
Article:
Barthorpe, R.J. orcid.org/0000-0002-6645-8482 and Worden, K. orcid.org/0000-0002-1035-238X (2017) On multi-site damage identification using single-site training data. Journal of Sound and Vibration, 409. pp. 43-64. ISSN 0022-460X

https://doi.org/10.1016/j.jsv.2017.07.038

Journal of Sound and Vibration 409 (2017) 43–64

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

On multi-site damage identification using single-site training data

R.J. Barthorpe *, G. Manson, K. Worden

Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, United Kingdom

ARTICLE INFO

Article history:
Received 7 December 2016
Received in revised form
30 June 2017
Accepted 21 July 2017
Available online 29 July 2017

Keywords:
Structural health monitoring
Support vector classification
Multi-site damage identification
Statistical pattern recognition

ABSTRACT

This paper proposes a methodology for developing multi-site damage location systems for engineering structures that can be trained using single-site damaged state data only. The methodology involves training a sequence of binary classifiers based upon single-site damage data and combining the developed classifiers into a robust multi-class damage locator. In this way, the multi-site damage identification problem may be decomposed into a sequence of binary decisions. In this paper Support Vector Classifiers are adopted as the means of making these binary decisions. The proposed methodology represents an advancement on the state of the art in the field of multi-site damage identification which require either: (1) full damaged state data from single- and multi-site damage cases or (2) the development of a physics-based model to make multi-site model predictions. The potential benefit of the proposed methodology is that a significantly reduced number of recorded damage states may be required in order to train a multi-site damage locator without recourse to physics-based model predictions. In this paper it is first demonstrated that Support Vector Classification represents an appropriate approach to the multi-site damage location problem, with methods for combining binary classifiers discussed. Next, the proposed methodology is demonstrated and evaluated through application to a real engineering structure – a Piper Tomahawk trainer aircraft wing – with its performance compared to classifiers trained using the full damaged-state dataset.

1. Introduction

Structural Health Monitoring (SHM) refers to the process of measuring and interpreting *in situ* data acquired from a structural system in order to objectively quantify the condition of the structure. SHM has the potential to offer substantial economic and life-safety benefits for aerospace and civil structures in applications as diverse as civil and military aircraft, civil infrastructure, energy generation and offshore structures. The three motivating aspects for SHM that are recurrently mentioned in the literature are: (1) the life-safety benefits achievable through being able to continuously monitor safety-critical components. (2) the economic benefits achievable through avoiding unplanned down-time and increasing the efficiency of inspection and maintenance, and (3) the ability to optimise newly designed structures for which the current health condition is known through monitoring [1]. The damage identification problem can usefully be considered as a hierarchical process, developed from that discussed by Rytter [2], that progresses from the detection of the occurrence of

* Corresponding author.
E-mail address: r.j.barthorpe@sheffield.ac.uk (R.J. Barthorpe).

http://dx.doi.org/10.1016/j.jsv.2017.07.038

damage; to localisation of where damage has occurred; to classification of the type of damage that is present; assessment of the damage extent; and finally prediction of the residual life of the system in light of the damage.

The focus of this study is on the second level of this hierarchy - localisation - and in particular the problem of identifying changes in the structure that have occurred at multiple locations. Multi-site damage identification represents an important and challenging problem in SHM but one that has received comparatively little dedicated attention in the literature, with the majority of approaches presented in the literature (see, for example, the extensive reviews in [3,4]) focusing on the identification of single-site damage. The impact of this restriction is clear given that an *in situ* structure would be expected, over time, to exhibit degradation from its baseline state concurrently at multiple locations. Any practical application of SHM should be able to account for this fact, with systems that do not do so potentially unable to accurately describe the true health state of the structure.

Approaches to Structural Health Monitoring may be broadly separated into two classes: *data-based* approaches that follow a statistical pattern recognition paradigm and rely solely on experimental data; and *model-based* approaches that make use of the predictions of a representative physics-based model of the structure, typically by applying a Finite Element (FE) model updating procedure. While the problem of multiple damage location from a purely data-based perspective has received very little dedicated attention in the literature, FE model updating is, at least in principle, appropriate to the multiple damage location task. The solution procedure typically involves minimisation of residuals between damage-sensitive *features* drawn from the experimental structure and the predictions of the model. Ruotolo and Surace [5] presented an early study of this type focusing on multiple damage location and assessment in a cantilever beams, with a weighted sum of modal features (natural frequencies, modal curvatures and mass-normalised modeshapes) adopted as the objective function to be minimised. A similar approach, the Multiple Damage Location Criterion (MDLAC), is presented by Contursi et al [6], with the cross-correlation between measured and theoretical natural frequencies adopted as the objective function and applied to multi-site damage. An incremental development of the MDLAC approach is presented in [7] with damage location pursued using Dempster-Shafer evidence theory, allowing fusion of the results of frequency MDLAC and modeshape MDLAC outcomes, and a micro-search genetic algorithm used to estimate damage extent.

Dealing with the uncertainty that arises from physical variability, experimental variability and model-form error presents a major issue for FE updating approaches. In recent years this has led to increased interest in methods that explicitly account for uncertainty, with both probabilistic methods (for example Bayesian updating [8,9]), and non-probabilistic methods (for example fuzzy updating [10]) methods demonstrating considerable success for single-site damage location. Nonetheless, challenges remain. The principal drawback of the approach is that developing a finite element model of a complex engineering structure sufficient to reliably predict response changes due to damage is a difficult and potentially expensive task, as evidenced by growth of interest in the fields of model validation and uncertainty quantification within structural dynamics. Secondly, a potentially onerous number of model executions may be required to fully explore the parameter space during the updating step, although approaches such as parameter subset selection may go some way towards addressing this [11]. It is noted that in cases where experimental examples have been used to illustrate approaches they have typically exhibited a low level of complexity, for example cantilever beams and truss-type structures. It is also noted that appropriately thorough consideration of the impact of numerical and experimental uncertainty remains comparatively rare.

Data-based methods – as typically applied – initially appear less well suited to the multi-class damage location task than model-based methods. Damage location via a statistical pattern recognition (SPR) paradigm requires the adoption of a supervised learning approach, thus requiring the gathering of data from the structure in both its undamaged and damaged state [12]. Localisation is pursued by evaluating the difference between the current state of the structure and an initial baseline state that is representative of the undamaged structure. A feature vector is thus formed that may be compared to previously acquired, labelled observations in order to diagnose the state of the structure by applying a classification algorithm. The major drawback of the data-based approach to SHM is the requirement to have access to structural data for all damage states of interest, which is rarely available even for single-site damage. However, a major advantage of this approach is in the handling of uncertainty. Provided the training set contains data representative of the variability that will be seen in the monitoring phase, this variability will be accounted for be any appropriately trained classifier. Manson et al [13] present one such successful example of single-site damage location based upon features drawn from transmissibility spectra, with classification performed using artificial neural networks (ANNs). It is noted that there is a caveat to the requirement for damage state data in cases where the adopted feature exhibits sensitivity to local damage, and some success has been reported using, for example, modeshape curvatures [14]. Nonetheless, such features exhibit disadvantages alongside advantages (susceptibly to experimental noise being an example in the case of modeshape curvatures) and there is clear motivation for developing methods for multi-site damage location that maintain generality of feature choice.

The requirement to have access to structural data for all damage states of interest presents a critical constraint when considering classification of multi-site damage. In the case that damage is believed to occur at only one of a finite discrete set of n locations, the requirement is that data should be acquired for each of the corresponding n damage states. In the more general case of damage occurring concurrently at more than one location, the naïve approach might be to gather damage data for all combinations of damage location. However, the number of states for which data would be required in order to cover all combinations would be 2^n (including the undamaged state), and so grows exponentially with the number of locations n. Even in a more restrictive case where classification of up to k damage locations from the full set of n locations is sought, the number of damaged states for which data would be required is $\sum_{i=1}^{k} n!/(i!(n-1)!)$. For illustration, consider as an example a structure with $n = 10$ damage states of interest, perhaps representing particular locations on the surface of a

continuous structure or discrete components of an assembled structure. In order to locate single site damage, a supervised learning algorithm would require data from all ten damage states, plus the normal condition. If, instead, data were required to train a classifier capable of locating damage at up to three sites, test data from 175 combinations of damage states would be required. A classifier capable of handling all 2^{10} combinations would require data from all 1024 damage states. As damaged state data for even single site damage is very rarely available, it would appear unwise to rely on the availability of damage data for all possible combinations of damage location in the general case. It is apparent that the lack-of-data problem, a major obstacle in the diagnosis of single-site damage, is a potentially critical issue in multiple-site damage location.

The contribution of the current study is to present a purely data-based approach to multi-site damage location that dramatically reduces the number of states for which training data is required. This is pursued by constructing a data-driven classifier that is capable of distinguishing occurrences of single-site damage, and evaluating the success of this classifier when presented with multiple-site damage data. The study is experimental in nature with the structure used for illustration being the wing of a Piper Tomahawk trainer aircraft. The features used are discordancy measures associated with transmissibility measurements, and Support Vector Machines (SVMs) are employed for classification. The layout of the paper is as follows: the experimental structure and data acquisition process are described in Section 2; an approach for reducing the gathered high-dimensional response data into a low-dimensional feature set is developed in Section 3; the performance of the selected feature set when presented with observations of multi-site damage is evaluated in Section 4; and the use of these features as the basis of a multi-site damage classifier is explained in Section 5. Results and conclusions are presented in Sections 6 and 7.

2. Experimental data acquisition

This section describes the experimental structure used in the study and the test sequence adopted. The raw data used in this study consisted of accelerance frequency response functions (FRFs) gathered from 15 locations on the structure in response to single point excitation. The accelerance FRFs are first converted into acceleration transmissibility spectra, described further in Section 3.2. A key motivation for the use of transmissibility spectra is that they can, in principle, be estimated without measuring the excitation applied to the structure. This makes them an attractive proposition for eventual practical application where load records are unlikely to be available. Finally, multivariate discordancy measures derived from sections of the transmissibility spectra are adopted as candidate damage sensitive features, described further in Section 3.3.

Throughout the study a great deal of attention was paid to test sequencing, as it is known that decisions made here can have a substantial bearing on the quality of the training dataset, and thus the performance of any classifier trained using this set. Care was taken throughout in order to produce a training dataset that is representative of the conditions under which classification will be sought. For the structure under study, it was felt that the boundary condition variability associated with the removal and replacement of inspection panels was likely to represent the most significant source of random variation in the response spectra. Accordingly, care was taken both to reduce this variability to as great an extent as possible by using a torque-controlled screwdriver and taking care over screw tightening, and to account for the outstanding variability in the recorded data through randomisation.

2.1. Experimental structure and data acquisition

The experimental structure considered in this study is an aluminium aircraft wing, tested in laboratory conditions. The main spar of the wing was bolted in cantilevered fashion to a substantial, sand-filled steel frame that was purpose built for the study as shown in Fig. 1. Achieving a truly rigid mounting is non-trivial. However, given the focus on monitoring *changes* in dynamic response as a result of the introduction of damage, it was felt that any small degree of flexibility in the mounting should have minimal effect on the outcome of the study. The wing is mounted upside-down in order to allow access to the inspection panels mounted on the underside of the wing. Mounting the wing in this fashion also ensures that it is preloaded in the correct direction.

Fifteen PCB 353B16 piezoelectric accelerometers were mounted on the upper (as mounted) surface of the wing using ceramic cement. The sensors are denoted $S_1 - S_{15}$, with their location plus those of the inspection panels and sub-surface stiffening elements (dotted lines) shown schematically in Figs. 2 and 3. The sensors were placed so as to form transmissibility paths between sensor pairs, with sensor S_{14} acting as a reference output for sensors S_1-S_5 and S_{15} performing this role for sensors S_6-S_{13}. The thirteen resulting paths are indicated in Fig. 3 and specified in full in Section 3. Placement directly above stiffening elements was avoided as it was believed that such locations would offer poor observation of localised changes in structural flexibility.

Experimental data acquisition was performed using a DIFA SCADAS III unit controlled by LMS software running on a desktop PC. All measurements were recorded within a frequency range of 0–2048 Hz with a resolution of 0.5 Hz. The structure was vertically excited with a band-limited Gaussian signal using a Gearing and Watson amplifier and shaker mounted on the underside of the wing. The shaker was connected to the wing via a stinger attached approximately 1 metre from the root end of the wing and positioned on the main spar. The excitation point was selected heuristically with the aim

Fig. 1. Piper Tomahawk aircraft wing.

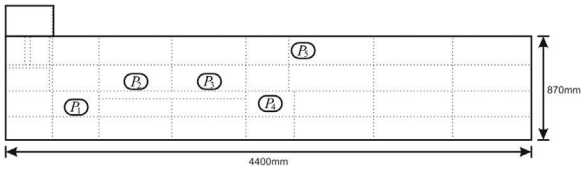

Fig. 2. Schematic of wing inspection panels and stiffening elements.

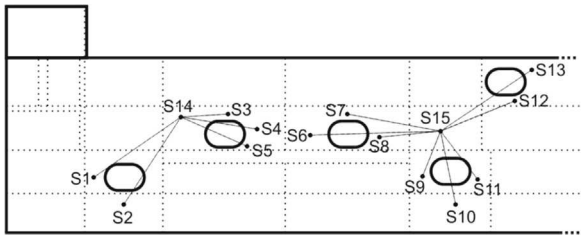

Fig. 3. Schematic of inspection panels, sensor locations and stiffening elements.

of efficiently delivering low frequency excitation to the structure under test. Five-samples averages were recorded in all cases as this was found to offer a good compromise between noise reduction and acquisition time.

2.2. Damage introduction

In order to introduce damage in a repeatable, and in some sense realistic, way advantage was taken of the presence of inspection panels on the underside of the wing. Five such panels, denoted P_1 to P_5, were employed in the present study, all of the same dimension and orientation. Through removal of these panels a gross level of damage could be introduced at each location. These damage states are denoted by $D_{\{.\}}$ where $\{.\}$ indicate the set of panels removed i.e. $D_{\{1,5\}}$ refers to the removal of panels P_1 and P_5. The undamaged case, with all intact panels in place, is referred to as N (normal condition). Superscripts are used to distinguish repeat realisations of the same damage state e.g. N^2 refers to the second realisation if the normal state. Examples of the normal state N and the damage state $D_{\{2\}}$ for panel P_2 are shown in Figs. 4 and 5 respectively.

The advantages of using the removal of a panel as a proxy for damage are that it is non-destructive, repeatable, and the primary effect of panel removal (a localised reduction in structural stiffness) is expected to be similar to the effect of introducing gross damage at the same location. The disadvantage is that the repeatability is not perfect. During preliminary

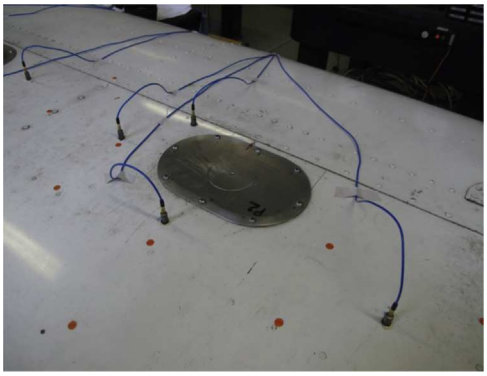

Fig. 4. Wing with Panel P_2 in place.

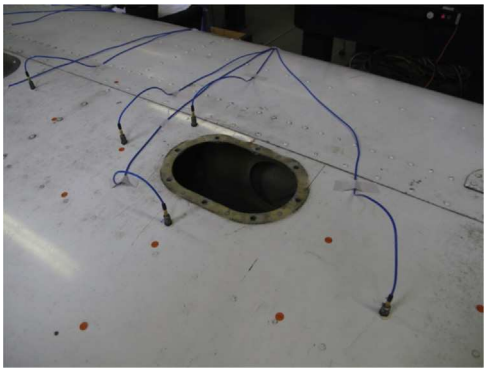

Fig. 5. Wing with Panel P_2 removed.

studies it was found that the removal and reattachment of the panel led to substantial variability in the FRF observations. This and other sources of variability were considered in greater detail prior to the specification of the test sequence in Section 2.4.

2.3. Sources of experimental variability

Consistent with the aims of the statistical pattern recognition (SPR) paradigm [12], consideration is given to potential sources of variability at the test planning stage. The ultimate aim is to develop a classifier that will generalise well to data unseen during training. By identifying and allowing for sources of variability at an early stage, the prospects of achieving good generalisation at a later stage may be increased. With regard to variability, the objectives of the current study are to reduce variability to whatever extent is possible, and to ensure the remaining variability is appropriately represented in the dataset. Preliminary testing indicated that the panel boundary conditions represented the dominant source of systematic experimental variability, with minimal effects from operational and environmental variation. This is in addition to random variation between runs (i.e. noise).

The inspection panels are attached using eight screws. It has been observed in previous studies that the measured FRFs may display marked sensitivity to the boundary condition variability that arises from the removal and reattachment of the panels [13]. The degree of variability observed can be problematic when attempting to construct a sensitive, robust classifier to identify damage states. In the experimental portion of the study, emphasis was placed upon gathering an improved representation of states that would be used for training and validating the developed classifier. To this end, systematic randomisation is applied in the development of the training set test sequence in order to incorporate the effect of panel removal and replacement in the data used for classifier training. In order to reduce the random variability associated with the panel boundary conditions, a torque-controlled electric screwdriver was used for removing and replacing the panels. The order in which the bolts were replaced was also taken into account, with a consistent sequence of attachment maintained throughout the test. These efforts to reduce random variability were consistently applied for both the training and testing sets. In practical applications, this level of consistency would not be expected. The training dataset should be structured so as to capture the true level of variability as fully as possible, in order that a classifier that generalises well to new data may be built.

Table 1
Datatset A test sequence.

Block	Run	Damage	Observations	Block	Run	Damage	Observations
1	1	N^1	5×20	6	11	N^6	5×20
	2	$D^1_{\{1\}}$	100		12	$D^2_{\{1\}}$	100
2	3	N^2	5×20	7	13	N^7	5×20
	4	$D^1_{\{2\}}$	100		14	$D^2_{\{2\}}$	100
3	5	N^3	5×20	8	15	N^8	5×20
	6	$D^1_{\{3\}}$	100		16	$D^2_{\{3\}}$	100
4	7	N^4	5×20	9	17	N^9	5×20
	8	$D^1_{\{4\}}$	100		18	$D^2_{\{4\}}$	100
5	9	N^5	5×20	10	19	N^{10}	5×20
	10	$D^1_{\{5\}}$	100		20	$D^2_{\{5\}}$	100

2.4. Test sequencing

Two tests were conducted, resulting in two data sets being available with which to develop and test classifiers. A test sequence was developed for each test to reflect the differing testing objectives. Randomisation and blocking were applied in the specification of the test sequences in order to account for the effects of measurement noise and panel boundary condition variability.

Dataset A comprises 1000 normal state observations and 1000 single-site damage state observations. The primary purpose of Dataset A was to provide training and validation data with which to develop a classifier, with the intention that this classifier then be capable of generalising to identifying multi-site damage for an independent test set. This is a demanding objective and particular attention was paid to full randomisation of the panel boundary condition in order to fully represent the variability that arises from this effect. *Randomisation*, as used here, refers to the removal and replacement of panels to ensure that the latent boundary condition variation (that which is present despite the use of a torque controlled screwdriver and care over the order of screw tightening) be represented in the dataset.

The test sequence used for gathering Dataset A is summarised in Table 1. The test consists of 10 *blocks*, each comprising a normal condition run and a single-site damage condition run. For each of the single-site damage states, each block contained 20 normal condition runs alternated with 20 damaged condition runs. Each *run* contains 100 observations of five-sample averaged response data. The use of blocks of testing allowed two levels of boundary condition randomisation to be introduced. In order to evaluate the variability arising from single-panel removal within each block, the panel of interest was removed and replaced after every 20 observations for the normal condition runs. Within each block only the panel of interest is removed - the remaining four panels remain in place. For the panel-off case, where there are no single-plate boundary condition effects to consider, a straight run of 100 observations was recorded. Between blocks, all five panels were removed and replaced. This introduced full randomisation of the boundary conditions. In order to add further information, blocks were executed for each of the five panels, and then the sequence repeated.

The test sequence used for gathering Dataset B is summarised in Table 2. The primary purpose of Dataset B was to provide a training set to allow the evaluation of the performance of the developed classifiers when presented with new data comprising normal, single-site and multi-site data. It contains all possible damage states (1 normal condition and 5 single-site damage states plus 26 multi-site damage states) plus a number of repetitions for each state with various degrees of randomisation, particularly of the normal condition.

The test was conducted in blocks comprising a number of normal condition runs and runs from one of the damaged state conditions. For each of the single-site damage states, each block contained 20 normal condition runs alternated with 20 damaged condition runs. For the multiple-site damage states, each run contained five normal and five damaged condition runs. Each run contains 10 observations of five-averaged data. Within each block only the panel (or panels) of interest are removed and replaced. For example in the first block, Run 1 compromised 10 observations of the normal condition N^1. Panel P_1 was then removed and Run 2, comprising 10 observations of the damaged condition D_1, was recorded. The panel was then replaced and Run 3, a further 10 observations of N^2, was performed.

Systematic randomisation of boundary conditions through removal and replacement of all panels between blocks was not pursued, although gradual randomisation was introduced as panels were removed and replaced during the test sequence. This is of some importance to any conclusions drawn with regard to the approach. The primary objective of Dataset B is to provide a training set to allow the evaluation of the performance of the developed classifiers when presented with new data comprising normal, single-site and multi-site data. It is not intended to be the basis of an exhaustive examination of how the developed classifiers perform across the full domain of possible observations, in the presence of variability. Any conclusions drawn must thus include the caveat that they are valid only for the data with which the classifiers were tested; and caution should be employed in drawing general conclusions on the basis of a limited testing dataset.

Total testing time was approximately 15 hours distributed over a five day period for Dataset A and 34 hours distributed

Table 2
Datatset B test sequence.

Block	Run	Damage	Observations	Block	Run	Damage	Observations
1	1–40	N^1	20 × 10	17	311-320	N^{17}	5 × 10
		$D_{\{1\}}$	20 × 10			$D_{\{1,2,4\}}$	5 × 10
2	41–80	N^2	20 × 10	18	321-330	N^{18}	5 × 10
		$D_{\{2\}}$	20 × 10			$D_{\{1,2,5\}}$	5 × 10
3	81–120	N^3	20 × 10	19	331-340	N^{19}	5 × 10
		$D_{\{3\}}$	20 × 10			$D_{\{1,3,4\}}$	5 × 10
4	121–160	N^4	20 × 10	20	341-350	N^{20}	5 × 10
		$D_{\{4\}}$	20 × 10			$D_{\{1,3,5\}}$	5 × 10
5	161–200	N^5	20 × 10	21	351-360	N^{21}	5 × 10
		$D_{\{5\}}$	20 × 10			$D_{\{1,4,5\}}$	5 × 10
6	201–210	N^6	5 × 10	22	361-370	N^{22}	5 × 10
		$D_{\{1,2\}}$	5 × 10			$D_{\{2,3,4\}}$	5 × 10
7	211–220	N^7	5 × 10	23	371-380	N^{23}	5 × 10
		$D_{\{1,3\}}$	5 × 10			$D_{\{2,3,5\}}$	5 × 10
8	221-230	N^8	5 × 10	24	381-390	N^{24}	5 × 10
		$D_{\{1,4\}}$	5 × 10			$D_{\{2,4,5\}}$	5 × 10
9	231-240	N^9	5 × 10	25	391-400	N^{25}	5 × 10
		$D_{\{1,5\}}$	5 × 10			$D_{\{3,4,5\}}$	5 × 10
10	241-250	N^{10}	5 × 10	26	401-410	N^{26}	5 × 10
		$D_{\{2,3\}}$	5 × 10			$D_{\{1,2,3,4\}}$	5 × 10
11	251-260	N^{11}	5 × 10	27	411-420	N^{27}	5 × 10
		$D_{\{2,4\}}$	5 × 10			$D_{\{1,2,3,5\}}$	5 × 10
12	261-270	N^{12}	5 × 10	28	421-430	N^{28}	5 × 10
		$D_{\{2,5\}}$	5 × 10			$D_{\{1,2,4,5\}}$	5 × 10
13	271-280	N^{13}	5 × 10	29	431-440	N^{29}	5 × 10
		$D_{\{3,4\}}$	5 × 10			$D_{\{1,3,4,5\}}$	5 × 10
14	281-290	N^{14}	5 × 10	30	441-450	N^{30}	5 × 10
		$D_{\{3,5\}}$	5 × 10			$D_{\{2,3,4,5\}}$	5 × 10
15	291-300	N^{15}	5 × 10	31	451-460	N^{31}	5 × 10
		$D_{\{4,5\}}$	5 × 10			$D_{\{1,2,3,4,5\}}$	5 × 10
16	301-310	N^{16}	5 × 10				
		$D_{\{1,2,3\}}$	5 × 10				

over an eleven day testing period for Dataset B. The observed temperature range was reasonably consistent between the two tests, being 19.2-21.9°C for Dataset A and 19.1-24.8°C for Dataset B. It was not expected that this range of temperatures would have a significant effect on the measured data. In practice, it is to be expected that damage identification would be conducted in the presence of operational and environmental variation, with temperature being among the key drivers in aerospace applications. The robustness of the method in the face of temperature variation was not studied in the current work. In general, three approaches are available for handing confounding influences in data-based SHM: seeking features that are sensitive to damage but insensitive to benign variation; pre-processing the data in such a way that the confounding influences will be filtered out; or including data representative of the confounding influences within the training set. Any practical application of the approach would involve one or more of these approaches to be adopted, and it is the opinion of the authors that every effort should be put into data normalisation where possible with any residual variation subsequently 'learnt' within the training process. This approach is equally valid whether training is conducted using single-site or multi-site data.

3. Feature extraction and selection

The aim of this section is to describe the reduction of the high-dimensional data recorded in Section 2 into a set of low-dimensional features suitable for the training of statistical classifiers in Section 5, and to evaluate the individual performance of the chosen feature set. The objectives of the feature selection and evaluation task may be summarised in several steps.

- Extraction of a candidate feature set from raw experimental data.
- Selection of a low-dimensional feature set containing features that are sensitive to single-site damage, using a training set drawn from Dataset A.
- Evaluation of these features for the identification of single-site damage and multi-site damage using a testing set drawn from Dataset B.

Each of these steps is covered in turn. The hypothesis of this section is that it may be possible to find features that offer a good level of discrimination between multiple-site damage states, despite only single-site damage states being available for feature selection.

3.1. Feature extraction

Damage leads to the dynamic response of the structure deviating from that observed when it is in its initial, undamaged condition. The structural response features considered in this study are transmissibility spectra. By comparing examples of undamaged and damaged spectra, regions of the spectra that are sensitive to particular damage states may be identified. These regions form the basis of features that may subsequently be used to train a statistical damage classifier. In the interests of developing a statistical classifier, it is desirable that the feature set used is of low dimension. Achieving a suitably concise feature set requires further condensation of the identified spectral region. In this study the additional data reduction is performed using a discordancy measure - the Mahalanobis squared distance (MSD) - between the newly presented and therefore potentially damaged state, and the previously recorded undamaged state. The resulting features are the discordancy values associated with damage sensitive regions of the transmissibility spectra. In Section 5, classifiers are trained using observations of the resulting concise set of damage sensitive features.

The raw data recorded in Section 2 comprised 15 accelerance FRFs, each containing 4097 spectral lines. The procedure undertaken for reducing this high-dimensional data to a low-dimensional set of features in this study is as follows. The FRF data are first converted into transmissibility spectra, which have been shown to perform well in similar damage identification scenarios [15,16,13,17]. In this study, log magnitude transmissibility spectra are employed, which exhibit further useful properties when employed for feature selection, as discussed below. Next, regions of the transmissibility spectra that are sensitive to damage at a particular location are selected. In this instance, feature selection is performed manually on the basis of a principled set of objectives and aided by visual tools. Finally, the identified low-dimensional spectral ranges are reduced to a scalar feature value through the application of the MSD measure.

3.2. Transmissibility data

The raw test data comprised real and imaginary components of the accelerance FRFs between the 15 response locations and the single force input from the shaker. Prior to further processing, this data was converted to transmissibility spectra using the relationship,

$$T_{ij}(\omega) = \frac{H_{ik}(\omega)}{H_{jk}(\omega)} \tag{1}$$

where $H_{ik}(\omega)$ is the FRF between input location k and response location i, $H_{jk}(\omega)$ is the FRF between input location k and response location j, and $T_{ij}(\omega)$ is the resulting transmissibility spectra between response location i and response location j. Log magnitude transmissibility spectra were generated for the 13 transmissibility paths $T_{1-5,14}$ and $T_{6-13,15}$.

The sensors were placed so as to form transmissibility 'paths' between sensor pairs. The sensors are arranged into two groups, the first covering panels P_1 and P_2 and the second covering panels P_3, P_4 and P_5. Each group contains a *reference* sensor which is common to all transmissibility paths within the group and a number of *response* sensors. For the first group the reference sensor is S_{14} and the response sensors are S_1-S_5; for the second group the reference sensor is S_{15} and the response sensors are S_6 to S_{13}, resulting in 13 transmissibility paths. Each panel has either two or three transmissibility paths associated with it. It is expected that the effects of removing a panel may observed in all spectra to some extent, but that the effect will be most apparent in the spectra associated with that panel.

3.3. Dimensionality reduction

The resulting transmissibility data is very high-dimensional. Each observation comprises 13 spectra, each of which contains 4097 spectral lines, giving a 53261-dimensional dataset. In order to alleviate the curse of dimensionality [18], some degree of dimensionality reduction must be pursued prior to applying pattern recognition methods. In this study, dimensionality reduction is realised as the selection of damage-sensitive regions of the transmissibility spectra, and the reduction of these multivariate regions to univariate features using a discordancy measure.

Selection of damage-sensitive windows of the spectra is the first stage in reducing the dimension of the recorded data. Further dimensionality reduction is achieved through outlier analysis using a discordancy measure. The discordancy measure employed is the Mahalanobis squared distance, which is a multivariate extension of the univariate discordancy measure. Discordancy measures allow deviations from normality to be quantitatively evaluated.

A brief summary of the technique is given here: the technique is described in [19] and validated for an experimental structure in [15,16,13]. For a multivariate data set consisting of n observations in p variables, the MSD may be used to give a measure of the discordancy of any given observation. The scalar discordancy value MSD_ξ is given by,

$$MSD_\xi = (\mathbf{x}_\xi - \bar{\mathbf{x}})^\top \Sigma^{-1} (\mathbf{x}_\xi - \bar{\mathbf{x}}) \tag{2}$$

where $\mathbf{x}_{\xi}$ is the potential outlier, $\bar{\mathbf{x}}$ is the mean of the sample observations and Σ is the sample covariance matrix. The MSD approach offers a measure of the extent by which a sample differs from the population. The mean and covariance matrix may be deemed either *inclusive* or *exclusive* measures dependent upon whether they are computed using data where outliers are already present. In this study an exclusive measure in adopted, with $\bar{\mathbf{x}}$ and Σ calculated using data from the structure in its undamaged condition only.

3.4. Feature selection from single-site training data

Feature selection from a spectrum can be done in several ways. It may be performed manually, on the basis of engineering judgement; algorithmically, primarily by applying some form of combinatorial optimisation; or a combination thereof. The three approaches are illustrated in [20], with outlier analysis performed on transmissibility spectra from an experimental structure subjected to damage. The aim was to select a set of nine features as inputs to a multi-layer perceptron (MLP) classifier. First a manual approach was taken to select 44 candidate features from the full spectral dataset, and from these a further reduction to the required nine features was made on the basis of engineering judgement. Secondly, a Genetic Algorithm (GA) was employed to select an 'optimal' subset of nine features from the manually selected set of 44: an example of a manual/algorithmic approach. Finally, a GA was run to select nine features from the full spectral dataset, with no prior selection of feature ranges. The broad conclusion was that the manual/algorithmic approach worked 'best' for the application presented, followed by the fully algorithmic approach and the fully manual approach. The great benefit of the fully algorithmic approach is in removing the requirement for an exhaustive manual search for candidate features.

In this section, the objective is to select features that are robustly indicative of the removal of panels using observations from Dataset A. Objectively optimal feature selection is not pursued in this instance, although this is discussed as an area of future work. The feature selection method is instead performed manually. This process is guided by a series of considerations, and aided by appropriate visualisations. In total, 10 features were selected for each of the five panels.

The aim of the feature selection step is to identify features that discriminate between the normal condition and the removal of a single panel, for each of the five panels in turn. For the present work, a predominantly manual feature selection approach was employed, guided by evaluation of the discordancy values of all possible 20-spectral-line windows of the spectra. As each spectrum contains 4097 lines, 4078 such windows existed for each of the 13 spectra, resulting in a total of 53014 candidate features. Discordancy values for each of these features were evaluated using the mean-averaged data recorded for each of the 20 runs in Dataset A. Plotting the discordancy values across the feature set for each of the structural states allows the sensitivity of the features to be visualised. An example for a limited range of the candidate feature set is given in Fig. 6(a). The two repetitions of removal of panel P_1 are shown as red dotted lines. The remaining 18 damage states (10 normal condition runs and eight other single panel removal runs) are shown as green dashed lines.

Four criteria were considered when assessing the strength of a candidate feature.

1. The feature should be sensitive to the removal of the specified panel.
2. In the interests of promoting robustness there should be consistency in the feature values between the two available repetitions.
3. Preferably, the feature should be insensitive to the removal of panels other than that specified. The success of the classifier is predicated upon the discriminative ability of the combined feature set; however, it is expected that if features that are capable of discriminating between damage classes individually can be identified, the discriminative ability of the combined set will also be increased.
4. Preferably, there should be a degree of consistency in the above considerations between the selected feature and those in the immediate vicinity. The feature selection process is guided by a desire to attribute some degree of confidence to the performance of the feature when presented with new data. It would appear, at the very least intuitively, that more confidence may be placed in a feature that is one of several contiguous features that perform well, rather than a feature that performs conspicuously better than its neighbours.

An extended list of candidate features was identified using the MSD plots, with between 15 and 20 features identified for each panel. Each of these features was manually assessed in order to reduce the feature set to a final size of 10 features per panel. Inspection was aided by employing visualisations of the form given in Fig. 6.

Fig. 6(a) gives an expanded view of MSD values for the identified feature and for those in its vicinity. The values plotted as red dotted lines are the discordancy values for the window when the panel of interest is removed (panel P_1 in the example plot in Fig. 6); the values plotted as green dashed lines are for the normal conditions and removals of the other four panels individually (i.e. panels P_2-P_5 in the example). The dashed vertical line denotes the selected feature, with the window index number corresponding to the index of the first spectral line in the 20-line window used to form the feature. Fig. 6 (b) displays the portion of the transmissibility spectra corresponding to the selected window for Dataset A. The window itself is denoted as lying between the dashed vertical lines. For clarity of presentation, the spectra are presented as shaded intervals bounded by the minimum and maximum measured spectral values for each class of data. In common with the upper plot, the spectra in red bounded by dotted lines correspond to single-site removal of the panel of interest; the spectra in green bounded by dashed lines denote the single-site removal of the other panels. The normal condition spectra are

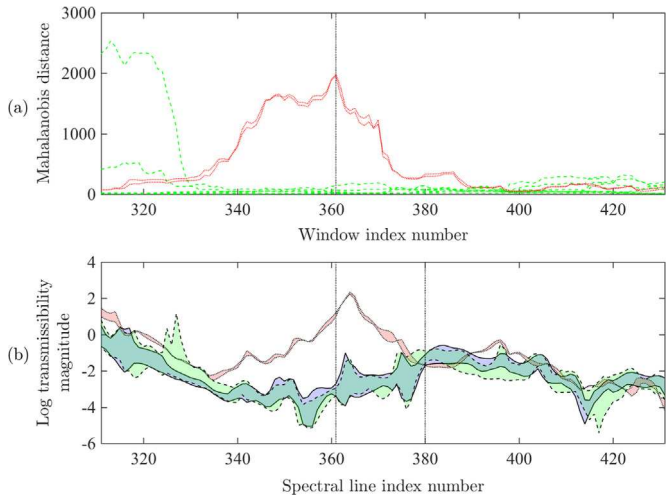

Fig. 6. Feature selection plots for Window Index Number 361, selected to identify removal of panel P_1. (a) The Mahalanobis distance values were used to discriminate between panel off (red, dotted) and panel on (green, dashed) conditions. (b) The transmissibility spectra illustrate the performance of the feature in discriminating between panel off (red, dotted), panel off (green, dashed), and normal (blue) conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

shown in blue and bounded by solid lines. The suitability of candidate features was manually evaluated on the basis of the selection criteria specified above. Through visual inspection of the full range of 4078 candidate windows for each of the 13 spectra, an initial feature set comprising 50 damage sensitive windows was selected.

Returning to the transmissibility plots after initially assessing the features on the basis solely of discordancy values allowed the information set available for making a final decision to be expanded. In several cases, considering the spectra led to some adjustment of the originally identified feature. One example of this is where a greater level of apparent robustness may be achieved through a small adjustment of the window. A second example is where the window returning the highest MSD values is found to contain multiple characteristic structures, such as peaks or troughs in the spectra. In such cases, it may be desirable for the window to span a single peak well, rather than partially span two or more such peaks.

As an example, consider the feature identified in Fig. 7. The original feature window, denoted by dotted vertical bars, contained two peaks. In the interests of promoting robustness in the feature set the window illustrated in Fig. 8 was preferred to that shown in Fig. 7. Despite a lower discordancy value than that for the initially identified feature, plus a lower level of consistency between the two observed repetitions, the feature shown in Fig. 8 was adopted due to the presence of only one characteristic in the defined window.

Through the process of assisted manual selection detailed above, the feature set was reduced from a candidate set of 53014 features to a final set of 50 features. The results of the feature selection exercise using single-site training data are given in Table 3. For conciseness, only the average discordancy value across the two repetitions is included. Discordancy values were a primary consideration for feature selection, and the values returned in these columns allow the reader some insight into the relative sensitivity of the features to damage on the allotted panel, and to the consistency in values between repetitions.

4. Feature evaluation for single-site and multi-site damage

In this section, the discriminatory performance of features selected using Dataset A are evaluated when these are applied to Dataset B. The objectives are two-fold. The first objective is to verify that the selected features are, individually, capable of discriminating between normal and single-site panel removals for a previously unobserved dataset. It is expected that the features should be capable of making this distinction with a level of accuracy approaching that achieved for the training set. The second objective is to evaluate the performance of the features in discriminating between the previously unobserved normal and multi-site panel removal data in Dataset B. The hypothesis here is that a good degree of discrimination might be achieved in the multi-site case, despite no multi-site data being used for feature selection or eventual training. The performance of the individual features is evaluated with the hope that this performance may approach that achieved for single-site cases. In Section 5 the requirement will be that the degree of discrimination of the combined feature set should be sufficient to permit the training of a family of classifiers that can robustly separate normal and damage states. In this study, evaluation is performed in incremental steps. Initially, a visual assessment of the transmissibility spectra in the vicinity of the selected features is made. This allows comment to be made on the general behaviour of the spectra when single- and

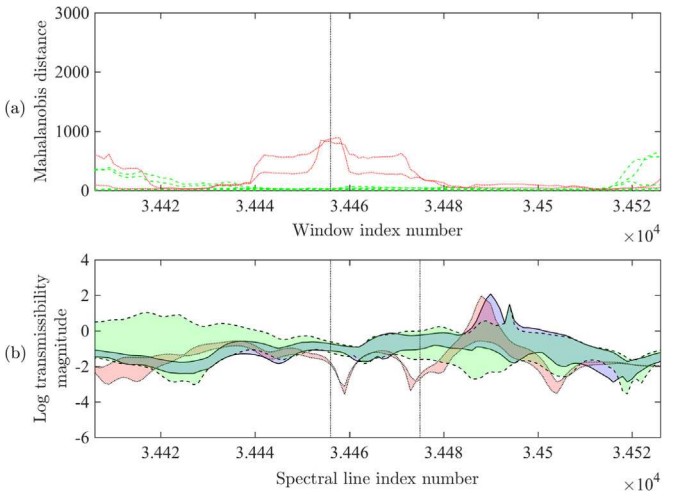

Fig. 7. Feature selection plots for Window Index Number 34456, proposed for identifying removal of panel P_4.

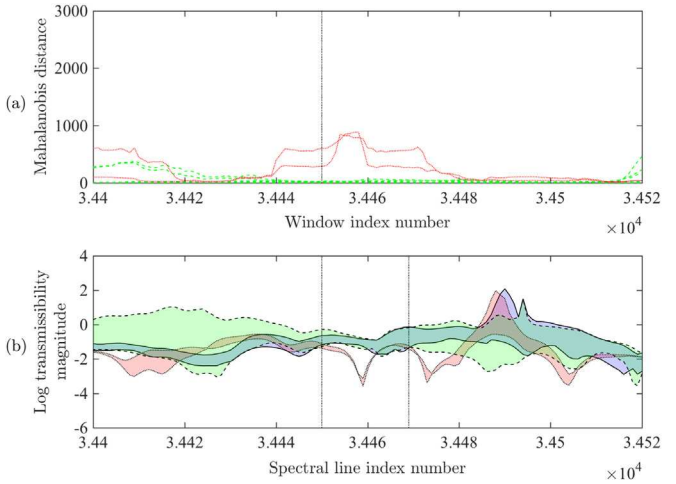

Fig. 8. Feature selection plots for Window Index Number 34450, selected to identify removal of panel P_4.

multi-site damage is introduced into the structure. Secondly, the discordancy values associated with each feature are evaluated for all 4600 observations contained in Dataset B. Finally, a summary statistic is applied to allow the discriminative capability of each feature to be concisely stated both for single-site and multi-site panel removal observations.

4.1. Qualitative feature evaluation

An initial assessment of the performance of the features selected in Table 3, when applied to the test set, was made through a qualitative examination of the spectra and through the calculation of discordancy values for every observation in the test set. First, spectral plots were produced to allow the behaviour of the transmissibility spectra in the region of the selected features to be assessed. In order to reduce the number of spectra to be visualised, averaging was applied in a manner consistent with that previously applied for feature selection. For each of the 62 runs (31 normal condition, five single-panel removals and 26 multiple-panel removals) named in the test sequence in Table 2, one mean-averaged spectra was calculated. To aid the clarity of presentation, the interval bounded by the minima and maxima of these averaged runs will be used for illustration in Figs. 9 and 10 in place of the full set of averaged spectra.

Secondly, the feature values returned for each of the 4600 observations in the testing set were quantitatively evaluated. The discordancy value MSD_{ξ} for each feature and for each observation is evaluated using Eq. (2), with the mean vector $\bar{\mathbf{x}}$ and covariance matrix $\boldsymbol{\Sigma}$ representing the normal condition data from Dataset A. Note that while this requires a large number of discordancy evaluations to be made (4600 observations for each of the 50 features in this case, requiring a total of 230000

Table 3
Selected feature set.

Panel	Feature	Path	Frequency [Hz]	Average Discordancy	Feature	Path	Frequency [Hz]	Average Discordancy
P_1	F_1	$T_{1,14}$	180.0–189.5	1968.1	F_6	$T_{1,14}$	1099.0–1108.5	591.0
	F_2	$T_{1,14}$	283.0–292.5	1081.9	F_7	$T_{2,14}$	133.5–143.0	1196.6
	F_3	$T_{1,14}$	336.0–345.5	1789.6	F_8	$T_{2,14}$	225.0–234.5	698.5
	F_4	$T_{1,14}$	386.5–396.0	2908.3	F_9	$T_{2,14}$	517.5–527.0	2246.6
	F_5	$T_{1,14}$	508.5–518.0	2063.5	F_{10}	$T_{2,14}$	643.0–652.5	555.5
P_2	F_{11}	$T_{3,14}$	102.0–111.5	41,717.9	F_{16}	$T_{4,14}$	1126.5–1136.0	11,683.2
	F_{12}	$T_{3,14}$	159.5–169.0	11,432.9	F_{17}	$T_{4,14}$	1386.5–1396.0	5523.6
	F_{13}	$T_{3,14}$	662.0–671.5	4287.4	F_{18}	$T_{5,14}$	269.5–279.0	16,984.1
	F_{14}	$T_{4,14}$	495.0–504.5	4986.8	F_{19}	$T_{5,14}$	494.5–504.0	2305.4
	F_{15}	$T_{4,14}$	913.0–922.5	9441.5	F_{20}	$T_{5,14}$	1387.0–1396.5	16,837.6
P_3	F_{21}	$T_{6,15}$	88.0–97.5	1914.2	F_{26}	$T_{7,15}$	940.5–950.0	1707.8
	F_{22}	$T_{6,15}$	208.5–218.0	749.1	F_{27}	$T_{8,15}$	352.5–362.0	2420.4
	F_{23}	$T_{6,15}$	945.5–955.0	2020.3	F_{28}	$T_{8,15}$	665.5–675.0	664.9
	F_{24}	$T_{6,15}$	1000.5–1010.0	2099.5	F_{29}	$T_{8,15}$	1140.0–1149.5	1912.9
	F_{25}	$T_{7,15}$	217.5–227.0	2533.2	F_{30}	$T_{8,15}$	1287.5–1297.0	2099.2
P_4	F_{31}	$T_{9,15}$	225.5–235.0	948.9	F_{36}	$T_{10,15}$	1149.5–1159.0	682.2
	F_{32}	$T_{9,15}$	310.0–319.5	2691.5	F_{37}	$T_{11,15}$	189.5–199.0	1218.8
	F_{33}	$T_{9,15}$	836.5–846.0	447.8	F_{38}	$T_{11,15}$	431.0–440.5	1473.9
	F_{34}	$T_{10,15}$	758.0–767.5	606.1	F_{39}	$T_{11,15}$	1090.5–1100.0	2090.8
	F_{35}	$T_{10,15}$	1003.5–1013.0	603.0	F_{40}	$T_{11,15}$	1250.5–1260.0	1325.7
P_5	F_{41}	$T_{12,15}$	359.5–369.0	1096.5	F_{46}	$T_{13,15}$	173.5–183.0	471.6
	F_{42}	$T_{12,15}$	513.0–522.5	201.6	F_{47}	$T_{13,15}$	308.0–317.5	386.3
	F_{43}	$T_{12,15}$	1110.0–1119.5	643.1	F_{48}	$T_{13,15}$	329.5–339.0	367.4
	F_{44}	$T_{12,15}$	1130.0–1139.5	462.0	F_{49}	$T_{13,15}$	369.0–378.5	1051.0
	F_{45}	$T_{13,15}$	1701.5–1711.0	203.8	F_{50}	$T_{13,15}$	1131.5–114.0	638.3

evaluations), comparative computational efficiency is maintained as the most computationally costly term - the inversion of the covariance matrix – need only be performed once per feature.

The transmissibility and discordancy outcomes for two of the 50 features are presented in Figs. 9 and 10. For each feature, four plots are given. The first plot presents the discordancy value data used for selecting features from Dataset A. Removal of the panel of interest is shown using red dotted lines and the removal of the other four panels are shown using green dashed lines. The second plot shows the portion of the transmissibility spectra corresponding to the chosen feature for each of the runs in Dataset A. The intervals spanning single-site removal of the panel of interest is denoted in red; single-site removal of the other panels is shown in green; and normal condition data is given in blue. The third plot shows transmissibility spectra from Dataset B for the chosen feature window. The colour scheme used is as for the second plot, but the intervals now span both single-site and multi-site panel removal. The final plot gives discordancy values for the feature, evaluated for all observations of Dataset B. The plot is divided according to the state of the structure using dashed vertical lines. Observations 1-2300 are for the normal condition; 2301-3300 correspond to single-panel removal; and 3301-4600 correspond to multiple-panel removal. These groups are further delineated using solid vertical lines. Damage states that include removal of the specified panel are highlighted in green.

Fig. 9 illustrates feature F_{32}, which performed well in identifying both single-site and multiple-site removals involving panel P_4. Fig. 9(b) illustrates the clear distinction between the removal of panel P_4 (in red) and the other states included in Dataset A, and this distinction led to the selection of the feature. Fig. 9(c) illustrates the same feature window for the transmissibility spectra of Dataset B. It is observed that the spectra behave in a very similar fashion to that found for Dataset A, both for the removal of panel P_4 alone and where panel P_4 was one of several removed. The clear distinction between the removal of panel P_4 and other states is maintained. This is reflected in the discordancy values illustrated in Fig. 9(d). The feature fires strongly for the observations of single panel removal, and similarly strongly for each of the states in which it was one of several panels removed (highlighted with a green background). This very encouraging level of performance was observed for the majority of the 50 identified features, and supports the hypothesis that a good degree of multi-site discrimination may be achieved.

Fig. 10 illustrates feature F_{49}, which performed less well in identifying removal of panel P_5. While the feature fired to a reasonable degree for the single-site removals, it was unable to distinguish all of the multiple site damage removals. It was in general found to be more challenging to identify features that performed well for panel P_5. However, the reason for feature F_{49} performing less well is likely to be attributable to the nature of the spectra captured within the identified feature

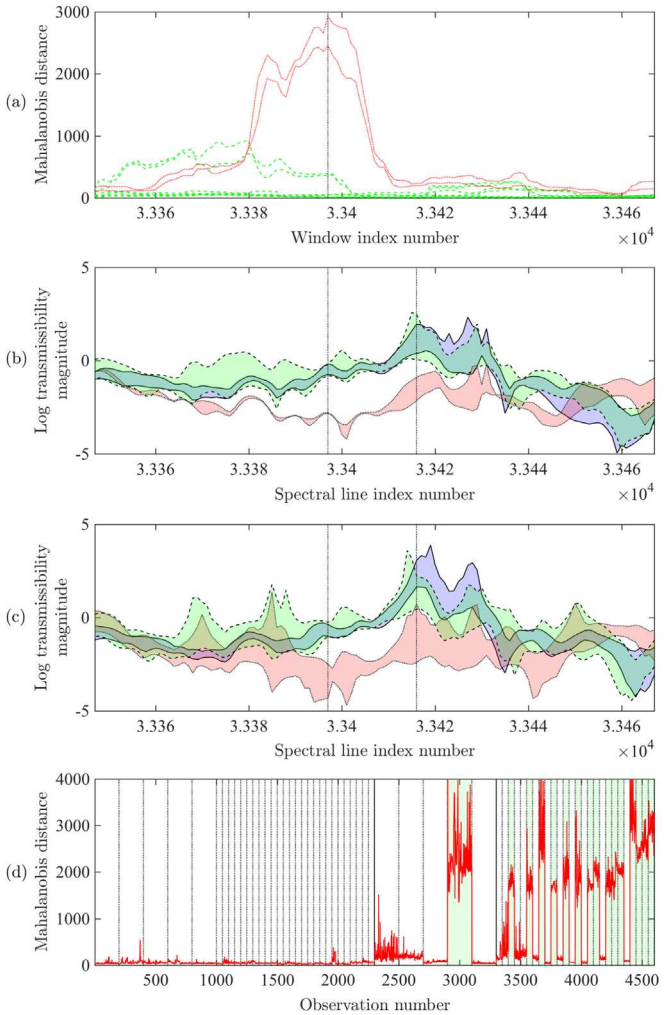

Fig. 9. Feature selection and evaluation plots for feature F_{32}, selected to identify removal of panel P_4. (a) Mahalanobis distance values between panel off (red, dotted) and panel on (green, dashed) conditions; (b) The transmissibility spectra for Dataset A for panel off (red, dotted), panel off (green, dashed), and normal (blue) conditions; (c) The corresponding transmissibility spectra for Dataset B; (d) The Mahalanobis distance values for the feature when applied to Dataset B, with conditions featuring removal of panel P_4 highlighted in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

window. The spectra comprise two 'peaks', only one of which demonstrates a marked shift as a result of damage. This was apparent to some degree from the observations of Dataset A in Fig. 10(b), but is more strongly evident in Fig. 10(c) for the observations of Dataset B.

The visual assessment of the selected feature windows suggests that there is similarity in the behaviour of spectra due to single-site and multi-site panel removal over some portions of the frequency ranges, and that this behaviour is reflected in the returned feature values. However, the visual assessment of a large number of features is a somewhat onerous task. A concise method for presenting the quantitative evaluation results for the selected features is instead introduced. It should be reiterated that the multiple-site damage data were not used for feature selection.

4.2. Quantitative feature evaluation

The evaluation process generates a large amount of graphical data, and a summary statistic was sought with which to evaluate the discriminatory effectiveness of the selected features. Receiver Operating Characteristic (ROC) curves offer such a measure, providing a simple benefit (true-positive rate) vs. cost (false-positive rate) analysis for two-class data. The area under the ROC curve (AUC) was adopted as a summary statistic for quantifying the discriminative ability of the selected features. A brief introduction to ROC curves and their application to the feature evaluation problem are given here. A more complete introduction to ROC analysis is given in [21].

The Receiver Operating Characteristic gives a comparison between the *true positive rate* (TPR) and *false positive rate* (FPR) operating characteristics for a given threshold level. Evaluating the ROC value for a range of possible threshold values

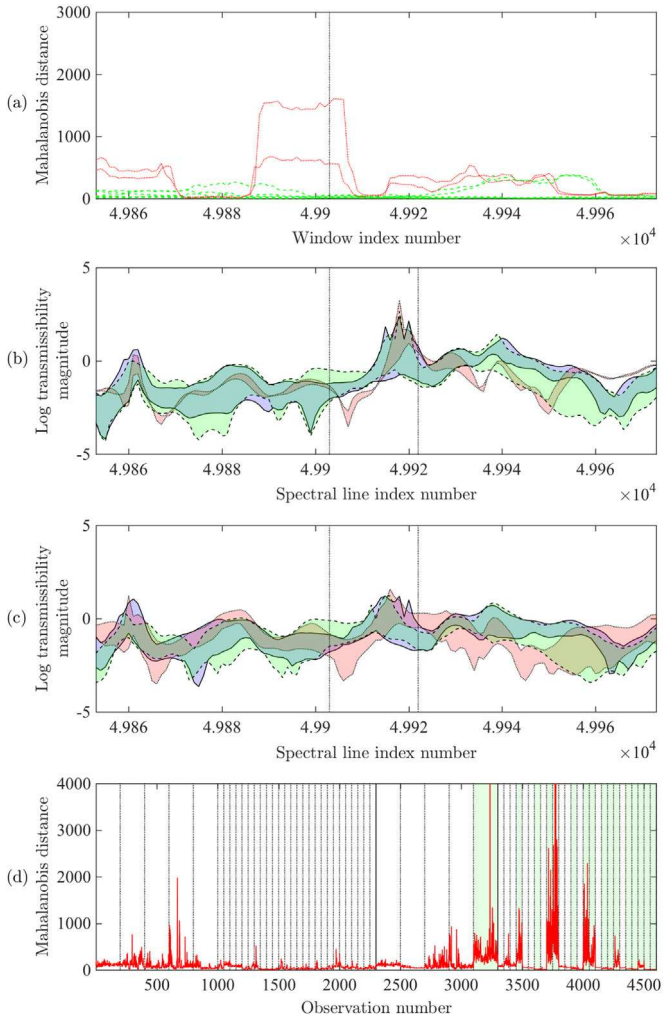

Fig. 10. Feature selection and evaluation plots for feature F_{49}. This feature was deemed to have performed comparatively poorly on the testing dataset.

enables the plotting of an ROC curve. The ROC curve may be employed to evaluate the discriminative ability of the data or to select a level for the threshold that is appropriate to the problem.

It is for the first of these functions that it is applied here. As an example, the curve in Fig. 11 offers an insight into the discriminatory performance of the feature F_{46}, intended to detect the removal of panel P_5. A negative class label implies the panel has not been removed and a positive class label implies that it has. The data used is from the multiple-site damage case. The dashed line corresponds to a classifier offering no discrimination. The further the curve corresponding to a classifier lies above and to the left of this line, the greater its discriminatory performance. In this case, the discriminatory performance of this feature for the data presented may qualitatively be described as good. A quantitative evaluation of the level of discrimination comes from considering the area under the ROC curve.

The area under the ROC curve (AUC) is a simple means of further summarising classifier performance. As the name suggests, the AUC is given by the area lying below the plotted Receiver Operating Characteristic curve (see Fig. 11). The AUC characteristic takes a value in the range [0 1] and has a useful interpretation as the probability that the classifier will rank a randomly selected positive instance higher than a randomly chosen negative instance. For the example of feature F_{46} given in Fig. 11, the AUC value when discriminating between panel-on and panel-off data for Dataset B was 0.778. This is taken to represent a comparatively poor level of discrimination between classes.

For the data considered, the two classes are *specified panel removed* and *specified panel not removed*. The AUC is adopted as a summary statistic for comparing the classification performance of the individual features in separating normal condition data first from single-site data and then from multi-site data.

For each of the 50 features identified using Dataset A, two AUC values were calculated. The first curve illustrates the ability of the feature to separate normal data from single-site damage data for Dataset B. The second illustrates the ability of the feature to separate normal data from multi-site damage data for Dataset B. The desired outcome is that the individual features would display a good degree of separability, evidenced by an AUC value of close to 1, for both cases. The values

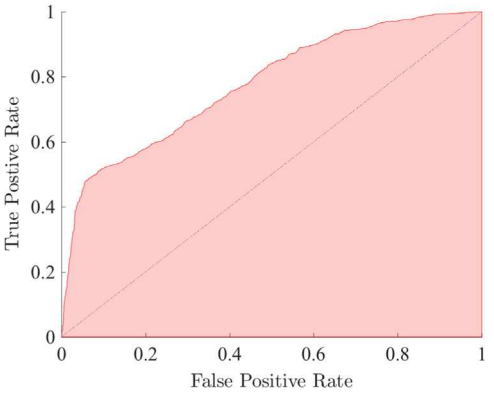

Fig. 11. Area under the ROC curve for feature F_{46}.

calculated for the discrimination of single-site and multi-site damage for each feature, plus a summary of the overall outcome for each panel, are presented in Table 4.

4.3. Summary of feature evaluation outcomes

The outcome of applying the features selected using the training dataset to the testing dataset may be summarised thus:

- A mean AUC value of 0.977 across all features was achieved for the single-site panel removal.
- A mean AUC value of 0.966 across all features was achieved for the multi-site panel removal.
- Perfect classification was achieved by 18 of the 50 individual features for single-site panel removal.
- Perfect classification was achieved by 19 of the 50 individual features for multi-site panel removal.

The overall performance of the identified features was deemed to be excellent, with some individual exceptions. Some of the features 'fired' unexpectedly when presented with normal condition observations. A small degree of inter-test variability between Datasets A and B is observed through comparison of normal condition spectra for these features. This finding serves to reiterate the importance of gathering a training set that is truly representative of the conditions that may be encountered.

It is also found that the removal of some of the panels is distinctly more easily detected than the removal of others, and that while some transmissibility paths provide a rich source of features, others offer little discrimination between states. It is notable that transmissibility paths comprising response sensors that are separated from the damage location by a stringer appear to be less successful than those for which the reference sensor is close to the damage. This leads to the suggestion that it may be beneficial to position sensors to be proximal to the expected damage location. In principle, regions of high stress could be predicted via numerical modelling. Of greater practical relevance in situations where such predictions are not available is the further suggestion that for stiffened-panel structures such as the aircraft wing considered, it may be beneficial to place at least one sensor within each panel (*panel* in this sense being a region of the wing top-sheet bounded on all sides by stiffening elements) that is to be monitored for damage. While these suggestions are perhaps somewhat unsurprising, this is nevertheless a useful demonstration of the importance of sensor placement, and gives some insight into how an extended network of sensors for detecting top sheet damage in similar structures may be formed.

Overall, it was found that the individual features selected on the basis of the training Dataset A (comprising normal and single-site panel removal data) performed very well when presented with single-site panel removal data from a previously unobserved dataset. This result, while expected, serves to validate the test sequencing introduced in Section 2.4.

Of primary interest, however, is the finding that the features also performed very well when presented with multi-site panel removal data, despite no multi-site data being used for feature selection. The multi-site performance of the classifiers averaged over the 50 selected features is very close to that when presented with single-site data, and is evidenced both through visual inspection and through applying quantitative measures. The conclusion drawn is that there is support for the stated hypothesis: for the structure states investigated, it has been possible to find features that offer a high degree of discrimination between multiple-site damage states despite only the single-site damage states having been observed. The significance of this result is the suggestion that the problem of the explosion in the number of damage states that must be observed in order to build a classifier capable of identifying multiple-site damage may be circumvented in some circumstances. It is, of course, not possible to draw general conclusions on the basis of a single case study, and further investigation into the validity of the approach is warranted.

It should be remembered that while evaluating features, the primary concern is that they should fire when damage occurs at the specified location and not fire when the structure is undamaged. It is of only secondary concern that the feature should not fire when damage is not present at the specified location, but is present at other locations on the

Table 4
Area under ROC curve results for Dataset B.

Panel	Feature	AUC Single	AUC Multi	Feature	AUC Single	AUC Multi	Mean AUC Single	Mean AUC Multi
P_1	F_1	0.994	0.989	F_6	0.902	0.893		
	F_2	1.000	1.000	F_7	1.000	1.000		
	F_3	0.896	0.922	F_8	0.999	0.998	0.970	0.970
	F_4	1.000	1.000	F_9	1.000	1.000		
	F_5	0.913	0.901	F_{10}	0.998	0.998		
P_2	F_{11}	1.000	1.000	F_{16}	0.994	1.000		
	F_{12}	1.000	1.000	F_{17}	1.000	0.999		
	F_{13}	1.000	1.000	F_{18}	1.000	1.000	0.999	1.000
	F_{14}	1.000	1.000	F_{19}	0.999	1.000		
	F_{15}	1.000	1.000	F_{20}	0.993	1.000		
P_3	F_{21}	0.998	0.999	F_{26}	0.999	0.998		
	F_{22}	0.986	0.984	F_{27}	1.000	1.000		
	F_{23}	0.996	0.997	F_{28}	0.947	0.983	0.991	0.993
	F_{24}	1.000	0.983	F_{29}	0.985	0.991		
	F_{25}	1.000	1.000	F_{30}	1.000	0.999		
P_4	F_{31}	1.000	1.000	F_{36}	0.990	0.999		
	F_{32}	1.000	1.000	F_{37}	0.999	0.999		
	F_{33}	0.979	0.988	F_{38}	0.994	0.997	0.993	0.997
	F_{34}	0.991	0.996	F_{39}	1.000	1.000		
	F_{35}	0.981	1.000	F_{40}	0.996	0.995		
P_5	F_{41}	0.999	0.998	F_{46}	0.972	0.986		
	F_{42}	0.846	0.760	F_{47}	0.991	0.778		
	F_{43}	0.966	0.983	F_{48}	0.975	0.976	0.933	0.867
	F_{44}	0.687	0.794	F_{49}	0.975	0.537		
	F_{45}	0.973	0.919	F_{50}	0.946	0.941		
						Overall:	0.977	0.966

structure. If such features are found, it is intuitively easier to interpret the patterns of firing features. However, the application of statistical pattern recognition methods is specifically intended to reveal underlying patterns within the observed data that would at the very least be challenging for a human observer to interpret.

Having identified features that are individually capable of discriminating between damage states to a good degree, focus moves to the application of statistical pattern recognition in order to employ the features for damage classification. The expectation is that combining the features should offer a further improvement in discriminatory performance. In the following section, the features identified form the basis of statistical classification using support vector machines.

5. Multiclass Support Vector Machines

The final stage of the SPR paradigm is the statistical modelling of the selected features. Multiclass support vector machines (SVMs) are investigated for this purpose. The hypotheses to be tested are that:

1. Multiclass SVMs may be an effective option for the classification of multi-site damage data, where observations of each class are available for training, and;
2. That a classifier trained using observations of normal and single-site damage data may offer a level of classification accuracy approaching that of a classifier trained using observations of all damage states. when applied to a testing set that includes multi-site damage observations.

To address these hypotheses two classifiers are constructed and tested. Classifier 1 is trained using data from the structure in all its damage states, including observations of multi-site damage (Dataset B). Classifier 2 is trained using data from the structure only in its normal and single-site damaged states (Dataset A). The objective for both classifiers is to achieve a high rate of correct classification when presented with a testing set of previously unseen data, which contains observations from the structure in single-site, multiple-site and undamaged states. The same architecture is used for both classifiers, and they differ only in the data used for training and validation.

SVMs have been demonstrated to possess several properties that make them well-suited to the damage identification task. They have been shown to generalise well from the small datasets often encountered in damage identification problems, and have the attractive property that they can support different classes of discriminant function (for example linear, polynomial or radial basis functions) without requiring substantial modification of the basic learning algorithm. The multiclass SVM strategy employed in this study is to train a 'one-against-one' binary SVM to identify the occurrence of damage, and an ensemble of 'one-against-all' binary SVMs to locate damage. The structure of the classifier system is described in the following section.

5.1. Classifier architecture

The classifiers used in this study were created using the MATLAB Support Vector Machine Toolbox. The classification architecture is based upon the binary Support Vector Machine (SVM) extended to the multi-class problem, with each 'classifier' in fact comprises an ensemble of 6 binary SVM classifiers. The first SVM (labelled SVM_1) seeks to separate damaged state data from normal state data and thus acts as a damage detection step. SVM_1 is unique among the 6 SVMs employed in being trained using all 50 features identified in Table 3; as all 50 features are sensitive to damage, it is logical to include them all when training the classifier. In principle down-selection could be pursued in order to reduce the dimension of the feature set, but this was not found necessary in the current study.

Five further binary SVMs (labelled $SVM_{2\text{-}6}$) seek to indicate whether removal has occurred for each of the five panels in turn. Each SVM seeks to class an individual location as 'undamaged' or 'damaged'- SVM_2 seeks to classify panel 1 as on or off, SVM_3 relates to panel P_2 etc. The classes and features used are summarised in Table 5. During initial development, it was found that the performance of the individual classifiers $SVM_{2\text{-}6}$ was substantially diminished if normal condition data was included in the training set. As such, training of these classifiers was conducted using a training set comprising damaged state data only. It was also found that greatly improved results were achieved for classifiers $SVM_{2\text{-}6}$ when only the 10 features selected for each panel were used for classification, instead of making use of all 50 features. This agrees with the expectation that the performance of these 'damage location' SVMs may be promoted by restricting them to a space containing only those features deemed sensitive to damage at the location of interest.

A key decision in the application of support vector classification is the choice of SVM kernel. The choice of kernel effectively acts as a form of regularisation during training, preventing overfitting to the training data. An initial survey of the feature data indicated that there was a clear argument for adopting a non-linear kernel. In the absence of specific knowledge on the form of the classifying hyperplane, the radial basis function (RBF) kernel was adopted in the discriminant function. This choice was made due to the low degree of parameterisation of the RBF kernel (only two hyperparameters must be set) and its consequent preference for selecting a smooth classification bound. In principle, cross validation could be employed to select between competing kernel options. This was not employed in the current work as it was felt that given the relatively high degree of separability of the data considered this would be unlikely to influence the quality of classification.

The data presented to the classifier are log discordancy values. Normalisation of the data is recommended to aid the conditioning of the optimisation problem. In this instance, each feature was normalised to the interval [0 1], with 1 being the maximum value of the feature observed in Dataset A. Where the outcomes of $SVM_{1\text{-}6}$ are contradictory, for example where SVM_1 indicates no damage has occurred but damage is nevertheless indicated by the location SVMs, precedence is given to the prediction of the detection classifier SVM_1. In cases that damage is indicated to have occurred to the structure, but damage is not identified at any of the five locations monitored, observations are labelled 'Unclassified'. Classifier 1 and Classifier 2 differ only in the data that is used in their development: all other factors (classifier architecture, validation procedure and features employed) are kept the same. With the structure of the classifiers specified, attention is turned to the data used to train, validate and test the classifiers.

5.2. Training, validation and testing sets

Two experimental datasets are used for developing and testing the classifiers. The first, referred to as Dataset A, comprises 1000 normal state and 1000 single-site damage state observations of the wing structure. The second dataset, referred to as Dataset B, comprised 2300 normal, 1000 single-site damage and 1300 multi-site damage state observations. Conducting supervised learning in a principled fashion necessitates separating the data into three, non-overlapping sets: the testing set, validation set and training set. Each serves a purpose in the development and testing of the classifier.

- The training set is used to set the values of the classifier parameters
- The validation set is used to set the values of the classifier hyperparameters.
- The testing set is used to verify that the developed classifier works for an independent set of observations.

For the RBF kernels employed in this study, two hyperparameters must be set: the misclassification tolerance parameter C, and the radial basis kernel width α. The validation step allows an informed decision as to the most appropriate values of

Table 5
Classes and features employed for binary SVMs.

SVM	Features employed	Positive class	Negative class
SVM_1	F_{1-50}	All Undamaged Data	All Damaged Data
SVM_2	F_{1-10}	Damaged, P_1 removed	Damaged, P_1 not removed
SVM_3	F_{11-20}	Damaged, P_2 removed	Damaged, P_2 not removed
SVM_4	F_{21-30}	Damaged, P_3 removed	Damaged, P_3 not removed
SVM_5	F_{31-40}	Damaged, P_4 removed	Damaged, P_4 not removed
SVM_6	F_{41-50}	Damaged, P_5 removed	Damaged, P_5 not removed

the hyperparameters to be made in order to meet a defined set of objectives. The principal objective for the current study is to maximise the probability of correct classification when applied to the validation set. In cases where perfect correct classification was achieved on the validation data, two further objectives were set: maximisation of the margin of separation; and minimisation of the number of support vectors required. It was found that the classifiers developed in the present study were comparatively robust to hyperparameter choice. As a result, for the purposes of this study, validation was pursued by generating a full factorial set of feasible hyperparameter combination, and manually selecting the best-performing hyperparameter pair for each SVM in turn. Further detail on the validation process may be found in [22]. It should be noted that, while not deemed necessary here, it is in general recommended that a formal cross-validation process be undertaken at the validation stage.

Datasets A and B are each separated into training, validation and testing sets, denoted A_{train}, A_{val} and A_{test} for Dataset A and B_{train}, B_{val} and B_{test} for Dataset B. The distribution of observations in each set is summarised in Table 6, where N refers to normal state observations; D_s refers to single-site damage observations; and D_m refers to multi-site damage observations. A consideration in separating the data was that the training and validation sets drawn from Datasets A and B be the same size so as not to favour either Classifier 1 (trained using Dataset B) or Classifier 2 (trained using Dataset A). Recall that Dataset A alone was used for feature selection.

5.3. Classifier 1

The first hypothesis considered is that SVMs may be an effective option for the classification of multi-site damage data, where observations of each class are available for training. This hypothesis is addressed using Classifier 1. Classifier 1 is trained using the training and validation sets B_{train} and B_{val}. These sets include observations of the structure in its normal, single-site damaged and multiple-site damaged states. The ability of the developed classifier to correctly identify damage is evaluated using the testing set B_{test}. The validity of the SVM classification approach for the presented multiple damage identification problem may thus be assessed. The performance of Classifier 1 is also taken as being a 'best case' benchmark, given the features selected and the modelling decisions made. It is against this benchmark that the performance of Classifier 2 may be measured.

5.4. Classifier 2

The second hypothesis considered is that a classifier trained using observations of normal and single-site damage data may offer a level of classification accuracy approaching that of a classifier trained using observations of all damage states when applied to a testing set that includes multi-site damage observations. This hypothesis is addressed using Classifier 2. Classifier 2 is developed using the training and validation sets A_{train} and A_{val}, which contain observations of normal and single-site damage only. The classifier is again evaluated using testing set B_{test}, which contains observations of all damage states. The performance of a classifier trained in the absence of observations of multi-site damage data when presented with observations from all damage states may thus be assessed. The performance of this classifier is compared to that of the 'best case' performance given by Classifier 1. In addition to being trained using a greatly reduced number of damage states, this classifier is trained and tested using data from two separate tests conducted several days apart. This is a somewhat severe test of the capacity of the developed classifier to generalise not only to newly presented data that may contain unseen states, but which may also be subject to some degree of inter-test variability. It is common in applications of statistical pattern recognition to real data for a single dataset to be used, and for interleaved data to be used for training, validation and testing.

6. Results

6.1. Classifier 1: multi-site damage

Classifier 1 represents the 'best case' scenario of having data available from all classes, and enables the validity of applying the SVM classification approach to a multiple damage identification problem to be evaluated. The classifier is trained using dataset B_{train}, which comprises 250 normal condition observations, 24 observations for each of the 5 single-site damage states and 5 observations for each of the 26 multiple-site damage states. The model hyperparameters are set using the validation set B_{val}.

Table 6
Separation of observations in Datasets A and B into training, validation and testing sets.

	Dataset A			Dataset B		
	N	D_s		N	D_s	D_m
A_{train}	250	250	B_{train}	250	120	130
A_{val}	250	250	B_{val}	250	120	130
A_{test}	500	500	B_{test}	1800	760	1040

The classifier is tested on the testing dataset B_{test}, which contains 1800 normal condition observations, 152 observations for each of the 5 single-site damage states and 40 observations for each of the 26 multiple-site damage states. The classification results are presented in confusion matrix form in Fig. 12 and summarised in Table 7.

The overall classification results are excellent. For two of the five single-site damage classes, a perfect (100% correct) rate of classification was achieved. The results are even more encouraging for the multiple-site damage classes, with a perfect rate of classification for 25 of the 26 classes. There was in fact a single misclassification from a total of 1040 observations of multiple-site damage data, with damage state $D_{\{1,2,3\}}$ misclassified as $D_{\{1,2,3,5\}}$. It is worth noting that every misclassification involved panel P_5. For the single-site damage, six observations of state $D_{\{2\}}$ were classified as $D_{\{2,5\}}$; four observations of state $D_{\{4\}}$ were classified as state $D_{\{4,5\}}$; and three observations of $D_{\{5\}}$ were classified as undamaged. Of the 1800 undamaged state observations, eight were classified as the removal of P_5. It was found in Section 3 that the selection of features for panel P_5 was more challenging than for panels P_1-P_4. This difficulty in identifying 'strong' features for panel P_5 manifests itself as an adverse effect, albeit limited in severity, on the performance of the classifier in relation to the panel. This highlights the critical dependence of the classifier performance on the quality of features that can be identified. This may be of concern in the identification of smaller damage extents than the panel removal case investigated in this study.

The performance of this 'best case' classifier gives confidence that the SVM architecture used is suitable for identifying multiple-site damage for the problem at hand. Attention now moves to whether such a classifier can be developed in the absence of training data for the multiple-site damage cases.

6.2. Classifier 2: multi-site damage

The true test of the classifier comes when it is applied to the testing set B_{test}. This dataset contains multi-site damage data, in addition to previously unseen observations of the structure in its undamaged and single-site damaged states. An additional confounding factor is that the data comes from a separate testing programme to that used for training: the classifier is being asked to generalise across tests, handling any inter-test variability that may have arisen. The full classification results are presented in confusion matrix form in Fig. 13. The results are summarised, in terms of the probability of perfect correct classification for the three sub-categories of structural states in Table 8

The classification outcomes presented in Fig. 13 and Table 8 are highly encouraging. For the structure in its normal state, there were no false indications of damage, and 1769 of the 1800 observations tested were correctly classified as undamaged. The remaining 31 observations were returned as 'unclassified', with SVM_1 indicating damage but $SVM_{2\text{-}6}$ not indicating that damage had occurred on panels P_1-P_5. This is taken as an indication that the classifier failed to generalise fully between tests, and may be attributable to the training set not fully representing the variability that may arise between tests. A second observation is that, while the objective of SVM_1 was damage identification, the 50 features used were selected based essentially upon damage location criteria. The criterion for selecting the features was that they should be indicative of damage on individual panels, rather than indicative of damage regardless of source.

The classifier performed exceptionally well for the single-site damage states, with only 5 non-correct classifications out of 760 observations. Of these, three were returned as 'unclassified' with SVM_1 indicating damage but SVM_5 failing to identify that panel P_4 had been removed. A further two observations were misclassified, with the classifier indicating the removals of panels P_2 and P_3 where in fact only panel P_2 had been removed. This was one of only two instances in which the classifier falsely indicated damage at any location. The other was for the removal of panels P_1 and P_2, for which the classifier indicated that panel P_5 had also been removed.

In total, there were 128 observations (out of a total of 1040 observations) for which the classifier trained solely on single-site data failed to perfectly identify multi-site damage states. Of these, 125 observations missed the removal of one panel but were otherwise correct. 118 of these observations were due to removal of panel P_5 having been missed; the remaining seven were due to panel P_2 being missed. The suggestion, as initially raised in Section 3, is that the feature set identified for indicating the removal of panel P_5 is comparatively 'weaker' than those for the remaining four panels. This will warrant further investigation if the approach is to be applied for less severe damage scenarios for which less discriminatory feature sets will be available. Alternative methods of feature selection and normalisation may be sought in such cases.

Overall, the success of the classifier when applied to multi-site data is encouraging. It appears that given a suitably discriminatory feature set, the SVM approach is capable of achieving an exceptionally high level of correct classification for single-site damage, and a very good level of classification for the much more challenging task of identifying multi-site damage.

7. Conclusions

This study presents a promising approach for handling multi-site damage location problems without requiring training data from all damage states. First, it is shown that a multi-class classifier based on binary support vector machines (SVMs) is a suitable approach for handling multi-class damage identification problems: for the experimental case presented, near perfect results were achieved when damaged-state data for all the structural states are incorporated in the training set. The second, and major, contribution of the study is to show that a classifier trained using *only* normal condition and single-site damage data may be capable of identifying the presence of multi-site damage with a high degree of accuracy. This has the

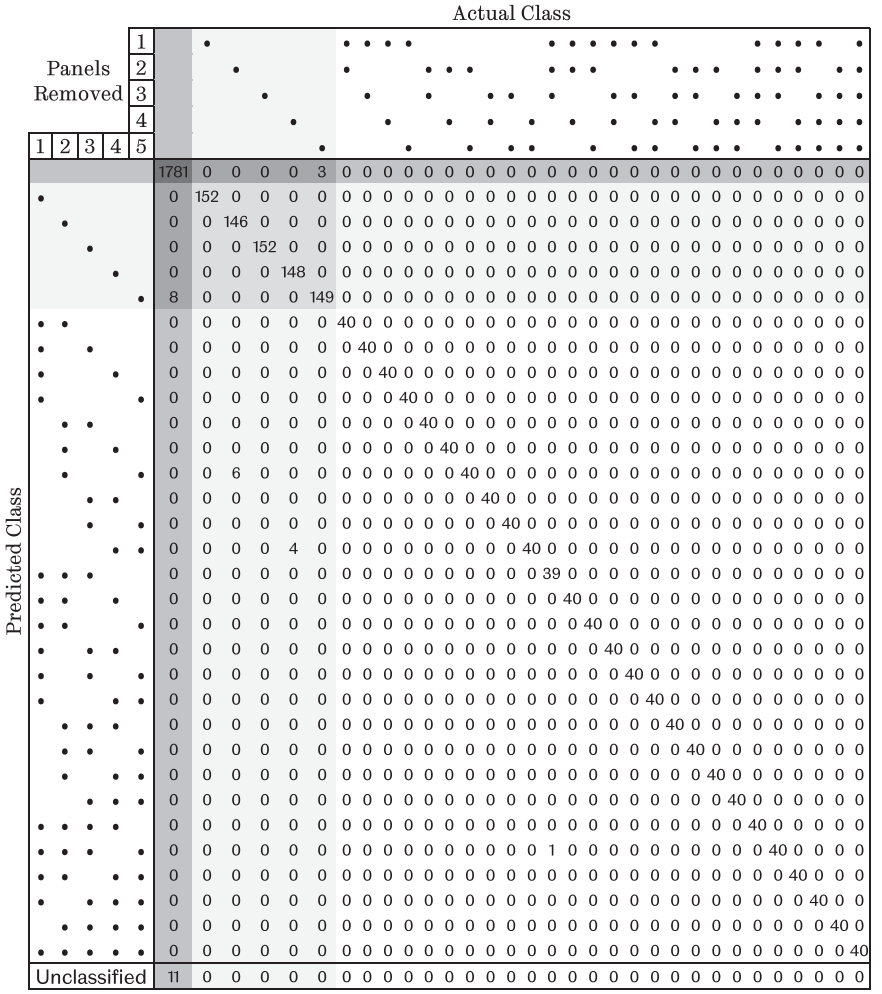

Fig. 12. Confusion matrix results for Classifier 1.

Table 7
Summary of results for Classifier 1 applied to testing set B_{test}.

Condition	Correctly classified	Incorrectly classified	Unclassified
Undamaged	98.94%	0.50%	0.61%
Single-site damage	98.29%	1.71%	0.00%
Multi-site damage	99.90%	0.10%	0.00%

potential to hugely reduce the number of states for which damaged condition data is required when training a data-based classifier.

The case investigated is challenging. However, the removal of inspection panels represents a relatively gross level of damage, and the set of damage locations are both small in number (five in total) and coarsely dispersed. A concern that remains is whether the degree of success observed can be maintained in more challenging scenarios, for example at lower levels of damage, higher numbers of locations and in the presence of environmental variability. This concern is promoted by the difficulty experienced in achieving perfect diagnosis for panel P_5. The features selected for identifying this panel were somewhat weaker than those found for other panels: this might be expected given that the structure is mounted in a cantilever configuration, with panel P_5 situated furthest from the mounting point and so in the region of lowest surface stress. It is likely that the true efficacy of the method will only become apparent when considering more challenging cases.

The first area warranting further study is whether general rules can be learnt regarding the behaviour of the structure. The approach relies upon the presence of features that are sensitive to the removal of one panel in isolation and which also 'fire' when that panel is one of several removed, but little regard has been given to the physical phenomena underpinning this behaviour. A possible explanation arises through considering the structure of the wing. In the case presented in this

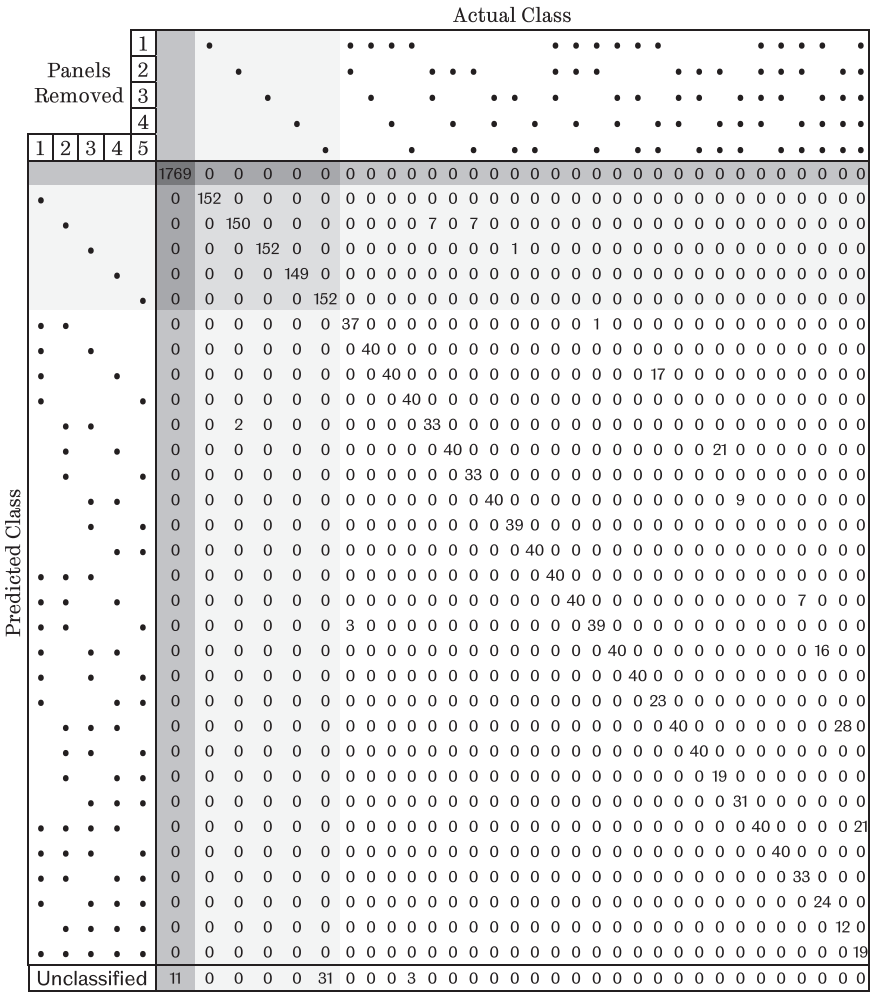

Fig. 13. Confusion matrix results for Classifier 2.

Table 8
Summary of results for Classifier 2 applied to testing set B_{test}.

Condition	Correctly Classified	Incorrectly Classified	Unclassified
Undamaged	99.28%	0.00%	1.72%
Single-site damage	98.34%	0.28%	0.39%
Multi-site damage	86.73%	13.27%	0.00%

study, transmissibility data were employed as raw data and reduced to low-dimensional features using methods from outlier analysis. The investigated section of the wing is made up of a thin aluminium topsheet mounted on a framework of more substantial stiffening elements. This results in the topsheet being divided into many rectangular regions of flexible material bounded on each side by more rigid elements. Engineering insight suggests that the coupling between these regions may be rather weak, with the effect that stiffness changes on individual regions would be expected to be most obvious in the response measurements taken from that region and less so in responses taken from other locations. It is conceivable that this may lead to a degree of independence between regions, which would, intuitively, aid the ability of features selected using single-site damage data to generalise to multiple-site damage states. While beneficial in this case, further investigation of features that display behaviour amenable to the presented approach is warranted. The broad conclusion to be drawn is that while the approach pursued in this study has been shown to be successful despite not requiring a law-based model for reference, its success is still reliant on having some insight into the physical underpinnings of the observed phenomena. For the case investigated, the nature of the structure may have played a role in 'localising' the behaviour of some features to particular, bounded regions of the structure, and consequently making the multiple damage location problem tractable. The priority for future work related to this portion of the work undertaken should be to establish

whether this notion is valid. If so, an extension may be to investigate whether other sub-structures that exhibit comparable similar behaviour exist.

Further areas for study are whether SVMs represented the most appropriate classifier choice; and how data condensation may best be applied to maximise classifier performance while minimising the quantity of data required for training. Several options for the architecture and validation of multi-class SVMs exist. Furthermore, comparison to other classification approaches was not pursued in this instance. While the particular SVM approach adopted worked well for the wing example, it is expected that the specifics of the classification approach adopted will grow in importance as the discriminative capacity of the feature set decreases and/or the number of considered damage locations grows. However, while comparison of options for multi-class identification may be among the areas that warrant further study, the performance of the classifier is essentially limited by the quality of information used for training. The priority for future work should be on the gathering of high quality, low dimensional discriminative data rather than the choice and tuning of the adopted classification tool. Work is ongoing to investigate the approaches for combined feature extraction and sensor placement optimisation for this purpose.

Finally, while the approach developed is intended to reduce the amount of data required for the multiple-location damage problem, the problem of sourcing even single-site damage data remains and is an open topic of research. Approaches being investigated include the use of 'pseudo-faults' such as mass addition; learning from populations of similar structures; and the use of 'hybrid methods' that make use of physics-based model predictions in the training process. It is the opinion of the authors that the latter of these options should be the primary focus of future work as long term these appear to offer the most promising route towards practicable multi-site damage identification. A further possibility is that the model could be used to conduct initial sensor placement optimisation and feature set evaluation, with gathering of training data subsequently conducted in an entirely data-based fashion.

Acknowledgement

The authors gratefully acknowledge the support of EPSRC Grant EP/E010849/1 Smart Sensing for Structural Health Monitoring (S3HM). Portions of this abridged work were first published in full form by R. J. Barthorpe in his Ph.D. thesis [22].

References

[1] C.-P. Fritzen, P. Kraemer, Self-diagnosis of smart structures based on dynamical properties, Mech. Syst. Signal Process. 23 (6) (2009) 1830–1845.
[2] A. Rytter, Vibration Based Inspection of Civil Engineering Structures, (Ph.D. thesis), Aalborg University (1993).
[3] S.W. Doebling, C.R. Farrar, M.B. Prime, D.W. Shevitz, Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review, Tech. rep., Los Alamos National Lab., NM (United States) 1996.
[4] H. Sohn, C.R. Farrar, F.M. Hemez, D.D. Shunk, D.W. Stinemates, B.R. Nadler, J.J. Czarnecki, A review of structural health monitoring literature: 1996–2001, Los Alamos National Laboratory, USA.
[5] R. Ruotolo, C. Surace, Damage assessment of multiple cracked beams: numerical results and experimental validation, J. Sound Vib. 206 (4) (1997) 567–588.
[6] T. Contursi, A. Messina, E. Williams, A multiple-damage location assurance criterion based on natural frequency changes, J. Vib. control 4 (5) (1998) 619–633.
[7] H. Guo, Z. Li, A two-stage method to identify structural damage sites and extents by using evidence theory and micro-search genetic algorithm, Mech. Syst. Signal Process. 23 (3) (2009) 769–782.
[8] M.W. Vanik, J. Beck, S. Au, Bayesian probabilistic approach to structural health monitoring, J. Eng. Mech. 126 (7) (2000) 738–745.
[9] S. Sankararaman, S. Mahadevan, Bayesian methodology for diagnosis uncertainty quantification and health monitoring, Struct. Control Health Monit. 20 (1) (2013) 88–106.
[10] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: a review, Mech. Syst. Signal Process. 56 (2015) 123–149.
[11] B. Titurus, M. Friswell, Damage detection using successive parameter subset selections and multiple modal residuals, Mech. Syst. Signal Process. 45 (1) (2014) 193–206.
[12] C.R. Farrar, T.A. Duffey, S.W. Doebling, D.A. Nix, A statistical pattern recognition paradigm for vibration-based structural health monitoring, Struct. Health Monit. 2000 (1999) 764–773.
[13] G. Manson, K. Worden, D. Allman, Experimental validation of a structural health monitoring methodology: Part iii. damage location on an aircraft wing, J. Sound Vib. 259 (2) (2003) 365–385.
[14] M. Cao, M. Radzieński, W. Xu, W. Ostachowicz, Identification of multiple damage in beams based on robust curvature mode shapes, Mech. Syst. Signal Process. 46 (2) (2014) 468–480.
[15] K. Worden, G. Manson, D. Allman, Experimental validation of a structural health monitoring methodology: Part i. novelty detection on a laboratory structure, J. Sound Vib. 259 (2) (2003) 323–343.
[16] G. Manson, K. Worden, D. Allman, Experimental validation of a structural health monitoring methodology: part ii. novelty detection on a gnat aircraft, J. Sound Vib. 259 (2) (2003) 345–363.
[17] N. Maia, J. Silva, A. Ribeiro, The transmissibility concept in multi-degree-of-freedom systems, Mech. Syst. Signal Process. 15 (1) (2001) 129–137.
[18] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
[19] K. Worden, G. Manson, N. Fieller, Damage detection using outlier analysis, J. Sound Vib. 229 (3) (2000) 647–667.
[20] G. Manson, E. Papatheou, K. Worden, Genetic optimisation of a neural network damage diagnostic, Aeronaut. J. 112 (1131) (2008) 267–274.
[21] T. Fawcett, An introduction to roc analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874.
[22] R.J. Barthorpe, On model-and data-based approaches to structural health monitoring, (Ph.D. thesis), University of Sheffield, 2010.