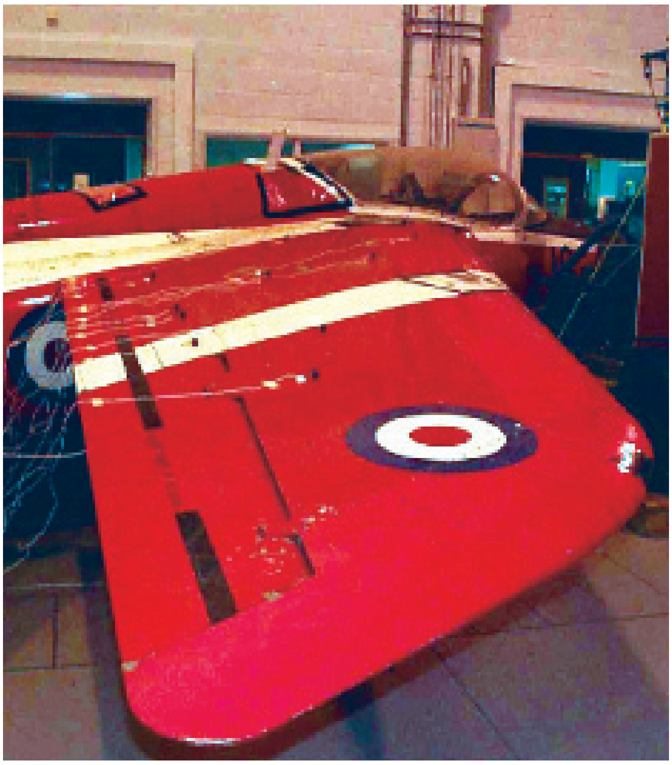
Genetic algorithms have been used successfully before in optimisation problems such as sensor placement[10] and feature selection for machine condition monitoring[11]. The work here centres on the feature selection stage of a previously proposed methodology aimed at Levels 2 and 3 of the damage identification hierarchy, namely damage location and damage severity.

The layout of the paper is as follows: the next section considers the experimental set-up and discusses the data acquisition process. This is then followed by sections on feature selection for the damage location problem, the first using engineering judgement and the second using genetic algorithms. The next two sections follow the same approach for the damage severity problem. The paper is completed with some discussion and conclusions.

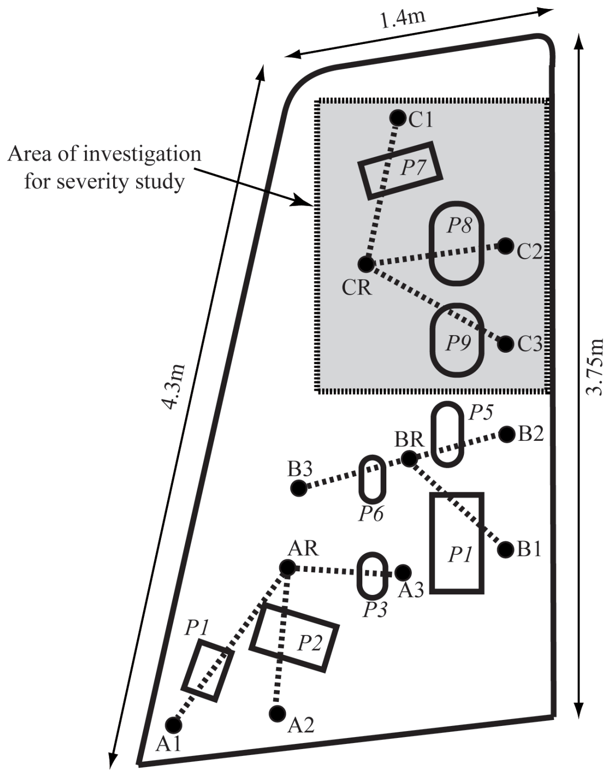
2.0 TEST SET-UP AND DATA CAPTURE

The experimental set-up and data capture procedure was very similar for both the damage location and the damage severity problems. The structure under investigation was a Gnat trainer aircraft and in particular its starboard wing as shown in Fig. 1(a). The fact that the actual structure was not allowed to be damaged led to the simulation of a fault as the sequential removal of inspection panels. Clearly, removal of inspection panels will not result in the same type of changes in the damage signals as the changes observed when true damage is present. That said, this approach allowed the structural health monitoring strategy to be demonstrated and it also had the advantage that each damage scenario was reversible and monitoring the repeatability of the measurements was feasible. Nine panels were used for the experiment and their choice was mainly based on their ease of removal and also the range of cover sizes.

The distribution of the inspection panels can be seen in the schematic of Fig. 1(b). The actual panel areas are given in Table 1. Panels P3 and P6 were deemed likely to be the most problematic as they are by far the smallest. Each panel was fixed to the wing by a number of screws, the number varying between 8 and 26. The screws were secured and removed with an electric screwdriver with a controllable torque. In the attempt to control variability in the fixing conditions the same torque setting was used throughout the test.



(a)



(b)

Figure 1. (a) Photograph and (b) schematic of Gnat trainer aircraft starboard wing showing numbering of inspection panels and transducers. The panel area that was used for damage severity problem is highlighted.

https://doi.org/10.1017/S0001924000002219 Published online by Cambridge University Press

Table 1
Area of inspection panels

Panel	Area (m^2)
P1	0·022
P2	0·050
P3	0·008
P4	0·080
P5	0·018
P6	0·008
P7	0·040
P8	0·047
P9	0·023

The base experimental measurements were transmissibilities across each panel as they had previously proved effective for the problems of novelty detection[6] and the extension to damage location[7] and damage severity[8]. The panels were split into three groups: A, B and C. Each group was allocated a centrally placed reference accelerometer (denoted AR, BR and CR respectively in Fig. 1b), together with three other accelerometers (A1, A2, A3 etc.). This resulted in their being a transmissibility *path* between a reference accelerometer and an associated accelerometer for each inspection panel. Note, the transmissibility is defined as the ratio $\frac{Y_1(\omega)}{Y_2(\omega)}$ 0, where $Y_1(\omega)$ (resp. $Y_2(\omega)$) is the acceleration frequency spectrum recorded at the reference accelerometer (resp. the acceleration frequency spectrum recorded at the other accelerometer). Although this network of accelerometers had the potential of forming many transmissibilities, in this study only those measured directly across each plate were used. The accelerometers used were standard piezo-electric accelerometers manufactured by PCB. For the damage location problem, the removal of all nine inspection panels was considered and measurements were recorded using all three trans-missibility groups, A, B and C. The damage severity problem, however, only considered panels P7, P8 and P9 and measurements were only recorded using the group C accelerometers.

In order to obtain the transmissibility measurements, the wing was excited using a Ling electrodynamic shaker attached directly below panel P4 on the bottom surface of the wing. A white Gaussian excitation was generated within the acquisition system and amplified using a Gearing and Watson power amplifier.

All the measurements were acquired using a DIFA Scadas 24-channel acquisition system controlled by LMS software running on a HP workstation. In all cases 1,024 spectral lines were measured between the frequencies 1,024 and 2,048Hz. Both real and imaginary parts of the function were obtained, however these were converted to magnitudes for feature selection and the phases were discarded.

2.1 Data acquisition for damage location

In total, 25 sets of measurements were made, two sets for each damage state (specific panel removal) and seven for the normal condition. This was done in order to investigate variability of the normal and damaged condition data. In all cases, each transmissi-bility was first obtained using 16 averages (to provide a clean reference to help with feature selection). Next, 100 measurements were taken sequentially using only one average. Over the full sequence of 25 configurations, this gave 700 one-shot measurements for the normal condition and 200 for each damage condition.

2.2 Data acquisition for damage severity

As previously mentioned, in the damage severity problem, damage was only investigated in panels P7, P8 and P9 and measurements were only recorded using the group C accelerometers. This is shown as the highlighted area in Fig. 1(b). In order to investigate damage severity in the absence of permission to actually damage the aircraft, it was decided that several copies of each of the three panels be made and damage introduced into the copies in a controlled manner. Each panel was reproduced four times. One of the copies was left undamaged to act as a reference, while the other three were subjected to three levels of damage: one eighth, one quarter and one half of the panel area was excised (symmetrically). During the experimental phase, data were also recorded for a complete removal of each panel, giving four different damage states for each panel. There are thus 12 different severity/location combinations for the damage which were labelled D1 to D12 (D1 to D4 refer to increasing damage in panel P7 etc.).

In all, 34 tests were carried out – ten normal condition tests were conducted and two tests for each of the 12 different severity/location combinations. Each test comprised 121 measurements: 120 one-shot transmissibility measurements followed by one 16-average transmis-sibility. Therefore, in total 240 one-shot transmissibilities were obtained for each damage state and 1,200 were taken for normal condition. This was considered a large enough sample to allow a principled statistical and neural analysis. The 16-average measure-ments were again used to try and identify regions of the frequency domain, which appeared to be sensitive to damage i.e. for feature selection. For the purposes of the current work, only the damage data D1 to D4 pertaining to damage in panel P7 was used.

The planned strategy had been to use the single-shot transmissibility functions obtained from the network of accelerometers on the structure to construct novelty detectors. However, due to excessive noise on these measurements, the novelty detectors did not perform sufficiently well. For this reason, and to have a sufficient amount of data for training the novelty detectors and avoid neural network generalisation problems, a bootstrapping technique was adopted to construct 120 16-average transmissibility measurements for each transmissiblity for each of the test conditions. The procedure was as follows:

1. For the first transmissibility (between accelerometers CR and C1) for the first normal condition set, randomly select 16 single-shot measurements from the 120 available and construct a 16-average measurement.
2. Repeat to give 120 16-average measurements.
3. Repeat process for the second and third transmissibilities (i.e. between CR and C2 and between CR and C3).
4. Repeat process for all other normal and damage conditions.

3.0 FEATURE SELECTION BASED ON ENGINEERING JUDGEMENT – DAMAGE LOCATION PROBLEM

In the original damage location analysis[7], the first stage was to establish which features could be used to individually detect damage in the plates. The detection algorithm used was the outlier analysis procedure described and validated in Refs 5 and 6. The main idea of the method is that after constructing a statistical model of the normal data, subsequent data will be tested to see if they are statistically consistent with the normal data. Inconsistent data vectors are taken to infer damage. More details may be found in the standard reference[12].

A feature was a region of the given transmissibility which separated unambiguously the normal condition data from the damage data. It was also decided that features should be classified as strong, fair or weak according to the following criteria:

- A strong feature is a region of the frequency range on which the normal data and damage data appear to be structurally different. Also, the damage data should be strongly separated compared to the spread of the normal data. See Fig. 2(a).

https://doi.org/10.1017/S0001924000002219 Published online by Cambridge University Press

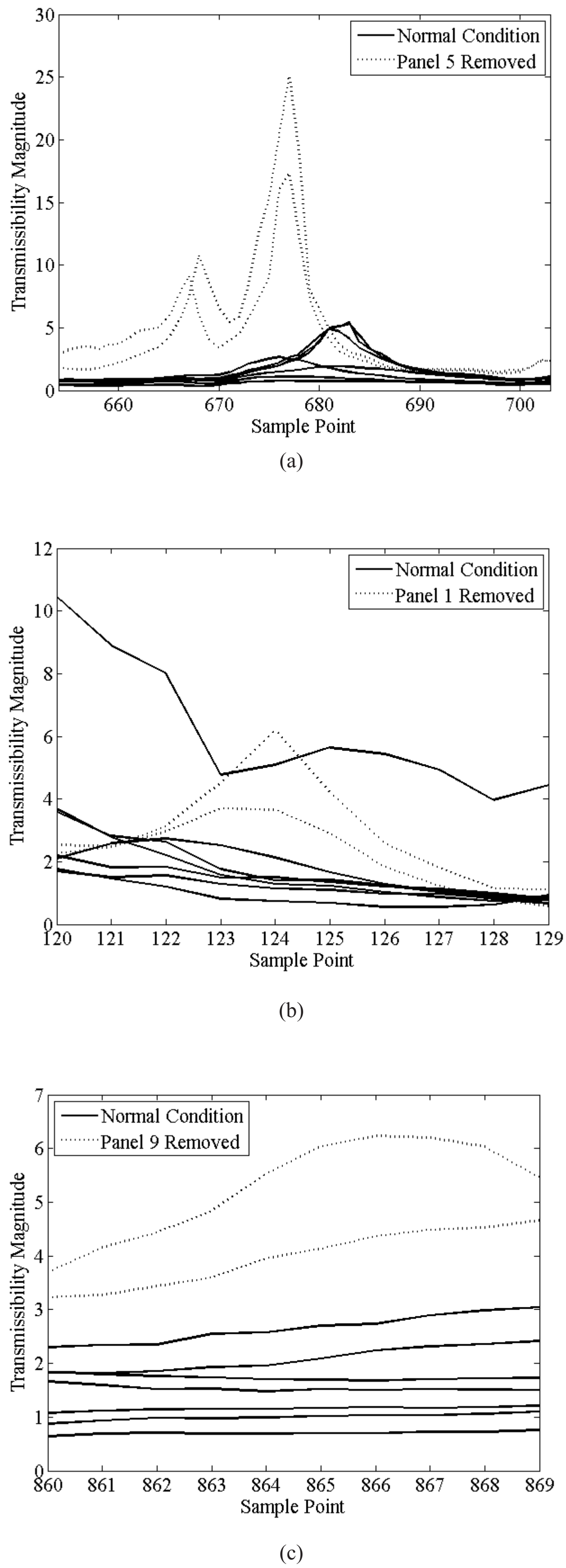


Figure 2. Examples of (a) a strong feature, (b) a fair feature and (c) a weak feature from different transmissibility functions.

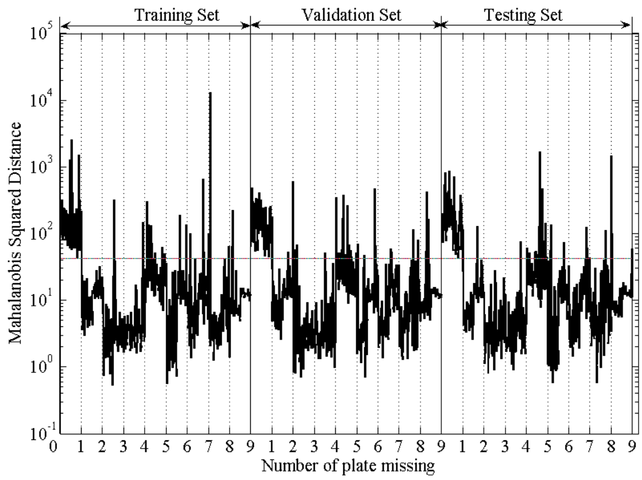


Figure 3. Outlier statistics for all damage states for the novelty detector trained to recognise panel 1 removal.

- A fair feature is a region of the frequency range on which the normal data and damage data appear to be structurally different or the damage data are strongly separated compared to the spread of the normal data. See Fig. 2(b).
- A weak feature is a region of the frequency range on which the normal data and damage data are separated. See Fig. 2(c).

In total, 44 features were assessed as strong or fair. For each of these features, a novelty detector was constructed using outlier analysis. Each detector gave a scalar variable or novelty index which could be tested against a threshold to indicate damage if it exceeded it. Figure 3 shows the novelty indices over all of the damage states from a novelty detector which was selected for its performance in identifying the removal of P1. The horizontal dashed line is the 99% confidence in identifying an outlier, calculated using a Monte-Carlo approach as outlined in Ref. 5. The *x*-axis labels the damage state and there are three repetitions as the data were divided into training, validation and testing sets for the neural network training.

The ultimate purpose of the whole procedure was to produce a damage location system. The algorithm chosen was a Multi-Layer Perceptron (MLP) neural network[13]. The idea was simply to map a set of the novelty indices obtained from the transmissibilities to the damage location (here, to which panel was removed). Nine novelty indices were formed from the best nine features. Although this selection process was informed by statistical evidence, it cannot be called objective as the chosen features were a subset selected from a larger feature set which were chosen based upon engineering judgement.

The outputs from the nine novelty detectors were used for the training of the neural network following the guidelines recommended in Ref. 13. The data were, as previously mentioned, divided equally into training, validation and testing set. The network structure consisted of nine input nodes (one for each novelty index) and nine output nodes (one for each damage class). The number of hidden nodes was determined during training using a process of cross-validation. In order to estimate the classification accuracy of each network the one of M strategy was employed. This means that each class was assigned a specific network output[13]. By using this strategy, if the network is presented with a test vector, the appropriate class is assigned after finding the highest output. With this procedure however, the network has in a sense been tuned to both the training and validation sets (as a part of the number of hidden units finding process) and therefore an independent testing set was required for proper verification of the network.

It transpired that, when the networks were trained, the one that gave the lowest validation error had ten hidden units and gave a misclassification error of 0·155, corresponding to a training error of 0·158. The best results were obtained after 150,000 presentations of data. The network weights were updated after each presentation, i.e. the training epoch was one. When the best network was tested, it gave a generalisation error of 0·135, i.e. 86·5% of the patterns were classified correctly. As anticipated, most errors were associated with panels 3 and 6 (the smallest), with some additional confusion between panel 8 and 9 removals. The results were very encouraging, but there was certainly room for improvement so a further attempt was conducted[9]. This time a genetic algorithm was used to select the best nine features out of the total 44 pre-selected, i.e. the GA performed a combinatorial optimisation. The best classification rate achieved was 98 % in the testing set. In the next section, this work will be extended to remove the engineering judgement entirely from the feature selection process. Instead, a genetic algorithm will be employed to choose the features from the transmissibility data.

4.0 FEATURE SELECTION USING A GENETIC ALGORITHM – DAMAGE LOCATION PROBLEM

For a more detailed analysis of genetic algorithms (GAs), the reader is referred to Refs 14 and 15. Briefly speaking, they are algorithms that imitate the mechanism of natural selection, meaning that out of a random initial 'population' of solutions, the 'fittest' have a greater chance of propagating all or part of their 'information' to subsequent generations. Genetic operations such as crossover and mutation are generally used to exploit good solutions and to explore new solutions respectively. In their most traditional form, each potential solution is encoded with binary bit-strings. To many practitioners, the structure of the genetic algorithm (GA) is fixed and the problem should be modified, if necessary, to conform to this structure. It is the authors' belief, in agreement with Michalewicz[15], that the problem should not be modified to fit the algorithm, but rather the genetic algorithms themselves should be modified to suit the problem at hand, leading to what Michalewicz terms "evolution programs". As far as the authors are concerned, these algorithms are simply a means to an end, namely, seeking the optimal value of some function and may be altered in any way which improves this process.

In the current problem, the most suitable encoding of the problem was to use strings of integer numbers. Each string was made up of nine parts with each part representing one of the nine novelty indices. Each of these parts consisted of three integer numbers: the first integer indicated which of the 18 (9 normal and 9 reciprocal) transmissibility functions was to be used to form the feature, the second integer specified the starting point (between one and 974) within the transmissibility function for the feature and the third integer specified the feature length (either 10, 20, 30, 40 or 50 lines represented as an integer between one and five). One final integer was included after the 27 (9 sets of 3) feature integers which was an integer between one and ten to control the random initialisation of the neural network weights. Following the attempt in the previous work[7] to simplify the feature selection process by separating the removed panels into groups of three (see Fig. 1(b)) and selecting the novelty indices from the relevant transmissibility functions only, the first of the three integers (controlling the transmissibility) was limited to 1 – 6, 7 – 12 and 13 – 18 in each group respectively.

The fitness value for each chromosome was simply the final classification rate in the validation set after the neural network training. In order for the potential parents to be selected, a roulette-wheel approach was used along with single point crossover[14]. Retaining the elite chromosome was another important aspect of the evolution program (in each generation the best chromosome was compared to the previous best one and changed or saved accordingly).

The mutation operator was the factor which departed most from a conventional genetic algorithm. Since each chromosome contained the encoded information about the three feature parameters, namely transmissibility number, starting point and window size, and in the initialisation integer, it was decided that four different mutation operators should be employed. A fifth one was added later after preliminary results showed it could be useful. In detail, the transmissibility function number and the window size were 'mutated' with a adjustment. The same procedure was followed for the seed. Larger mutations (subsequently referred to as 'long jump') were permitted for the starting point was as this was an integer between 1 and 974. The fifth mutation also acted on the starting point, but with only a slight modification of between –10 and +10 (subsequently referred to as 'short jump').

The optimisation was run several times. The final GA settings that were used were population size equal to 20, crossover probability 0·8, length, seed and transmissibility mutation probability 0·2, starting point mutation 'long jump' 0·1 and 'short jump' 0·4. The standard MLP neural network was trained with 25 cycles. Its structure was fixed to nine nodes in the input and output layer and with ten hidden units, the same structure as selected in the Engineering judgement approach. The software package that was used to train the neural network was NETLAB[16] and not the same in-house package as in Refs 7 and 9. The optimisation results proved to be very encouraging for such a large optimisation problem with a 100 generation run often returning a classification rate of 93 to 94% on the validation set. This is still slightly lower than the 99% found in the validation set of the combined engineering judgement and optimisation approach used in Ref. 9, but a lot better than 86·5 % from the engineering judgement only approach used in[7]. However, using the above GA settings and allowing the algorithm to run for 300 generations (approximately one hour on current PC) returned validation classifications rates around 95 to 96%. This is mainly a result of the low number of the neural network training cycles, which avoids overtraining. Finally, when the Evolution Program was left running for 689 generations, it was able to achieve a validation classification rate of 98·3%, as shown in Fig. 4, which returned a corresponding testing set classification rate of 96·9%.

It is not a straightforward matter to assess which of the nine features was predominantly associated with the neural network correctly classifying a certain panel removal. The features which were selected by the genetic algorithm to give the 98·3% validation classification rate were examined. Normal data was compared to data from each of the panel removals and a subjective label of strong, fair or weak were attached to each panel removal/feature combination, according to the criteria discussed in the previous

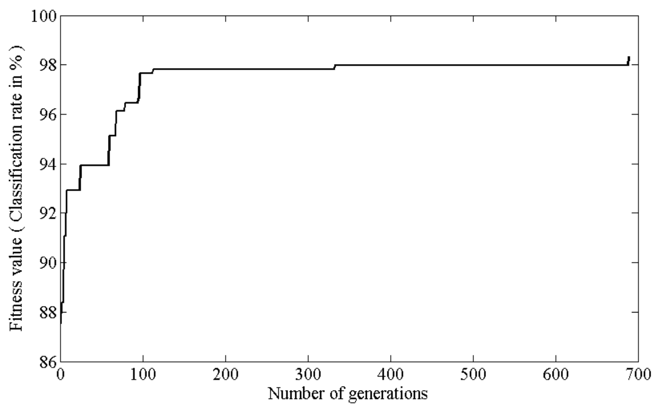


Figure 4. Fitness over number of generations in a run when the GA was not restricted to a specified number of generations. After 689 generations the algorithm reached a 98·3% classification rate.

https://doi.org/10.1017/S0001924000002219 Published online by Cambridge University Press

Table 2.
Features that were selected by the GA (with 98·3 % in Fig. 4) to detect each of the nine panels. The * denotes reciprocal. Symbols F, S and W mean fair, strong and weak respectively.

		Feature Number (Transmissibility Function in brackets)								
		1 (1*)	2 (3)	3 (1*)	4 (6*)	5 (4*)	6 (4)	7 (8*)	8 (8)	9 (8)
Panel	*P1*	F								
	P2			F/S						
	P3								W	
	P4					S				
	P5		F/S	F/W				S		
	P6									
	P7									
	P8							F/S		
	P9			F						

section. Table 2 shows which features seem to relate to which panel removal. As may be expected, in many cases there is no single strong feature and the neural network achieves its goal via combinations of multiple features. That said, feature three seems to be important in detecting the removal of panels P2, P5 and P9.

To summarise, a genetic-algorithm approach to feature selection has resulted in significant improvement in classification rates over the engineering judgement approach and almost as high a classification rate as the Engineering judgement/combinatorial optimisation approach. The most important aspect of the GA approach is the fact that it avoids the extremely time-consuming visual selection of features without compromising on performance. In the next two sections, the technique will be extended to the problem of damage severity.

5.0 FEATURE SELECTION BASED ON ENGINEERING JUDGEMENT – DAMAGE SEVERITY PROBLEM

This section consider the features which were selected to assess damage severity in panel P7 (as shown in Fig. 1b) as part of a previous study[8]. In that study, the issue of damage severity was considered for all three panels, P7, P8 and P9, in the highlighted region of Fig. 1(b). It was found that the greatest testing set error occurred with panel P7 and so panels P8 and P9 will not be discussed further in this work.

Five features were highlighted from transmissibility between accelerometers CR and C1 as being potential features for detecting damage in panel P7 using Engineering Judgement. These were spectral lines 71-80, 110-119, 558-577, 880-919 and 930-949 of the 1,024 line transmissibility plot. Novelty detectors were formed for each of these features by using the 120 16-average bootstrapped measurements from the first eight normal condition sets as training data. The other two normal condition sets and all the damage condition sets were used to form the testing sets. Outlier analysis was again used to detect when an observation in the testing set was statistically unlikely to have been generated by the same mechanism as those in the training set. Once the damage was successfully located using the damage localisation classifier, the outputs from the five novelty detectors were then passed to the level three neural network to give a measure of the damage severity.

The network was a multi-Layer perceptron (MLP) neural network. The idea was to try to map the novelty indices obtained from the transmissibilities to the damage severity. The neural network is supplied with the values of the five novelty indices at the input layer and required to predict the damage severity at the output layer.

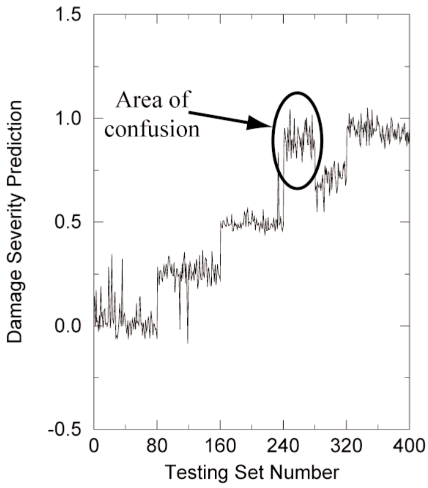


Figure 5. Original neural network prediction of damage severity for panel P7 using features selected by engineering judgement.

For the network structure, the input layer had five nodes, one for each novelty index, and the output layer had a single node, for the severity value (0 indicating no damage and 1 indicating the panel is missing). The number of hidden nodes for each case was determined during training as for the location. Instead of the network structure being determined based on the lowest misclassification error, it was based upon the lowest mean squared error between the predicted damage severity and the actual damage severity.

The network with the lowest mean squared error on the validation set (4·7%) had five hidden units with the best results obtained after 120,000 presentations of data. which returned a mean square error of 5·6% on the testing data set. The neural network prediction is shown in Fig. 5. As can be seen, the first repetition of the 0·75 damage case has returned severity values close to 1·0 indicating that the network is confusing this with the plate missing situation. It should be noted that the damage conditions have been relabelled 0·25, 0·50 and 0·75 in order to spread them out evenly. This amounts to a nonlinear transformation and may not be the most satisfactory.

At this point, it should be mentioned that the neural network training was conducted using an in-house package which used gradient descent to perform the weight optimisation. To provide a fair comparison with the features selected by genetic algorithm, the outputs from the novelty detectors trained on the Engineering judgement features were used as inputs to a multi-layer perceptron neural network which optimised the network weights using the superior scaled conjugate gradient technique. This was conducted using the NETLAB software package[16]. During this study it was found using the logarithmic values of the novelty indices resulted in lower network errors and so this will be repeated for the engineering judgement features described above.

A total of ten runs were conducted with different random initialisation seeds and the best run gave a validation set mean square error of 1·89% and a corresponding testing set error of 1·45%. The neural network prediction for the testing set is shown in Fig. 6. A significant improvement from Fig. 6 may be observed, especially in the first repetition of the 0·75 damage.

In the next section, it will be seen if the results shown in Fig. 6 may be improved upon by conducting the feature selection process using a genetic algorithm.

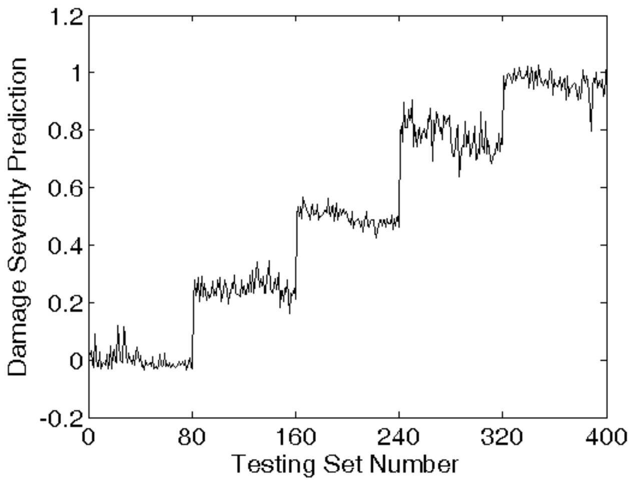


Figure 6. NETLAB neural network prediction of damage severity for panel P7 using logarithmic valued features selected by engineering judgement.

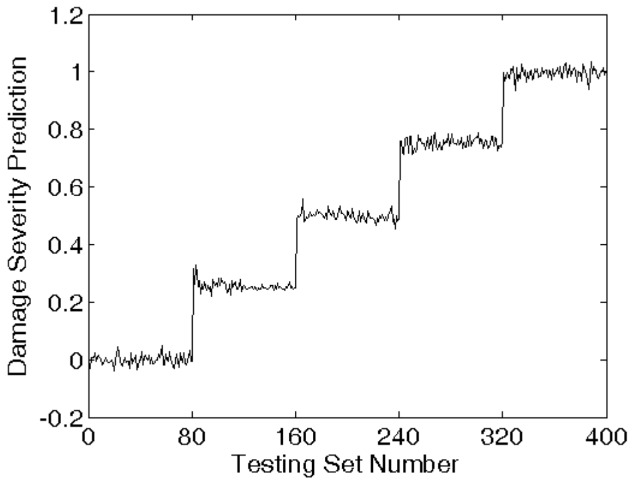


Figure 7. NETLAB neural network prediction of damage severity for panel P7 using logarithmic valued features selected by 100 generation genetic algorithm.

6.0 FEATURE SELECTION USING A GENETIC ALGORITHM – DAMAGE SEVERITY PROBLEM

In the previous section, a significant improvement was observed in the severity prediction capability simply by using logarithmic novelty indices and a more sophisticated neural network package. That said, there is still room for improvement and, in this section, the genetic algorithm approach to feature selection will be investigated to see if this may further improve the severity prediction capabilities of the technique.

For the problem at hand, the issue of encoding potential solutions was as detailed for the damage location problem discussed above with the simplification that the severity problem does not require the specification of transmissibility numbers (as there was only one transmissiblity function used). Each feature would be encoded by the starting spectral line integer and a window size (this was fixed to being 10, 20, 30, 40 or 50 lines). At the end of each chromosome another integer between one and ten was added which controlled the random initialisation seed for the neural network training. These meant a five feature chromosome would be encoded using 11 integers, five of these would be between one and 974 to represent the start line, another five which would be between one and five for the window size and the final integer between one and ten for the network seed.

On this occasion, the fitness values were calculated by taking the reciprocal of the mean square error of the validation set prediction. Parental selection was conducted using a weighted roulette-wheel approach and single-point crossover was used. Options of the injection of new blood and retaining the elite chromosome were also included. Rather than use the usual single mutation operator, the authors have chosen again to employ multiple mutation operators, as in the damage location case.

The optimisation was run five times using the logarithmic novelty index data. The GA settings were population size = 20, 100 generations, 0·6 crossover probability (single-point), 0·2 for each of four mutations, 20% new blood and retention of the elite chromosome. The run which gave lowest validation error of 0·25% returned a chromosome with very different features (lines 46-85, 117-166, 176-215, 613-662 and 925-944) from those chosen using engineering judgement. A mean square error of 0·26% was returned when the testing data set was propagated through the neural network trained on these features and Fig. 7 shows the neural network prediction for the testing set. The lower error is evident in the more 'staircase-like' structure of the plot.

Due to time considerations, a relatively small number of optimisation runs were conducted with subjectively chosen parameters. It was comforting to note however, that the mean validation error of the five runs was 0·28% and the largest error of the five runs was 0·38%, all a significant improvement on the 1·89% using the Engineering judgement features. That said, there is probably room for further improvement, if it were desired – this has been confirmed by validation error of 0·17% during a run where 500 generations were used.

7.0 DISCUSSION AND CONCLUSIONS

A structural health monitoring strategy is often envisaged using a data to decision flowchart. Within this process the feature selection stage seems to be arguably the most crucial. Therefore it requires significant time and effort and it is generally conducted with a large degree of subjectivity. Genetic algorithms or evolution programs present a rather promising and high potential solution to optimisation problems. They also offer automation and satisfactory objectivity to this critical phase of a structural health monitoring programme.

The work in this paper showed their use on the same methodology for a damage diagnostic in three levels, detection, location and severity. The results were very good for the first two levels even though they did not exceed those found in[9]. However, here the features were selected by the GA from a completely random point. It is also important that it is unlikely that the GA parameters that were used are the optimal parameters, but they seemed robust to small changes. In a similar approach for the severity problem, a reduction in testing set error from 1·45%, using a network trained with features selected by engineering judgement, to 0·26%, with GA-selected features was found.

Throughout this work, Level 2 and 3 damage identification was conducted in a supervised manner (i.e. assuming that data pertaining to all damage states were available), as is necessary with neural network classifiers. In reality, one will not have access to damage data and therefore recourse may need to be made to high-fidelity models of the structure and relevant damage models. Alternatively, it may be possible to develop approaches of extracting 'damage-like' features from a structure without actually damaging the structure – it is towards overcoming this problem that the authors are concentrating a large degree of effort.

https://doi.org/10.1017/S0001924000002219 Published online by Cambridge University Press

In order to return to the present work, it should be stated that, if the field of structural health monitoring is to fully make the transition from the laboratory to real, complex structures, the key process of feature selection must be done in an automated, objective manner. It is hoped that the current work contributes to that goal.

REFERENCES

1. DOHERTY, J.E. Nondestructive evaluation, Chapter 12 in *Handbook on Experimental Mechanics,* 1987, KOBAYASHI, A.S. (Ed), Society for Experimental Mechanics.
2. DOEBLING, S.W., FARRAR, C.R. and PRIME, M.B. A summary review of vibration-based damage identification methods, 1998, *The Shock and Vibration Digest*, **30,** pp 91-105.
3. PAI, P.F. and JIN, S. Locating structural damage by detecting boundary effects, 2000, *J Sound and Vibration,* **231,** (4), pp 1079-110.
4. RYTTER, A. Vibration-Based Inspection of Civil Engineering Structures, PhD Thesis, 1993, Department of Building Technology and Structural Engineering, University of Aalborg, Denmark.
5. WORDEN, K., MANSON, G. and ALLMAN, D.J. Experimental validation of a structural health monitoring methodology, Part I: Novelty detection on a laboratory structure, 2003, *J Sound and Vibration,* **259,** (2), pp 323-343.
6. MANSON, G., WORDEN, K. and ALLMAN, D.J. Experimental validation of a structural health monitoring methodology, Part II: Novelty detection on a Gnat aircraft, 2003, *J Sound and Vibration,* **259,** (2), pp 345-363.
7. MANSON, G., WORDEN, K. and ALLMAN, D.J. Experimental validation of a structural health monitoring methodology, Part III: Damage location on a Gnat aircraft wing, 2003, *J sound and Vibration,* **259,** (2), pp 365-385.
8. MANSON, G., WORDEN, K. and ALLMAN, D.J. Experimental validation of a damage severity method, 2002, Proceedings of 1st European Workshop on Structural Health Monitoring, Paris, France, pp 845-852.
9. WORDEN, K., HILSON, K. and MANSON, G. Genetic optimisation of a neural Damage Locator, Submitted to *J Sound and Vibration*.
10. WORDEN, K. and BURROWS, A.P. Optimal sensor placement for fault detection, 2001, *Engineering Structures,* **23,** pp 885-901.
11. JACK, L.B. and NANDI, A.K. Genetic algorithms for feature selection in machine condition monitoring with vibration signals, 2000, IEE Proc – *Vis Image Signal Process,* **147**, (3), pp 205-212.
12. Barnett, V. and LEWIS, T. *Outliers in Statistical Data*, 1994, John Wiley and Sons, Chichester.
13. TARASSENKO, L. *A guide to Neural Computing Applications,* 1998, Arnold.
14. Goldberg, D.E. *Genetic Algorithms in Search, Optimization and Machine Learning,* 1989, Addison-Wesley Publishing, Reading, MA.
15. MICHALEWICZ, Z. *Genetic Algorithms + Data structures = Evolution Programs,* 1996, Springer.
16. NABNEY, I.T. *NETLAB: Algorithms for Pattern Recognition,* 2001, Springer.

https://doi.org/10.1017/S0001924000002219 Published online by Cambridge University Press