Novelty detection across a small population of real structures: A negative selection approach

Novelty detection across a small population of real structures: A negative selection approach

Giulia Delo[1], Massimiliano Mattone[1], Cecilia Surace[2], Keith Worden[3]

[1] Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy
[2] Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Turin, Italy
[3] Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

E-mail: giulia.delo@polito.it

Abstract. Vibration-based Structural Health Monitoring (SHM), exploits a variety of approaches for novelty detection. In particular, many data-based methods try to recognise patterns by exploiting analogies with the human body's natural defences at a cellular level. These algorithms fall within the Artificial Immune System (AIS) class and can be chosen, according to their peculiarities, to solve specific problems in diverse application areas. This study investigates the damage-detection process in different operational conditions, obtained by applying structural modifications to a laboratory-scale aeroplane, which follows the geometric features of the GARTEUR benchmark project. Damage identification is performed by exploiting the Negative Selection Algorithm (NSA), already applied by some of the authors on numerically-simulated case studies, and chosen for its capability of self/non-self discrimination under varying operational or environmental conditions. The research is expanded by using sparse autoencoders for feature dimensionality reduction. The method is applied to an experimental dataset acquired via Scanning Laser Doppler Vibrometer (SLDV) measurements, to identify consistent damage-sensitive features from the frequency response functions, and to obtain a reliable fault-detection performance.

1. Introduction

Several Structural Health Monitoring (SHM), approaches have been developed to ensure the safety and reliability of engineering structures, according to different levels of insight. The first objective is to identify the presence of anomalies or damage in the behaviour of a target system or structure. For this purpose, vibration-based methods are often adopted to assess damage conditions, exploiting their wide applicability and low intrusiveness. Vibration-based SHM often requires the use of a large variety of data correlated to different environmental and operational conditions to ensure robustness and to aid the interpretation of results. Regardless, the scarcity of such data often limits the extent of their applicability, which can be improved by adopting population-based approaches. In the SHM framework, computational models frequently draw inspiration from biological systems and the human immune system for solving optimisation issues for data-based novelty detection or pattern recognition. These models fall within the Artificial Immune System (AIS) class, which comprises Negative Selection Algorithms (NSA), the Immune Network Model, and Clonal Selection [1].

The NSA is a family of algorithms relying on the immune system's capability to discriminate between self and non-self cells. It has already been applied for SHM unsupervised novelty detection, adopting time-series features [2]. Subsequently, Surace & Worden explored the use of NSA for features based on transmissibility functions [3]. These damage-sensitive features can be extracted from the structure's dynamic response only. Thus, they could be used for in-flight monitoring, when any excitation is unknown [4]. Furthermore, by examining different paths for evaluating transmissibilities, these functions could not only indicate the presence of the anomaly but also gain insight into its location. An important feature of the negative selection approach is the possibility to consider different configurations associated with normal structural behaviour, which may occur because of changes in environmental or operational conditions; it was examined in [3], using two numerical case studies regarding an offshore platform and an aircraft wing.
This study aims to verify the algorithm applicability to an experimental case study regarding a steel aeroplane laboratory structure. Experimental modal analysis has been performed using a Scanning Laser Doppler Vibrometer (SLDV), which allows one to acquire the dynamic response on a dense grid of measurement points. The Frequency Response Functions (FRFs) obtained have been used to calculate Transmissibility Functions (TFs) between points of the monitored structure. Two normal conditions of the aeroplane have been analysed considering two diverse configurations of the winglets, and damage has been simulated in several locations on the structure by the introduction of concentrated masses. Negative selection algorithms have been examined in the literature, and some doubts emerged concerning their suitability for specific applications [5, 6, 7]. Some concerns relate to the dimensionality of the feature space. Accordingly, a sparse autoencoder is proposed for reducing the feature dimensionality before generating the detectors.
The layout of the paper is as follows. Section 2 presents the negative selection algorithm and the approach for feature dimensionality reduction. Section 3 describes the experimental case study regarding the aeroplane's dynamic response in different normal conditions (NCs), and damaged conditions (DCs), and the novelty detection results. Section 4 provides a discussion and some conclusions.

2. Methodology

The following section presents the main routine developed to implement the novelty detection algorithm.

2.1. Negative Selection Algorithm

The Negative Selection algorithms are biologically inspired one-class classifiers introduced by Forrest et al. [8]. They are inspired by the vertebrate adaptive immune system and its ability to distinguish self and non-self cells (pathogens), and consequently attack the latter. The immune system includes specific white blood cells, i.e., the lymphocytes, which are differentiated into B-cells and T-cells. T-cells are lymphocytes provided with T-cell receptors (TCR), also known as recognition units. They mature in the thymus to learn how to recognise antigens and consequently trigger an immune response. The process occurs via a randomly-based adaptive rearrangement, proliferation, and negative selection. The cells not involved in recognising antigens are destroyed via apoptosis, and new lymphocytes are generated. Thus, the negative selection produces a continuous and adaptive turnover, guaranteeing a remarkable capability of identifying antigens using a limited number of antibodies.
NSAs form a family of unsupervised algorithms which develops following these concepts. They consist of two main phases. First, there is the generation of detectors, those features selected so that they do not match the initial features (or normal condition) to cover the complementary space with respect to the normal condition feature space. The NSA features can be explained as hyperspheres, and the detectors are those hyperspheres surrounding the self-set. The self-set,

S, is examined via a self-matching rule to determine the distance between the different normal conditions, λ_s, and to establish the S radius, r_s, which is used as the matching threshold. The value of r_s should be sufficiently greater than λ_s to avoid false positives, but not too great, to avoid false negatives. The NSA then generates pseudo-random candidate vectors and compares them to the self-set to keep only those which do not fall within the matching threshold and computes each one's radius, r_d^i, similarly as previously outlined. This phase, called the *censoring phase*, ends when the required number of detectors is reached. Afterwards, there is the *monitoring phase*, in which it is analysed whether the novelty detectors bind with the new candidate features for identifying anomalies or damage. The NSA algorithm is explained in the flowchart shown in Fig 1.

The NSA is frequently applied to binary features. Instead, this study adopts the Real-Value Negative Selection (RVNS). The features are experimentally-obtained transmissibility functions subjected to appropriate encoding for dimensionality reduction. The RVNS is less explored than the NSA involving binary features, as the detector search space is continuous, and the generation of detectors is more difficult. However, it enhances interpretability, and it is necessary for some applications [9].

The effectiveness and sustainability of this method have been discussed in [5, 6], and some concerns arose regarding the runtime complexity, especially for high-dimensional features, the estimation of the right number of detectors, and their coverage of the non-self space. Therefore, a simple way to assess the detector's coverage consists of introducing a limit to the maximum number of consecutive attempts that fall in the already-covered area, as done in [9].

As previously outlined, the algorithm computes the radius, r_d^i of each detector in D. This radius could be considered a hyperparameter that may be adjusted to improve space coverage. However, in this study, the V-detector approach [10] is adopted, i.e., the radius is variable to enhance the NSA performance. This type of algorithm strongly depends on the choice of matching rule(s) used in the censoring and monitoring phases. The rule choice varies according to the type of feature being analysed. Considering ℓ-dimensional vectors, either the Euclidean distance or the Mahalanobis distance could be used. In this study, the cosine similarity is used, as in [3], as it ignores the magnitude difference between the two vectors:

$$\mathrm{sim}(\mathbf{x},\mathbf{y}) = \frac{\sum_{i=1}^{l} x_i y_i}{\sqrt{\sum_{i=1}^{l} x_i^2 \sum_{i=1}^{l} y_i^2}} \tag{1}$$

The cosine similarity is equal to 1 when the vectors are alike, and, consequently, the distance between two vectors is given by:

$$\mathrm{dist}(\mathbf{x},\mathbf{y}) = 1 - \mathrm{sim}(\mathbf{x},\mathbf{y}) \tag{2}$$

2.2. Feature dimensionality reduction

The first concerns about the negative selection approach regard the curse of dimensionality [6]. Therefore, the original self-set is subject to dimensionality reduction to address algorithm runtime issues and simplify classification.

The noise-contaminated transmissibility functions are used to train a sparse autoencoder, a particular symmetric neural network, which consists of an encoder and a decoder. The encoder learns how to represent the input feature by mapping it on a low-dimensional hidden layer, or bottleneck, and the decoder learns how to reconstruct the original feature vectors. This architecture can also be helpful for different applications, such as feature extraction, image compression or denoising, and novelty detection. When the autoencoder training is complete, the encoder is used to extract low-dimensional features, which can be used as input in the NSA, to generate the self-set S, and the monitored set M.

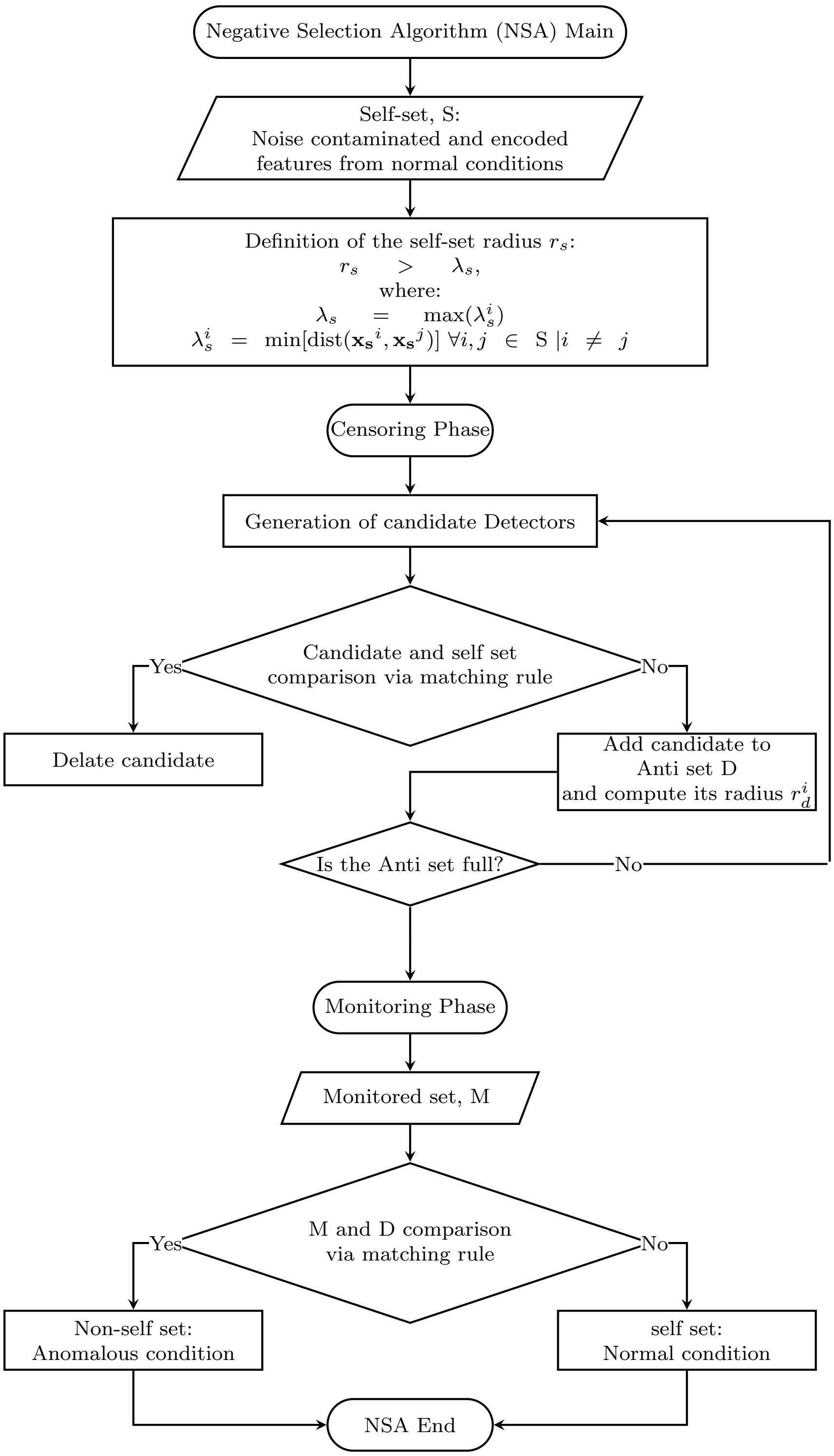

Figure 1. Negative Selection Algorithm flowchart.

The simultaneous training of the encoder and decoder consists of minimising a loss function representing the difference between the original feature and the reconstructed one. In sparse autoencoders (SAE), the loss function also includes a sparsity regularisation term (to limit the number of neurons that are activated at the same time), and a weight regularisation term. SAE uses sparsity, in addition to reducing the hidden-layer dimension, to create an information bottleneck, thus forcing the model to learn the essential characteristics of the input.

3. Experimental case study

3.1. Case study characteristics and experimental modal analysis

In order to verify the applicability of the NSA to an experimental case study, a steel laboratory structure has been considered. This is part of a broader population of similar structures, which mainly follows the geometry of the benchmark study conducted by the Structures and Materials Action Group (SM-AG19) of the Group for Aeronautical Research Technology in EURope (GARTEUR) [11]. The structure examined, simulates the dynamic behaviour of an aeroplane, and consists of a thin-section box fuselage, a single rectangular plate for realising the wings, and shorter plate elements for the winglets, the vertical tail and the horizontal tail. The fuselage's length is 1.50 m, the wing span is 2.00 m, and all plate elements have a 8.00 mm thickness. The aeroplane top view is shown in Fig 2.

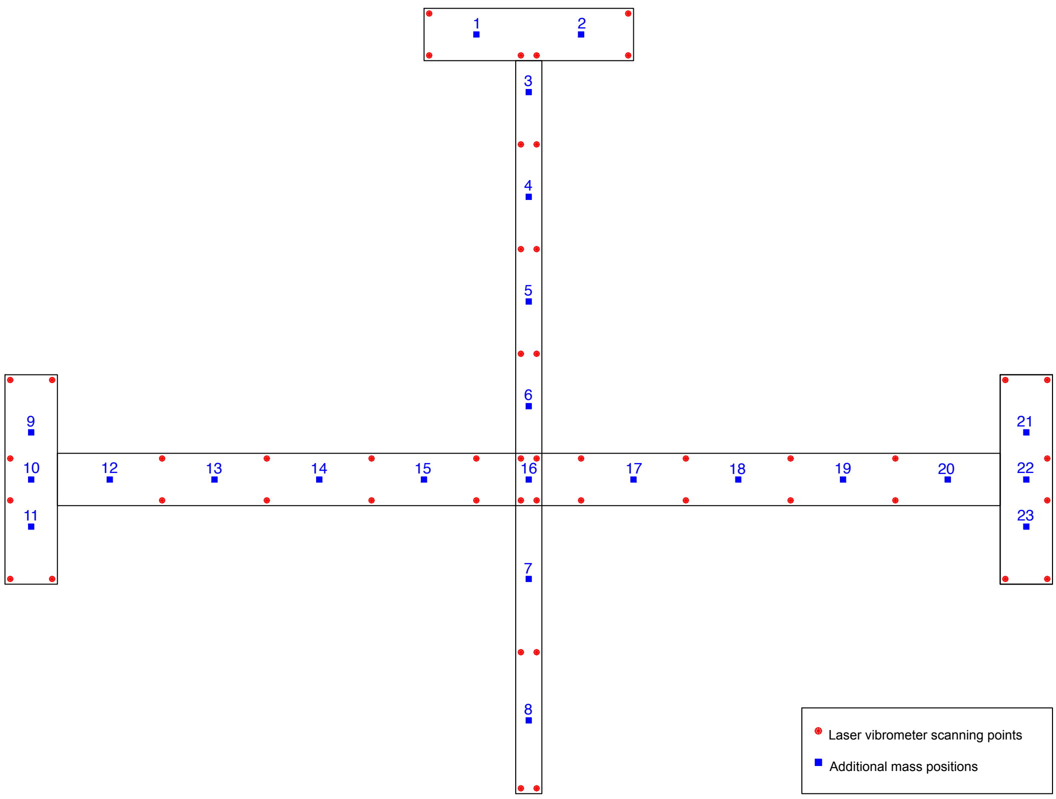

Figure 2. Garteur aeroplane top view, showing laser-vibrometer scan points in red, and simulated-damage locations in blue.

The experimental tests have been carried out in the "LAQ-AERMEC Aeromechanical Structural Systems" laboratory of the Department of Mechanical and Aerospace Engineering, Politecnico di Torino. The same structure has been analysed in two different configurations for investigating the novelty detection approach under varying operational conditions. These configurations have been achieved by performing an experimental modal analysis on the structure with and without the winglets. The aeroplane has been suspended, simulating free boundary conditions, and has

been excited by an electrodynamic shaker, which has been connected to the aeroplane using a stinger at the rear of the fuselage. The applied signal is a 1 V Amplitude Periodic Chirp in the 0-100 Hz frequency range.
The vibration response of the structure has been measured via a Polytec PSV-I-500 SLDV for identifying the natural frequencies and the main vibration modes. Experimental modal analysis based on SLDV allows one to quickly scan the velocity of a sufficiently dense grid of points. Subsequently, it is possible to obtain the response in terms of a time history, or in the frequency domain, compute the FRFs, extract modal parameters, and calculate a broad set of TFs. Tests have been repeated for each configuration, simulating damage via applying additional masses (around 2% of the aeroplane mass) in multiple positions over the entire structure. It can be assumed that true damage causes a change in the structure's stiffness, which can be simulated in Finite Elements models by locally reducing the Young's modulus [12]. Therefore, it is common to experimentally simulate damage by adding discrete masses, as it is a simple and reversible way of introducing changes in the dynamic properties of the structure. Indeed, the additional masses cause an eigenfrequency decrease, which is equivalent to the effect of a local stiffness loss [13, 14]. As shown in Fig 2, 23 damage locations have been tested in the configuration with the winglets, while 19 damage locations were considered for the no winglets configuration, i.e., excluding the four locations corresponding to winglet damage (see points 9, 11, 21, 23 in Fig 2).

3.2. Application of the algorithm and results

Among the set of experimentally-computed TFs, the NSA is applied to the transmissibility between the wingtips. It is a sufficiently wide feature to guarantee global damage sensitivity. Furthermore, the lower stiffness of the wings, compared to the fuselage, implies that the behaviour of the wings is the main actor in the entire dynamic response of the aeroplane. The original features consisted of 3200 spectral lines; however, only 2000 spectral lines were analysed. It was possible to neglect the final part of the spectrum, above the highest measured natural frequency, and the initial portion, because, for frequencies approaching zero, spectra are affected only by the structure's rigid body motions.
The self-set S, includes 1000 feature vectors obtained by making 500 identical copies of the TFs corresponding to the undamaged structure for each operational condition. Each feature has been contaminated differently with samples of multiplicative Gaussian noise of unitary mean and standard deviation, s=1%. The resulting set is used for autoencoder training. Afterwards, the trained encoder is used to extract eight-dimensional features from the noise-contaminated normal condition vectors for assembling the low-dimensional self-set S. Similarly, 100 copies of feature vectors from each normal and damaged condition of the two configurations are contaminated by noise and subjected to the encoder for assembling the monitored set M. The NSA is applied to the low-dimensional self-set S, to compute the detectors ${\mathbf{x_d}}^j$. Subsequently, the match between each monitored feature ${\mathbf{x_m}}^i$, and the set of detectors is evaluated via a novelty index NI, which is defined as follows:

$$\mathrm{NI}(i) = \min_j[\mathrm{dist}({\mathbf{x_m}}^i, {\mathbf{x_d}}^j) - r_d^j] \tag{3}$$

A negative value of NI implies that the corresponding feature is novel or anomalous. The Novelty index is shown in Fig 3. The novelty index is shown for the monitored set, M, including the normal and damaged conditions of the two configurations investigated. For each configuration, the first 100 points represent the indices of the polluted normal conditions, and 99.5% are correctly identified as not anomalous. Data from 101 to 2000 present the novelty indices of the damaged conditions in the no-winglet configuration, and data from 2101 to 4400 represent the damaged conditions in the winglet configuration. It can be observed that the novelty index is negative, and all these conditions are correctly classified as novelties. For further detail, it should be noted that the damaged conditions are ordered to have the first two DCs on the tail, followed by the 6 DCs relative to the two portions of the fuselage observed by the vibrometer,

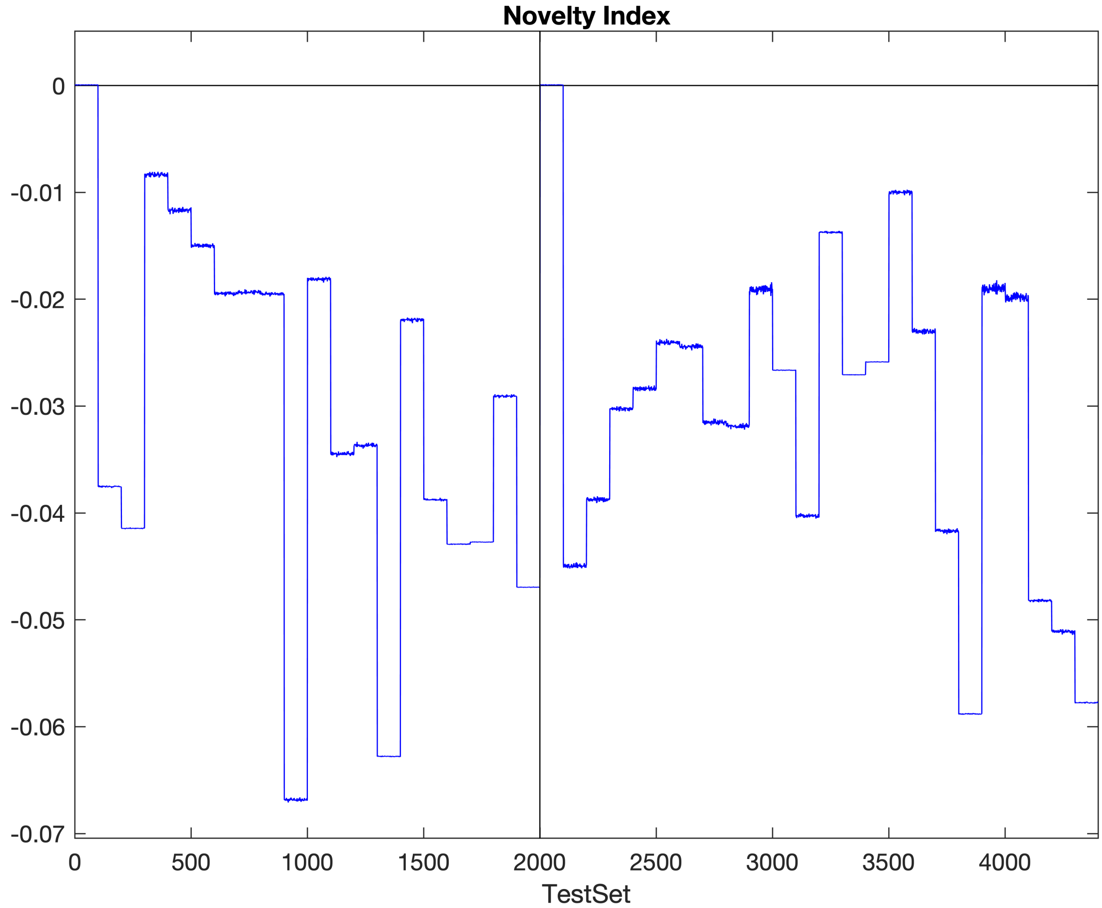

Figure 3. Novelty Index NI, for the monitored set M. The feature vectors regard the transmissibilities between the wingtips. The first 2000 features correspond to the configuration with no winglets, the subsequent ones correspond to the configuration with winglets.

and finally, the DCs of the wings from left to right, as shown in Fig 2.
By considering the transmissibility function along the fuselage, finding corresponding results can be more complex. This issue may be because of the lower sensitivity of the fuselage response to the simulated damage. Thus, achieving suitable damage-sensitive features may present a demanding challenge. Focussing on the wings, it is also possible to analyse more localised features, such as the one presented in Fig 4, which shows the novelty index corresponding to a transmissibility function between two points near the wingtip on the left side. By considering the transmissibility between two neighbouring points close to the wingtip, the feature is less sensitive to damage applied to the fuselage. Therefore, it can not be used for global damage assessment. However, the algorithm properly identifies damage in proximity of the area covered by the feature, i.e., data from 1701 to 2000, corresponding to the wingtip in the first configuration, and data from 3901 to 4300, corresponding to the wingtip in the second configuration. This characteristic suggests the possibility of exploiting a set of local transmissibilities to obtain further information on the damage conditions. A symmetric result can also be obtained by studying the novelty index corresponding to the transmissibilities between two points near the wingtip on the right side, as shown in Fig 5.

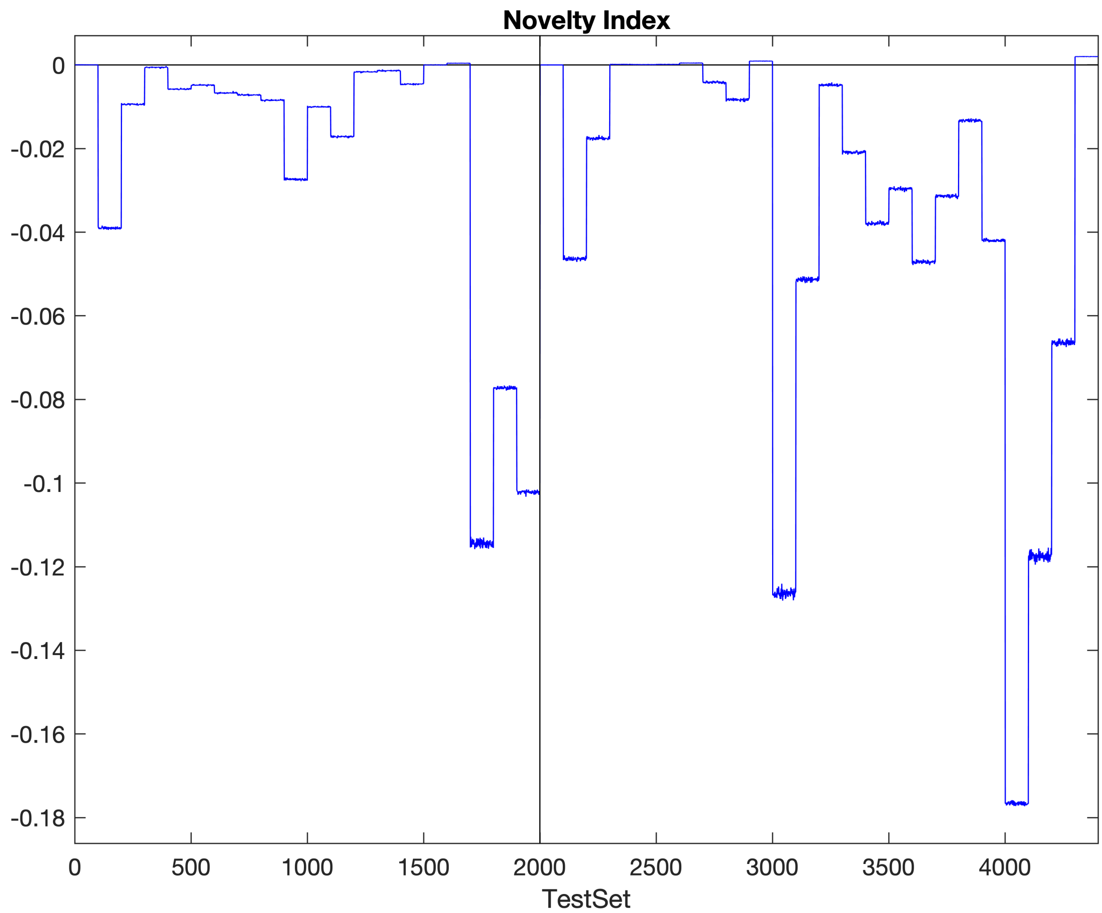

Figure 4. Novelty Index NI, for the monitored set M. The feature vectors regard the transmissibilities between two points near the wingtip on the left side. The first 2000 features correspond to the configuration with no winglets. The subsequent ones correspond to the configuration with winglets.

4. Discussion and conclusion

This study analyses the capabilities of the negative selection algorithm for identifying multiple damage conditions from two different normal configurations of the same structure. A real-value negative selection algorithm is proposed, which adopts vibration-based features, i.e., TFs, as proposed in [3]. This methodology takes into account the operational or environmental variations to which structures may be subjected.

Despite concerns about the applicability of this family of algorithms [6], NSA provides satisfactory damage-identification performance for the analysed features, and the high runtime issue is solved using autoencoders for dimensionality reduction. However, the algorithm critically depends on the choice of some parameters, such as the radii of the hyperspheres, in the self-set and the detectors set. Different approaches could be explored to derive them, and various distance measurements could be used, depending on the source features. In addition, the space-filling efficiency of detector hyperspheres, as pointed out by [9], is non-trivial to quantify. Thus, it is an issue that could be further addressed.

Given the features' dimensionality reduction, the NSA is extended to an experimental case study. The feature vectors are extracted from the experimental modal analysis of a laboratory-scale aeroplane model in two different configurations. These are considered two initial investigations to

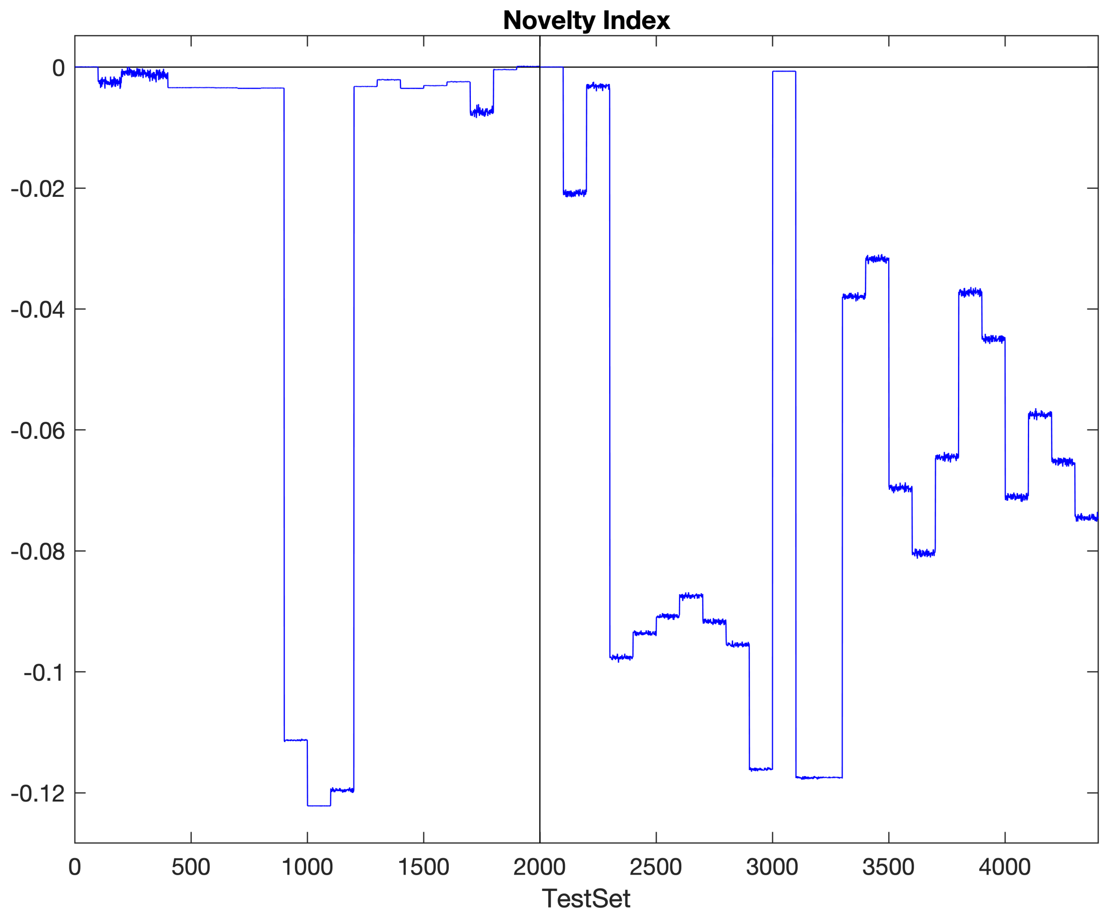

Figure 5. Novelty Index NI, for the monitored set M. The feature vectors regard the transmissibilities between two points near the wingtip on the right side. The first 2000 features correspond to the configuration with no winglets. The subsequent ones correspond to the configuration with winglets.

extend the study to similar structures and additional operational conditions. It is observed that the transmissibility function between the wingtips enables clustering of all damage conditions. Furthermore, by using a more local feature (e.g., a feature vector measured between two neighbouring points on the wing), it can be appreciated how, although the NSA loses the ability to identify all damage conditions, it can detect the anomalies located close to the measured points. Therefore, the choice of features remains of primary concern, but combining different feature vectors could help in retrieving further information about the anomaly's location.

Acknowledgements
The authors of this paper gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) via grant references EP/W005816/1 and EP/S001565/1.

References

[1] Ji Z and Dasgupta D 2007 *Evolutionary Computation* **15** 223–251 ISSN 1063-6560, 1530-9304 URL `https://direct.mit.edu/evco/article/15/2/223-251/1265`

[2] Dong Y, Sun Z and Jia H 2006 *Mechanical Systems and Signal Processing* **20** 1461–1472 ISSN 08883270 URL `https://linkinghub.elsevier.com/retrieve/pii/S0888327005000105`

[3] Surace C and Worden K 2010 *Mechanical Systems and Signal Processing* **24** 1114–1128 ISSN 08883270 URL `https://linkinghub.elsevier.com/retrieve/pii/S0888327009003069`

[4] Worden K, Manson G and Allman D 2003 *Journal of Sound and Vibration* **259** 323–343 ISSN 0022460X URL `https://linkinghub.elsevier.com/retrieve/pii/S0022460X02951680`

[5] Stibor T, Mohr P, Timmis J and Eckert C 2005 Is negative selection appropriate for anomaly detection? *Proceedings of the 7th annual conference on Genetic and evolutionary computation* (Washington DC USA: ACM) pp 321–328 ISBN 978-1-59593-010-1 URL `https://dl.acm.org/doi/10.1145/1068009.1068061`

[6] Stibor T and Timmis J 2007 Comments on real-valued negative selection vs. real-valued positive selection and one-class SVM *2007 IEEE Congress on Evolutionary Computation* (Singapore: IEEE) pp 3727–3734 ISBN 978-1-4244-1339-3 978-1-4244-1340-9 URL `http://ieeexplore.ieee.org/document/4424956/`

[7] Freitas A A and Timmis J 2007 *IEEE Transactions on Evolutionary Computation* **11** 521–540

[8] Forrest S, Perelson A, Allen L and Cherukuri R 1994 Self-nonself discrimination in a computer *Proceedings of 1994 IEEE Computer Society Symposium on Research in Security and Privacy* pp 202–212

[9] Ji Z and Dasgupta D 2005 Estimating the detector coverage in a negative selection algorithm *Proceedings of the 7th annual conference on Genetic and evolutionary computation* pp 281–288

[10] Ji Z and Dasgupta D 2004 Real-valued negative selection algorithm with variable-sized detectors *Proceedings of the 6th annual conference on Genetic and evolutionary computation* (Springer) pp 287–298

[11] Link M and Friswell M 2003 *Mechanical Systems and Signal Processing* **17** 9–20 ISSN 08883270 URL `https://linkinghub.elsevier.com/retrieve/pii/S0888327002915347`

[12] Friswell M, Penny J and Garvey S 1998 *Computers & Structures* **69** 547–556 ISSN 0045-7949 URL `https://www.sciencedirect.com/science/article/pii/S0045794998001254`

[13] Behmanesh I and Moaveni B 2015 *Structural Control and Health Monitoring* **22** 463–483

[14] Papatheou E, Manson G, Barthorpe R J and Worden K 2010 *Journal of Sound and Vibration* **329** 2349–2366