Unsupervised damage clustering in complex aeronautical composite structures monitored by Lamb waves: An inductive approach

Amirhossein Rahbari [a], Marc Rébillat [a,*], Nazih Mechbal [a], Stephane Canu [b]

[a] PIMM laboratory, ENSAM-CNRS-CNAM-HESAM, Paris, France
[b] LITIS laboratory, INSA Rouen, Rouen, France

ARTICLE INFO

ABSTRACT

Structural Health Monitoring (SHM), i.e. the action of monitoring structures in real-time and in an automated manner, is a major challenge in several industrial fields such as aeronautic. SHM is by nature a very high dimensional data-driven problem that possesses several specificities when addressed as a machine learning problem. First of all data in damaged cases are rare and very costly as the generation of damaged data is not always possible and simulations are not reliable especially when dealing with complex structures. SHM is thus by nature an unsupervised problem. Furthermore, any incoming sample should be instantaneously clustered and handcrafted damage indexes are commonly used as a first dimension reduction step due to large datasets to be processed. As a consequence, unsupervised dimensionality reduction (DR) techniques that project very high dimensional data into a two or three-dimensional space (such as t-SNE or UMAP) are very appealing in such a context. However, these methods suffer from one major drawback which is that they are unable to cluster any unknown incoming sample. To solve this we propose to add inductive abilities to these well know methods by associating their projection bases with Deep Neural Networks (DNNs). The resulting DNNs are then able to cluster any incoming unknown samples. Based on those tools, a SHM methodology allowing for unsupervised damage clustering with dimensionality reduction is presented here. To demonstrate the effectiveness of the method, results of damage classification on large experimental data sets coming from complex aeronautical composite structures monitored through Lamb waves are shown. Furthermore, several DR techniques have been benchmarked and recommendations are derived. It is demonstrated that the use of raw Lamb wave signals instead of the associated damage indexes is more effective. This non-intuitive result helps to reduce the gap between laboratory research and the actual start-up of SHM activities in industrial applications.

1. Introduction

Nowadays, human life is highly dependent on safe mechanical structures and infrastructures related to industrial fields such as aeronautics/aerospace engineering, civil engineering, or energy. Thus there is a need to develop innovative damage monitoring techniques and preventive maintenance methods allowing for safer operation and longer service life (Farrar and Worden, 2012). The action of monitoring structures in real-time and in an automated manner is referred to as Structural Health Monitoring (SHM) and is one of the most active research topics in that field (Bhuiyan et al., 2015; Di Sante, 2015; Mitra and Gopalakrishnan, 2016; Papaelias et al., 2016). More precisely SHM refers to damage monitoring processes relying on automated and periodical structural response measurements during the service life of the structure by permanently incorporating sensors and actuators into the structure. The SHM process classically includes the following steps:

damage detection, localization, classification and quantification (Worden et al., 2007). This scientific field is now progressing toward operational service and recent surveys have shown that even reluctant industry areas are now convinced that SHM is the key technology that will enable the transition from schedule-driven maintenance to condition-based maintenance.

More precisely, the considered structures under study are geometrically complex aeronautics structures made up of composite materials and used in airplanes nacelles. Damage monitoring of such structures can be achieved by Lamb waves (Su et al., 2006; Raghavan and Cesnik, 2007; Su and Ye, 2009) excited by piezoelectric elements bonded to the surface of the composite structure under study and used both as sensors and actuators. The piezoelectric elements used as actuators generate a propagating Lamb wave within the inspected structure and these waves are then received by the other piezoelectric elements that work in sensor mode. By comparing the signals in the current state with signals recorded in an undamaged reference state, it is possible

---

to extract echos coming from damages appearing within the structure and thus to monitor the structure (Fendzi et al., 2016a,b; Lizé, 2018).

SHM of aeronautic composite structures can naturally be cast in a statistical pattern recognition framework. Data acquisition by sensors placed along the structure under study is the merely sufficient requirements to build an underlying model of the monitored structure. Technically speaking, SHM is achieved through four successive steps: operational evaluation, data acquisition, damage indexes extraction, and underlying model learning (Farrar and Worden, 2012). While the operational evaluation and data acquisition steps are nowadays standards, the damage indexes extraction step is more delicate. Damage indexes extraction can be performed in the time domain (root mean square, skewness, peak to peak amplitude, residual energy...), in the frequency domain (maximum frequency, energy of FFT, maximum of the spectrum...), or in the time–frequency domain (wavelet coefficients, ...) (Staszewski and Robertson, 2006; Bouzenad et al., 2019; El Mountassir et al., 2016). A tentative list of some effective damage indexes that can be extracted from raw signals is for example provided in Lizé (2018). Manual tuning by highly experienced operators are the key to be able to design robust and efficient damage indexes. This constitutes a severe drawback of the data-based approach and emphasizes the need for efficient and robust features selection tools (Chandrashekar and Sahin, 2014; Cai et al., 2018). It should be added that in practice, data in damaged cases are rare and very costly and thus rarely or very scarcely available in a SHM context. Furthermore, as SHM has to be performed in real-time, any incoming new sample should be quasi-instantaneously clustered in order to help the end-user to infer a decision with respect to maintenance actions. These constitute the specificities of the SHM problem for aeronautic composite structures when addressed through a data-based approach. Fig. 1 provides a general overview of this classical point of view and highlights the importance of damage indexes selection in the data-based approach to SHM.

SHM of aeronautic composite structures has already been addressed in the literature through such a data-based approach depicted in Fig. 1. The damage localization and quantification steps of SHM have for example been carried out by decomposing the Lamb wave signals using a Proper Generalized Decomposition (PGD) algorithm which is then used to train neural networks in a supervised machine learning framework (Borate et al., 2020). Features automatically extracted by means of genetic algorithms and able to handle Lamb waves for SHM problems have been proposed (Harvey and Todd, 2014). Lamb wave data has also been processed using dynamical wavelet fingerprints and has been demonstrated that features extracted from wavelet analysis of Lamb wave signals can be extremely relevant for SHM purposes (Miller and Hinders, 2014; Hinders and Miller, 2020). Damage detection by means of Lamb waves by selecting the most relevant and discriminant features with wireless applications in mind has also been proposed (Park et al., 2010). Signal post-processing techniques focusing on the extraction of features, on the application of unsupervised dimension reduction algorithms (principal component analysis — PCA, and nonlinear PCA — NL-PCA) and on their subsequent correlation with the damage state are also reported (Pavlopoulou et al., 2016). In Pozo et al. (2016), the use of principal component analysis projections instead of the entire measured response of the structure has been investigated. The article (Tibaduiza et al., 2013) is concerned with the practical use of Multiway Principal Component Analysis (MPCA), Discrete Wavelet Transform (DWT), Squared Prediction Error (SPE) measures and Self-Organizing Maps (SOM) to detect and classify damages in composite structures. Additionally, the work carried out in Tibaduiza-Burgos and Torres-Arredondo (2015) highlights the potentialities of data-driven modeling within the concepts of sensor data fusion, feature extraction and pattern recognition for Lamb wave signals. The methodology is furthermore experimentally demonstrated on an aircraft skin panel and fuselage panel for which several damage scenarios are analyzed. By the same team, another article (Torres-Arredondo et al., 2014) is concerned with the experimental validation of a structural health monitoring methodology for damage detection and identification. Three different data-driven multivariate algorithms are considered to obtain the baseline pattern. These are based on PCA, ICA and hierarchical NL-PCA.

When addressing SHM as a machine learning problem, dimension reduction techniques in addition to damage indexes extraction methods thus appears as very appealing as they could efficiently reduce run time and resource requirements. Using such approaches, the data coming from the extracted damage indexes for a given structural health state can be projected toward a lower dimensional space (ideally containing two or three dimensions for evident visualization purposes) where monitoring can be performed (Gharibnezhad et al., 2015; Torgerson, 1952; Hinton and Roweis, 2003). In recent years, various dimension reduction algorithms of this kind have been proposed such as Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018) or t-Distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten and Hinton, 2008) and its derivatives (Van Der Maaten, 2014; Pezzotti et al., 2016). Such dimension reduction methods can turn any non-linear clustering or classification problem with arbitrarily high dimensional samples into a simple two or three dimension clustering problem. Contrary to more standard dimension reduction algorithms previously used in the SHM literature (Borate et al., 2020; Park et al., 2010; Pavlopoulou et al., 2016; Pozo et al., 2016; Tibaduiza et al., 2013; Tibaduiza-Burgos and Torres-Arredondo, 2015; Torres-Arredondo et al., 2014), these algorithms rely on a probabilistic definition of distance between samples and try to preserve as much as possible short-scale distances rather than long-scale ones in both high and low dimensional space. The unsupervised abilities of t-SNE and UMAP allow them to discriminate autonomously various unseen damaged states among the observed states. However these techniques suffers from drawbacks, including slow run time on big data-set, hard to tune parameters that yield completely different embedding maps and inability to map any incoming and unknown sample (an issue known as the out-sample problem). To rectify this for this SHM problem we propose to add inductive abilities to these well know methods by associating their projection bases with Deep Neural Networks (DNNs). The resulting DNNs are then able to cluster any incoming unknown samples. Such an approach has already been proposed for other kind of applications (Roman-Rangel and Marchand-Maillet, 2019; Espadoto et al., 2019) but never assessed for SHM applications which are rather specific as stated previously.

In this article, the high potential impact of t-SNE and UMAP extended with DNNs for SHM purposes is thus explored. More precisely, the motivations of the current study presented here are:

- *Dimension reduction (DR) algorithms benchmarking on SHM data*: many applicable dimension reduction technique from PCA to state of the art algorithms (t-SNE, UMAP) are benchmarked on an unsupervised damage clustering task in complex aeronautical composite structures monitored by Lamb waves in order to highlight the potential of t-SNE and UMAP for SHM. A solution to efficiently tune the DR algorithms parameters is also proposed.

- *Usefulness of damage indexes*: Dimension reduction techniques are tested in two ways in order to assess whether or not the computation of damage indexes is useful for clustering purposes.

  1. Approach I: the unsupervised DR technique is used in order to project the data from the handcrafted damage indexes space toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps, as shown in Fig. 2,
  2. Approach II: the unsupervised DR technique is used as both a damage indexes extractor and a dimension reduction tool fed directly by raw signals that finally projects data toward an $\mathbb{R}^2$ or $\mathbb{R}^3$ map as shown in Fig. 3.

- *DNNs extension of t-SNE and UMAP to learn projection and solve the out-sample issue*: A learning approach based on DNNs is adapted from Roman-Rangel and Marchand-Maillet (2019), Espadoto et al.

Fig. 1. Overview of the classical data-based approach for SHM.

(2019) to catch the hidden behavior of t-SNE and UMAP projections and to address the out-sample problem. The underlying neural networks are trained on a small subset of data provided by the DR techniques to learn the projection from high dimensional space to low dimensional space and are then used to process any incoming new sample.

The paper is organized as follows: Section 2 introduces the benchmarked DR algorithms and associated quantitative criterions. Then in Section 3, the simulated and experimental data sets coming from complex aeronautical composite structures monitored by Lamb waves which are used in this article are introduced. The benchmark results on the experimental and simulated data sets results and the comparison of approaches I and II are then presented and discussed in Section 4. The DNNs extension of UMAP and t-SNE projections designed to circumvent the out-sampling problem is then assessed in Section 5. Finally a discussion and a conclusion are presented in Section 6. Generally speaking, this paper explores the use of powerful clustering methods to SHM and is hence participating in the ongoing industrialization of the SHM process.

2. Methodology

2.1. Dimension reduction algorithms used for benchmarking

Unsupervised algorithms are of main interest for SHM applications as stated previously. However supervised algorithms are useful to provide references in terms of performances for benchmarking purposes. As a consequence, the following DR algorithms will be considered:

- Principal Component Analysis (PCA) (Pearson, 1901) as one of the most common techniques applied in a wide range of disciplines, and especially in SHM (Tibaduiza et al., 2013; Manson et al., 2003)
- Multi-dimensional scaling (MDS) (Torgerson, 1952) as adequate to process data sets with a large number of features
- ISOMAP (Tenenbaum, 1998; Tenenbaum et al., 2000) and Locally Linear Embedding (LLE) (Roweis and Saul, 2000) as powerful non-linear dimension reduction techniques.
- t-distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten and Hinton, 2008) and Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018) as the last state of art dimensional reduction methods.
- Linear discriminant analysis (LDA) (Tharwat et al., 2017) method, belonging to supervised dimensionality reduction techniques, as a reference for benchmarking previous unsupervised techniques.

In what follows, clustering performances of the above-mentioned unsupervised algorithms (PCA, MDS, ISOMAP, LLE, t-SNE and UMAP) will be compared with the performances of one supervised algorithm (namely LDA) that will serve as a reference. The expected advantages of UMAP and t-SNE over the other algorithms include no restriction in the number of samples, limitless embedding dimension, unlimited ambient dimension, robust performance, coverage of a wide range of manifolds. With respect to SHM applications, the main drawback of these algorithms is their inability to map any incoming and unknown sample (issue known as the out-sample problem) as what they learn

is a low dimensional map that has similar probabilistic properties than the high dimensional map but not the transformation allowing to project one new sample belonging to the high dimensional space to its projection living in the low dimensional one. Furthermore, the tuning of perplexity for t-SNE and of the number of considered neighbors for UMAP is extremely important. It is thus necessary to find an adequate value for these parameters that make a trade-off between the local and global features of the expected manifolds. Finally, the distance metric chosen in order to define the neighborhood of a point is another important tuning parameter.

2.2. Quantitative criterions to compare clustering algorithms performances

Quantitative criterions are furthermore needed to evaluate the clustering performances provided by the different algorithms (or by the different parameters selected by the users for a given algorithm) in order to perform a fair comparison between algorithms or among parameters. The standard Calinski–Harabasz (Caliński and Harabasz, 1974), the Silhouette scores (Rousseeuw, 1987) and the computation time are here retained here as key performance indexes.

The Calinski–Harabasz (CH) score is a variance ratio criterion defined as *"the ratio of the between-cluster dispersion and the within-cluster dispersion"*. It is computed according to Eq. (1) where B refers to between-cluster variance matrix, W refers to the within-cluster variance matrix, N is the total number of samples and K the number of clusters. A high value of the CH score illustrates dense and well-segregated clusters.

$$CH = \frac{tr(B)}{tr(W)} \times \frac{N - K}{K - 1} \tag{1}$$

The Silhouette score is the mean of all the silhouette coefficients associated with each point i. It corresponds to the difference between the mean distance with points belonging to the same cluster ($a(i)$, cohesion) and the mean distance with points from neighboring clusters ($b(i)$, separation). A negative score indicates that the point is on average closer to a neighboring cluster than from its own cluster. A positive score indicates that the point is i = on average closer to its own cluster than from any other neighboring clusters. Its value is normalized between -1 and 1. This score is mathematically computed according to Eq. (2).

$$S = \frac{1}{N} \sum_{i=1}^{N} \frac{b(i) - a(i)}{max\{b(i), a(i)\}} \tag{2}$$

Computation time is the last key performance index retained for dimension reduction algorithms comparison. In order to assess it, the tested dimension reduction algorithms are implemented and available in `sklearn`. The original version of UMAP implemented in Python has been used. Neural network provided through Keras that is one of the powerful distributions of TensorFlow has been used. Computations have been run on a computing station equipped with 128 Go of random access memory and with 2 Intel Xeon ES 2630 processors running at 2.6 GHz, each with 6 cores and with 15 MB of cache memory.

Fig. 2. Overview of approach I: the unsupervised dimensionality reduction (DR) technique is used in order to project the data from the handcrafted damage indexes space toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps.
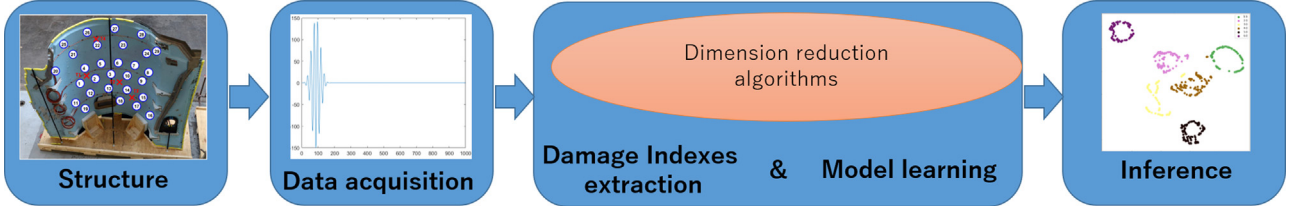

Fig. 3. Overview of approach II: the unsupervised DR technique is used as a damage indexes extractor and a dimension reduction tool fed by raw signals that projects data toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps.

3. Aeronautical composite structures monitored by lamb waves

This section describes the SHM numerical and experimental data sets that will be used here as benchmarks to assess the ability of dimension reduction algorithms to cluster data in an unsupervised manner, as a way to demonstrate that damage indexes computation is not necessary for clustering, and finally to illustrate the deep neural networks extension of t-SNE and UMAP based generic SHM process presented here.

3.1. Structural health monitoring of composite structures using lamb waves

3.1.1. Lamb wave generation and reception using piezoelectric elements

Lamb waves are one of the most effective approaches to monitor autonomously aeronautic composite structures (Su et al., 2006; Raghavan and Cesnik, 2007; Su and Ye, 2009). To achieve that goal, piezoelectric elements are bonded to the surface of the composite structure under study and can be used both as sensors and actuators. The piezoelectric elements used as actuators generate a propagating wave within the inspected structure and these waves are then received by the other piezoelectric elements that work in sensor mode. By comparing the signal in the current state with the signal recorded in an undamaged reference state, it is possible to extract echos coming from damages appearing within the structure and thus to monitor autonomously the structure (Fendzi et al., 2016a,b; Lizé, 2018). The main practical steps involved in Lamb waves SHM are described in the following sections.

3.1.2. Data acquisition framework

The excitation signal sent to the piezoelectric elements considered as actuators is a 5 cycles burst with an excitation frequency f_0 which depends on the considered database and an amplitude of 10 V. The excitation frequency is selected to promote one propagation mode over another. Typically, the mode S_0 is promoted over the mode A_0 as it propagates faster in the considered structures (Su and Ye, 2009; Fendzi et al., 2016a). In each phase of the experimental procedure, one piezoelectric element is selected as the actuator and the other act as sensors. All the piezoelectric elements act sequentially as actuators. Resulting signals are then simultaneously recorded by the other piezoelectric elements and consist of a certain number of data points sampled at a given sampling frequency. Signals are acquired several times in both the healthy (reference) and damaged (unknown) states.

3.1.3. Denoising procedure and X_{RAW} dataset

In order to be processed, the raw data that have been previously collected are first denoised (Fendzi et al., 2016a) by means of a discrete wavelet transform up to the order 4 using the db40 wavelet. Those signals are then filtered around their excitation frequency f_0 using a continuous wavelet transformation based on Morlet/Gabor wavelets and with a scale resolution equal to 20. The objective of this pre-processing step is to perform a bandpass filtering around the excitation frequency f_0 by means of wavelets. The scale parameter can be sought as an image of the bandwidth of the retained bandpass filter over the frequency range of interest. Here, choosing it equal to 20 is common as it provided convenient results in past studies (Su and Ye, 2009; Fendzi et al., 2016a,b). In those studies, an optimization algorithm based on the Shannon entropy was applied to select the best wavelet family as well as the best scale parameters to maximize the information contained within the wavelet decomposition. The diaphonic part present in the measured signals (i.e. the copy of the input signals that appears on the measured signal due to electromagnetic couplings in wires) has been previously eliminated based on the knowledge of the geometrical positions of the piezoelectric elements and of the estimated waves propagation speed in the material.

The concatenation of all the signals $s_{ij}(r(k))[n]$ for the time sample $n \in [1 : N]$, the repetition $r(k) \in [1 : R_U(k)]$ of the kth $\in [1 : K]$ unknown case or $r(k) \in [1 : R_R]$ of the reference case, the path $i \rightarrow j$, constitute the datasets X_{RAW} to be clustered by the dimension reduction algorithms. An example of such a dataset is presented in Eq. (3). Each line corresponds to one sample to be clustered and the columns to the associated features. There are R_R signals for the reference case (first horizontal block of Eq. (3)) and $R_U(k)$ signals for the kth unknown case (other horizontal blocks of Eq. (3) in Box I).

3.1.4. Damage indexes computation and X_{DI} dataset

The underlying idea of this step consists in comparing a set of signals acquired at an unknown state with a one acquired at a reference state (Fendzi et al., 2016a) by computing damage sensitive features denoted as Damage Indexes (DIs) in order to perform a first dimension reduction step. A set of $L = 26$ DIs are computed. The 26 damage indexes $\{DI^l\}_{l \in [1:L]}$ being computed are defined in Table 2 in Appendix A. The damage indexes DI^l_{ij} are computed for each path between a piezoelectric actuator $i \in [1 : I]$ to a piezoelectric sensor $j \in [1 : J]$. The concatenation of all the damage index $DI^l_{ij}(r_R, r_U(k))$ for the repetition $r_R \in [1 : R_R]$ of the reference case, the repetition $r(k) \in [1 : R_U(k)]$ of the kth $\in [1 : K]$ unknown case or $r(k) \in [1 : R_R]$

$$X_{\mathrm{RAW}} = \begin{bmatrix} s_{11}(1)[1] & \dots & s_{11}(1)[N] & \dots\dots & s_{IJ}(1)[1] & \dots & s_{IJ}(1)[N] \\ \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots \\ s_{11}(R_R))[1] & \dots & s_{11}(R_R)[N] & \dots\dots & s_{IJ}(R_R)[1] & \dots & s_{IJ}(R_R)[N] \\ \hline s_{11}(1)[1] & \dots & s_{11}(1)[N] & \dots\dots & s_{IJ}(1)[1] & \dots & s_{IJ}(1)[N] \\ \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots \\ s_{11}(R_U(1))[1] & \dots & s_{11}(R_U(1))[N] & \dots\dots & s_{IJ}(R_U(1))[1] & \dots & s_{IJ}(R_U(1))[N] \\ \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots \\ \hline s_{11}(1)[1] & \dots & s_{11}(1)[N] & \dots\dots & s_{IJ}(1)[1] & \dots & s_{IJ}(1)[N] \\ \dots & \dots & \dots & \dots\dots & \dots & \dots & \dots \\ s_{11}(R_U(K))[1] & \dots & s_{11}(R_U(K))[N] & \dots\dots & s_{IJ}(R_U(K))[1] & \dots & s_{IJ}(R_U(K))[N] \end{bmatrix} \quad (3)$$

Box I.

of the reference case, the path $i \rightarrow j$, constitute the datasets X_{DI} to be clustered by the dimension reduction algorithms. An example of such a dataset is presented in Eq. (4). Each line corresponds to one sample to be clustered and the columns to the associated features. There are $R_R(R_R-1)/2$ difference signals for the reference case compared to itself (first horizontal block of Eq. (4)) and $R_R \times R_U(k)$ difference signals for the kth unknown case compared to the reference case (other horizontal blocks of Eq. (4) given in Box II). As a consequence, the datasets X_{DI} naturally contain more samples than the datasets X_{RAW}.

3.2. Available databases

The databases used for this study consist of one simulated database and two experimental databases. The numerical database is considered as an academic dataset on which dimension reduction algorithms can be benchmarked with respect to the noise level. The experimental datasets correspond to actual geometrically complex aeronautic structures and are considered as challenging datasets for SHM applications.

3.2.1. Simulated database

For the simulated database, a $[0°/45°/23°/0°]$ composite laminate where the mechanical properties of each ply described in Table 1 is considered. A set of 5 piezoelectric elements (Noliac NCE51), each with a diameter of 20 mm and a thickness of 0.1 mm, are surface-mounted on the composite plate. An illustration of the plate and sensor placement is shown in Fig. 4. Numerical simulations are conducted using SDTools (Balmes, 2020). Squared elements with dimension 2 mm × 2 mm were used for the meshing. The time step for the transient simulation is chosen as 0.3 ms and leads to a sampling frequency of 3.33 MHz. The damage has a circular shape with a 20 mm diameter. The damage is represented by a local reduction in material properties in the damaged area. The rigidity is reduced step by step from the healthy state toward a totally damaged state resulting in one healthy reference case and 18 damaged cases. It is worth mentioning that this FEM model was previously validated through experiments (Fendzi et al., 2016a). After the simulation, a white Gaussian noise is added to the simulation results for each path between a given actuator and a given sensor in order to simulate experimental noise. Several realization of this noise constitute an equivalence to the experimental repetitions. A central frequency of $f_0 = 120$ kHz is used with SNR values equal to 0 dB, 10 dB and 20 dB dB. Here SNR stands for Signal to Noise Ratio and the value 0 dB refers to maximum amount of noise pollution (the energy of the noise being equal to the energy of the signal). The X_{RAW} database contains 2850 samples, each one associated with 25 000 features and divided into 19 different classes. The X_{DI} database contains 416 175 samples, each one associated with 650 features and divided in 19 different classes. The databases that have been generated are detailed in Table 3 in Appendix B and will be referred to as SNR0, SNR10 and SNR20 in the following of the paper.

Table 1
Mechanical properties of one ply of the chosen composite material.

Density (g/m³)	Thickness (mm)	E_{11} (GPa)	E_{22}	E_{33} (GPa)	G_{12} (GPa)	v_{12}
1554	0.28	60	40	8.1	4.8	0.03

3.2.2. Experimental database #1 : Fan Cowl (FC) structure

The geometrically complex aeronautics structure under study consists here in the fan cowl part of a nacelle of an Airbus A380. This structure is 1.5 m in height for a semi circumference of 4 m and is made of composite monolithic carbon epoxy material. It has been equipped with 30 piezoelectric elements manufactured by NOLIAC (diameter of 25 mm) and possesses many stiffeners delimiting various areas as shown in Fig. 5 *[Left]*. The excitation signal sent to the PZT element is a 5 cycles burst with an excitation frequency of $f_0 = 200$ kHz and with an amplitude of 10 V. The excitation frequency is selected to promote the mode S_0 over the mode A_0 as it propagates faster. The Lamb wave propagation speed within the material is estimated at around 5200 m/s for the S_0 mode. In each phase of the experimental procedure, one PZT is selected as the actuator and the other act as sensors. All the PZTs act sequentially as actuators. Resulting signals are then simultaneously recorded by the other piezoelectric elements and consist of 1000 data points sampled at 1 MHz. Signals were acquired 100 times in both the healthy (reference) and damaged (unknown) states. One healthy case and 7 damage cases have been considered. The X_{RAW} database contains 800 samples, each one associated with 900 000 features and divided into 8 different classes. The X_{DI} database contains 123 675 samples, each one associated with 23 400 features and divided into 8 different classes. The database that has been generated is detailed in Table 4 in Appendix B and will be referred to as FC in the following of the paper.

3.2.3. Experimental database #2: Inner fixed structure

The other geometrically complex aeronautics structure under study is the inner fixed structure part of a nacelle of an Airbus A380. This structure is 2.5 m in height for a semi-circumference of 3 m and is made of aluminum hexagonal honeycombs sandwiched between two layers of composite monolithic carbon epoxy material. It has been equipped with 29 piezoelectric elements manufactured by NOLIAC (diameter of 25 mm) as shown in Fig. 5 *[Right]*. The excitation signal sent to the PZT element is a 5 cycles burst with an excitation frequency of $f_0 = 100$ kHz and with an amplitude of 10 V. The excitation frequency is selected to promote the mode S_0 over the mode A_0 as it propagates faster. The Lamb wave propagation speed within the material is estimated at around 3000 m/s for the S_0 mode. In each phase of the experimental procedure, one PZT is selected as the actuator and the other act as sensors. All the PZTs act sequentially as actuators. Resulting signals are then simultaneously recorded by the other piezoelectric elements and consist of 1000 data points sampled at 1 MHz. Signals were acquired 100 times in both the healthy (reference) and damaged (unknown)

$$X_{\mathrm{DI}} = \begin{bmatrix}
\mathrm{DI}_{11}^1(1,1) & \cdots & \mathrm{DI}_{IJ}^1(1,1) & \cdots\cdots & \mathrm{DI}_{11}^L(1,1) & \cdots & \mathrm{DI}_{IJ}^L(1,1) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(1,R_R) & \cdots & \mathrm{DI}_{IJ}^1(1,R_R) & \cdots\cdots & \mathrm{DI}_{11}^L(1,R_R) & \cdots & \mathrm{DI}_{IJ}^L(1,R_R) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(R_R-1,R_R) & \cdots & \mathrm{DI}_{IJ}^1(R_R-1,R_R) & \cdots\cdots & \mathrm{DI}_{11}^L(R_R-1,R_R) & \cdots & \mathrm{DI}_{IJ}^L(R_R-1,R_R) \\
\mathrm{DI}_{11}^1(1,1) & \cdots & \mathrm{DI}_{IJ}^1(1,1) & \cdots\cdots & \mathrm{DI}_{11}^L(1,1) & \cdots & \mathrm{DI}_{IJ}^L(1,1) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(1,R_U(1)) & \cdots & \mathrm{DI}_{IJ}^1(1,R_U(1)) & \cdots\cdots & \mathrm{DI}_{11}^L(1,R_U(1)) & \cdots & \mathrm{DI}_{IJ}^L(1,R_U(1)) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(R_R,R_U(1)) & \cdots & \mathrm{DI}_{IJ}^1(R_R,R_U(1)) & \cdots\cdots & \mathrm{DI}_{11}^L(R_R,R_U(1)) & \cdots & \mathrm{DI}_{IJ}^L(R_R,R_U(1)) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(1,1) & \cdots & \mathrm{DI}_{IJ}^1(1,1) & \cdots\cdots & \mathrm{DI}_{11}^L(1,1) & \cdots & \mathrm{DI}_{IJ}^L(1,1) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(1,R_U(K)) & \cdots & \mathrm{DI}_{IJ}^1(1,R_U(K)) & \cdots\cdots & \mathrm{DI}_{11}^L(1,R_U(K)) & \cdots & \mathrm{DI}_{IJ}^L(1,R_U(K)) \\
\cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots & \cdots \\
\mathrm{DI}_{11}^1(R_R,R_U(K)) & \cdots & \mathrm{DI}_{IJ}^1(R_R,R_U(K)) & \cdots\cdots & \mathrm{DI}_{11}^L(R_R,R_U(K)) & \cdots & \mathrm{DI}_{IJ}^L(R_R,R_U(K))
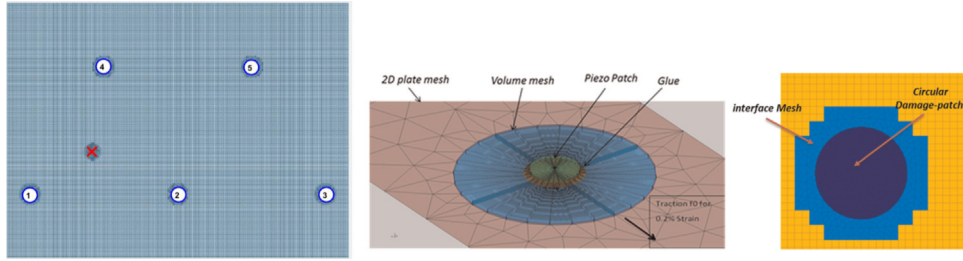\end{bmatrix} \quad (4)$$

Box II.

Fig. 4. *[Left]* Overview of the geometrical configuration for the simulated database. Mesh details around the PZT element *[Center]* and the damage *[Right]*.
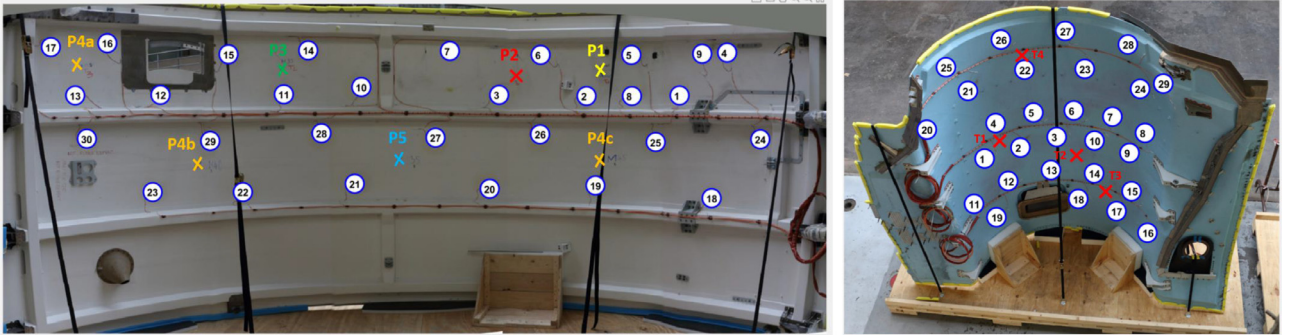

Fig. 5. Overview of the geometrical configuration of the Fan Cowl Structure (FC) *[Left]* and the Inner Fixed Structure (IFS) *[Right]* experimental datasets.

states. One healthy case and 5 damage cases have been considered. The X_{RAW} database contains 600 samples, each one associated with 841 000 features and divided in 6 different classes. The X_{DI} database contains 54 950 samples, each one associated with 21 866 features and divided in 6 different classes. The database that has been generated is detailed in Table 5 in Appendix B and will be referred to as IFS in the following of the paper.

4. Dimension reduction algorithms benchmarking for SHM

4.1. Benchmark approach

The first contribution presented here consists of benchmarking dimension reduction algorithms ranging from PCA to t-SNE and UMAP with respect to their clustering abilities as assessed by the CH, silhouette scores, and computation time (see Section 2). The unsupervised dimension reduction algorithms considered here are: PCA, MDS, ISOMAP, LDA, LLE, UMAP and t-SNE. Furthermore, LDA (a supervised dimension reduction technique) is used here to provide a reference value for comparison purposes.

Approach I where the unsupervised dimension reduction techniques are used in order to project the data from the handcrafted damage indexes space toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps, as shown in Fig. 2, is firstly investigated. The X_{DI} datasets computed from the numerical and experimental datasets described in Section 3 are considered. The previously mentioned dimension algorithms are all used for clustering purposes when fed by X_{DI}. They are first parametrized using a small subset of X_{DI} as described in the next section.

The second contribution addressed here is related to the usefulness of handcrafted damage indexes that are commonly used in SHM using

Lamb Waves due to the large amount of raw data to be processed. Processing directly such an amount of data was not achievable before but is now feasible using dimension reduction algorithms such as UMAP or t-SNE. In order to address this issue, Approach II where the unsupervised dimension techniques are used as both a damage indexes extractor and a dimension reduction tool fed directly by raw signals that finally project data toward a $\mathbb{R}^2$ or $\mathbb{R}^3$ map as shown in Fig. 3 is investigated. The X_{RAW} datasets computed from the numerical and experimental datasets described in Section 3 are considered. The previously mentioned dimension algorithms are all used for clustering purposes when fed by X_{RAW}. They are also parametrized using a small subset of X_{RAW} as described in the next section.

For both approaches, the influence of noise on clustering results provided by the various algorithms is assessed using the numerical datasets SNR0, SNR10, and SNR20. The ability of the various algorithms to process realistic experimental data is assessed on the two experimental datasets FC and IFS. The CH score (which is a variance to ratio score), the silhouette score (which is a combination of cohesion and separation), and the computation time are afterward used to conclude regarding dimension reduction algorithms performances and to compare them.

4.2. Parameters tuning

The previously mentioned dimension reduction algorithms need some hyper-parameters to be selected before applying dimension reduction. Adequate tuning is especially important for UMAP and t-SNE as different hyper-parameters selection may yield to completely different embedding dimension representation. For example, given the same input data and different tuning parameters for UMAP, the results provided in Fig. 6 illustrate a large overlap of clusters for some parameters selection while the clusters can be extremely well-distinguished if parameters are selected adequately (Manhattan distance metric and number of neighbors equals to 20 in this case).

In order to efficiently tune those parameters a subsampling based approach is used here based 10 % of available data picked randomly and uniformly among the various classes. This subset of data is then fed to UMAP and t-SNE projecting data in a 2D space with various parameters selection. For UMAP, the number of neighbors (from 5 to 70 with steps of 5) and the distance metric (Euclidean, Manhattan, Chebyshev, Cosine, or Correlation) are varied. For t-SNE, the perplexity (from 5 to 70 with steps of 5) and the distance metric (Euclidean, Manhattan, Chebyshev, Cosine or Correlation) are varied. On the basis of these computations, three graphs providing computation time, silhouette coefficient and CH score as a function of distance metrics and perplexity or number of neighbors are then generated. This figure then helps the user to select the adequate hyper-parameters.

An example of such a graph for the tuning of UMAP for the FC dataset and using approach II is shown in Fig. 7. From this figure, it can be seen that for this precise case, computation time is not depending on the number of neighbors but is influenced quite largely by the distance metric, with Euclidean and Cosine distance metrics being faster. The silhouette score is influenced by both hyperparameters. The Manhattan distance metric however provides silhouette scores that are much larger than any other distance metric. Furthermore, a local maximum of this score is obtained for a number of neighbor equals to 20. The same trends are observed on the CH score. For this example and on the basis of this analysis, the chosen hyperparameters for UMAP are a Manhattan distance metric with a number of neighbors equals to 20. The same analysis is performed for all the other datasets, for both approaches, and for the t-SNE algorithm.

An important issue is now to validate the fact that clustering results obtained on 10 % of the datasets with the selected parameters are also valid when considering the whole datasets. In order to illustrate this point, Fig. 8 presents three results obtained on the FC dataset, for approach II and the UMAP algorithm. On the left figure, results

obtained after tuning on 10 % of the dataset and in 2D are presented. On the center figure, results obtained on the whole dataset and in 2D are presented. On the right figure, results obtained on the whole dataset and in 3D are presented. This figure validates the fact that the result from sub-sampling can be reliably applied to the complete data set in 2D or 3D in the sense that UMAP still provides well-separated clusters for all cases. Similar results are obtained for all the datasets, algorithms, and approaches validating the hyperparameters selection methodology. The hyperparameters chosen for all the datasets, for t-SNE and UMAP algorithms, and for the approaches I and II are summarized in Table 6 located in Appendix C.

4.3. Approach I: dimension reduction projects damage indexes X_{DI} to $\mathbb{R}^2$ or $\mathbb{R}^3$

After parameter selection according to Section 4.2, a comparison between the different dimension reduction techniques previously mentioned (PCA, MDS, ISOMAP, LDA, LLE, t-SNE and UMAP) can be performed. For this benchmark, the dimension reduction techniques are applied to the simulated and experimental databases X_{DI} according to approach I where the unsupervised dimensionality reduction techniques are used to project the data from the handcrafted damage indexes space toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps as illustrated in Fig. 2.

For all cases at hand, ISOMAP failed in detecting the right manifolds and provides mixed clustered centered around the origin. MDS provides results that are similar to the ones provided by PCA when the number of samples is kept sufficiently low, but falls into infeasible singular value decomposition problem for a high number of samples and was consequently not able to reach a result. Like ISOMAP, LLE could not handle the right structure manifold and gives a few centered clusters distributed around the origin. For those reasons, ISOMAP, MDS and LLE algorithms are discarded and only the results provided by LDA, PCA, t-SNE and UMAP are displayed in Fig. 9.

Fig. 9 firstly shows that, as expected, LDA provides very adequate clustering results, with dense and well-separated clusters. One can notice that still the results provided by LDA are not perfect. For the SNR0 case, LDA is able to learn some sort of relation to segregate data but there is still a large overlap between clusters. For the FC case, several classes are well clustered whereas some still remain fully mixed. Given those partial failures of a supervised algorithm in those cases, one can hardly expect the unsupervised ones to perform well on those examples. PCA provides results that are acceptable in comparison with LDA for SNR0, SNR10, and SNR20. However, for the experimental cases, the results provided by PCA are not really exploitable as no clear structures pop out from the resulting low dimensional space. t-SNE shows an unstable behavior for SNR0 and SNR10 providing clusters mixed and centered around the origin. With respect to the SNR20 and IFS cases, adequate results are provided by t-SNE. UMAP always shows very appropriate results and segregates the clusters well, even for SNR0 and SNR10. For the IFS data-sets, it is interesting to notice that PCA and even LDA fail, but that t-SNE and UMAP work well and detect adequately the clusters. Finally for the FC case, all of the algorithms failed, including LDA. This might be due to badly chosen DIs that remove too much relevant information from the measured signals.

A quantitative analysis of the clustering results achieved by the various algorithms when fed with X_{DI} can then be achieved by analyzing computation time, silhouette score, and CH score. These scores are presented in Fig. 10 for the different algorithms, different datasets and different low-dimensional spaces ($\mathbb{R}^2$ or $\mathbb{R}^3$). By looking at computation times, it can be observed that on the simulated datasets (SNR0, SNR10 and SNR20), UMAP and t-SNE needs more resources than PCA or LDA. Interestingly, this gap in computation time between algorithms decreases by one or two orders of magnitudes when applied to the experimental datasets FC and IFS. It can be furthermore observed that the level of noise (SNR) has a large effect on silhouette score. Indeed, silhouette score increases for all algorithms with the SNR.
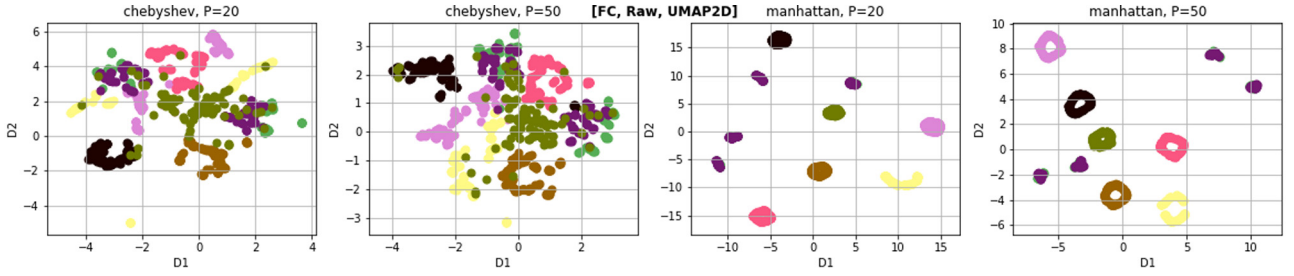
Fig. 6. Example of tuning results for UMAP applied on the FC dataset using the approach II in 2D. *[Left]* UMAP projections using Chebyshev metric with number of neighbors of 20 and 50, *[Right]* UMAP projections using the Manhattan metric with number of neighbors of 20 and 50. The optimal parameters retained here are Manhattan distance metric and number of neighbors equals to 20 in this case.
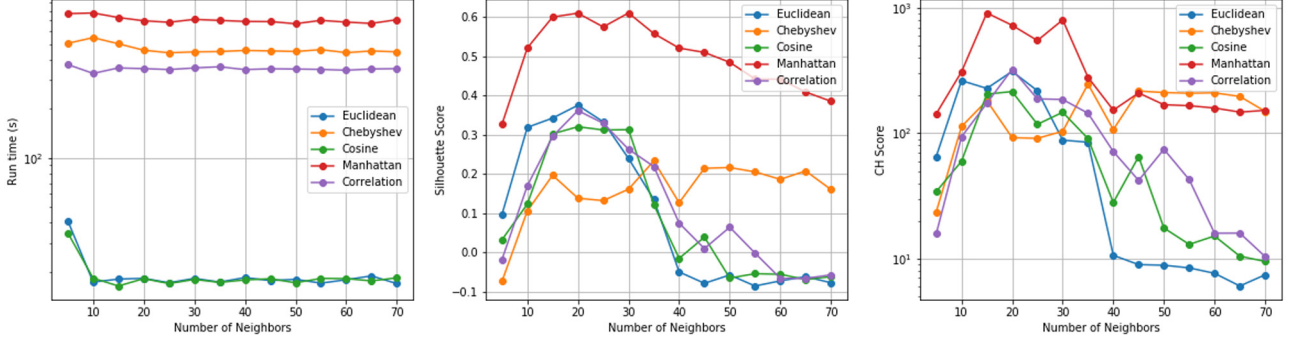

Fig. 7. Overview of visual graph for tuning UMAP parameters selection for the FC datasets using the approach II.
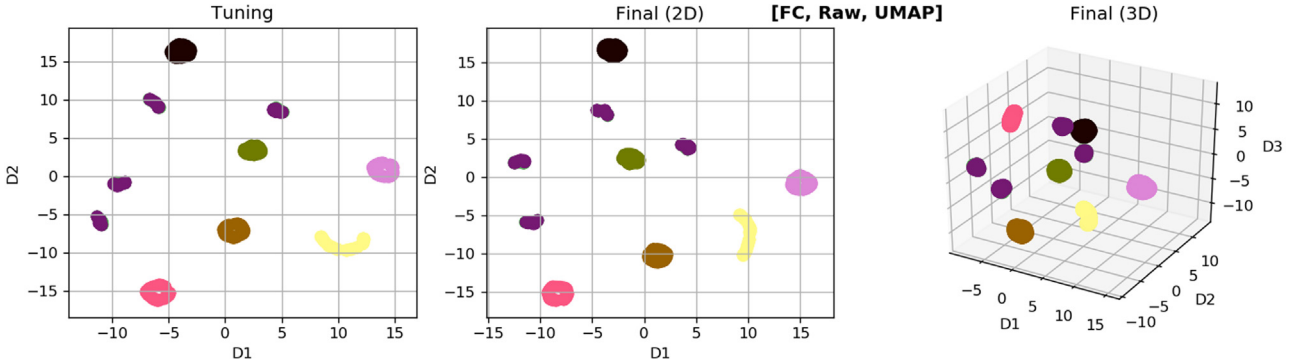

Fig. 8. Example of the extrapolation of tuning of UMAP applied to the FC database using the approach II: *[Left]* Tuning of UMAP on 10% of the dataset, *[Center]* UMAP in 2D on the complete dataset, and *[Right]* UMAP in 3D on the complete dataset.

This is in agreement with the fact that clustering noisy data is harder than clustering clean data. Among the benchmarked algorithms, UMAP generally provides a higher value for silhouette score and even equals LDA for the SNR10 dataset. Regarding CH score, the performances of UMAP and t-SNE are stable over the various datasets and are in general comparable to PCA performances and as expected lower than LDA performances. Finally, when comparing the computation time and score obtained for the $\mathbb{R}^2$ or $\mathbb{R}^3$ cases, no significant effect of the embedding dimension can be observed.

4.4. Approach II: dimension reduction projects raw signals X_{RAW} toward $\mathbb{R}^2$ or $\mathbb{R}^3$

After parameter selection according to Section 4.2, another comparison between the different dimension reduction techniques previously mentioned (PCA, MDS, ISOMAP, LDA, LLE, t-SNE and UMAP) can be performed. For this part, the dimension reduction techniques are applied to the simulated and experimental databases X_{RAW} according to approach II where the unsupervised DR technique is used as both a damage indexes extractor and a dimension reduction tool fed directly

by raw signals that finally projects data toward a $\mathbb{R}^2$ or $\mathbb{R}^3$ map as already explained in Fig. 3. Even if the number of samples is smaller, the number of features contained within the database X_{RAW} in comparison with the datasets X_{DI} is really larger and learning and clustering for this kind of database is highly challenging. One should also notice that approach II ignores damage indexes computation and thus saves plenty of time and resources in comparison with approach I.

For all cases at hand, ISOMAP and LLE again failed. MDS again provided results that are similar to the one provided by PCA when the number of samples is kept sufficiently low, but falls into infeasible singular value decomposition problem for high number of samples. For those reasons, ISOMAP, MDS and LLE algorithms are again discarded and only the results provided by LDA, PCA, t-SNE and UMAP are shown in Fig. 11.

Fig. 11 firstly shows that, again and as expected, LDA provides very adequate clustering results, with dense and well-separated clusters. One can notice that still the results provided by LDA are not perfect. For the SNR0 case, LDA is now able to learn a way much better than when relying on the damage indexes. Furthermore, the performances of LDA for the experimental IFS and FC cases severely
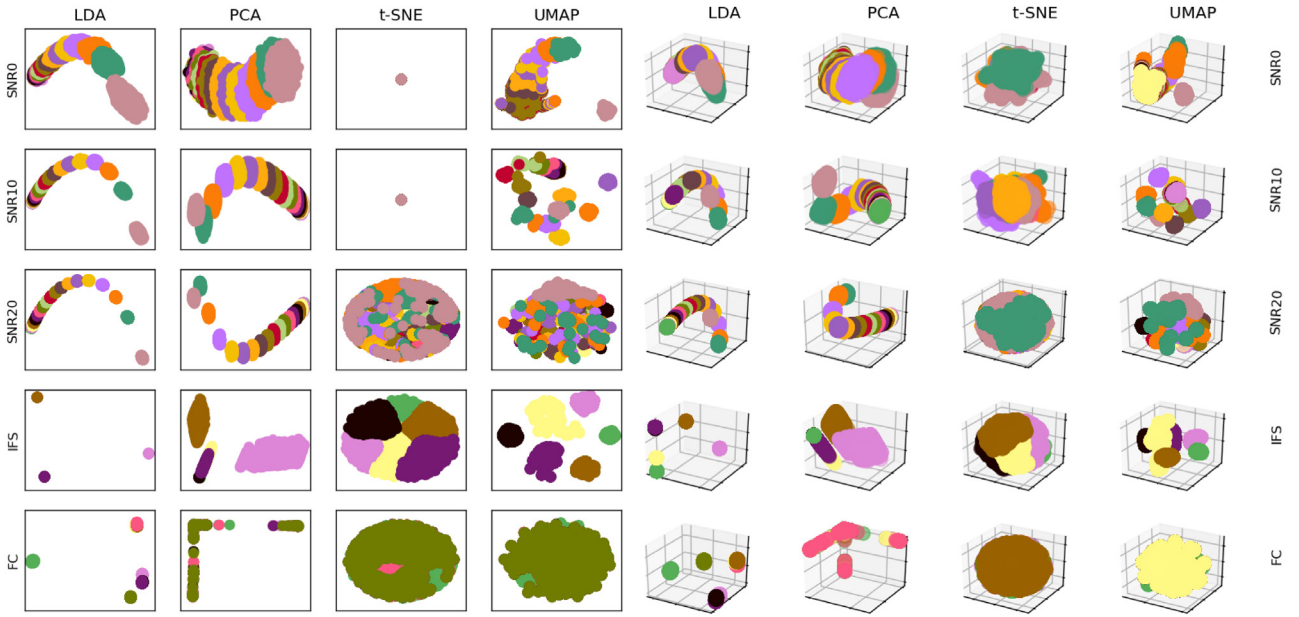
Fig. 9. Result of approach I where the unsupervised dimensionality reduction techniques are used to project the data from the handcrafted damage indexes space X_{DI} toward $\mathbb{R}^2$ *[Left]* or $\mathbb{R}^3$ *[Right]* maps (see Fig. 2) for all the considered datasets.
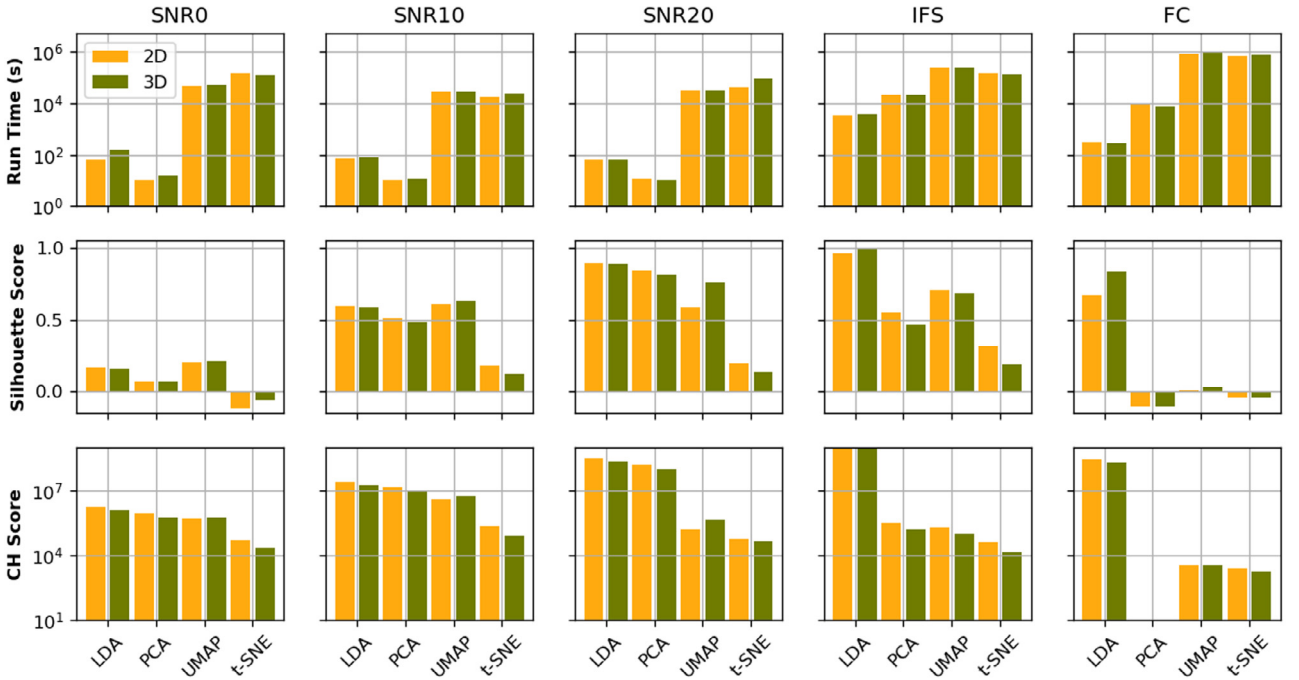

Fig. 10. Scores of approach I where the unsupervised dimensionality reduction techniques are used to project the data from the handcrafted damage indexes space X_{DI} toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps (see Fig. 2) for all the considered datasets.

degrades. Clusters are not well separated and are sometimes mixed. Given those partial failures of a supervised algorithm in those cases, one can again hardly expect the unsupervised ones to perform well on those examples. PCA provides results that are comparable to LDA for SNR0, SNR10 and SNR20. However, for the experimental cases, the results provided by PCA are not exploitable as no clear structures pop out from the resulting low dimensional space. t-SNE shows much better results than previously for SNR0 and SNR10 providing clusters similar to LDA and PCA. With respect to the SNR20, IFS and FC cases, extremely encouraging results are provided by t-SNE. UMAP always shows very appropriate result and segregates the clusters well whatever the considered case. For the IFS and FC data-sets, it is interesting to

notice that PCA and even LDA fails, but that t-SNE and UMAP work well and detect adequately the clusters. Finally for the FC case, t-SNE and UMAP previously failed but are now able to correctly cluster data and even overperform LDA. This might be due to the inability of LDA to manage a too large number of features and to the intrinsic ability of t-SNE and UMAP to detect hidden relevant dimensions within the datasets.

A quantitative analysis of the clustering results achieved by the various algorithms when fed with X_{RAW} can then be achieved by analyzing computation time, silhouette score, and CH score. These scores are presented in Fig. 12 for the different algorithms, different datasets and different low-dimensional spaces ($\mathbb{R}^2$ or $\mathbb{R}^3$). By looking at
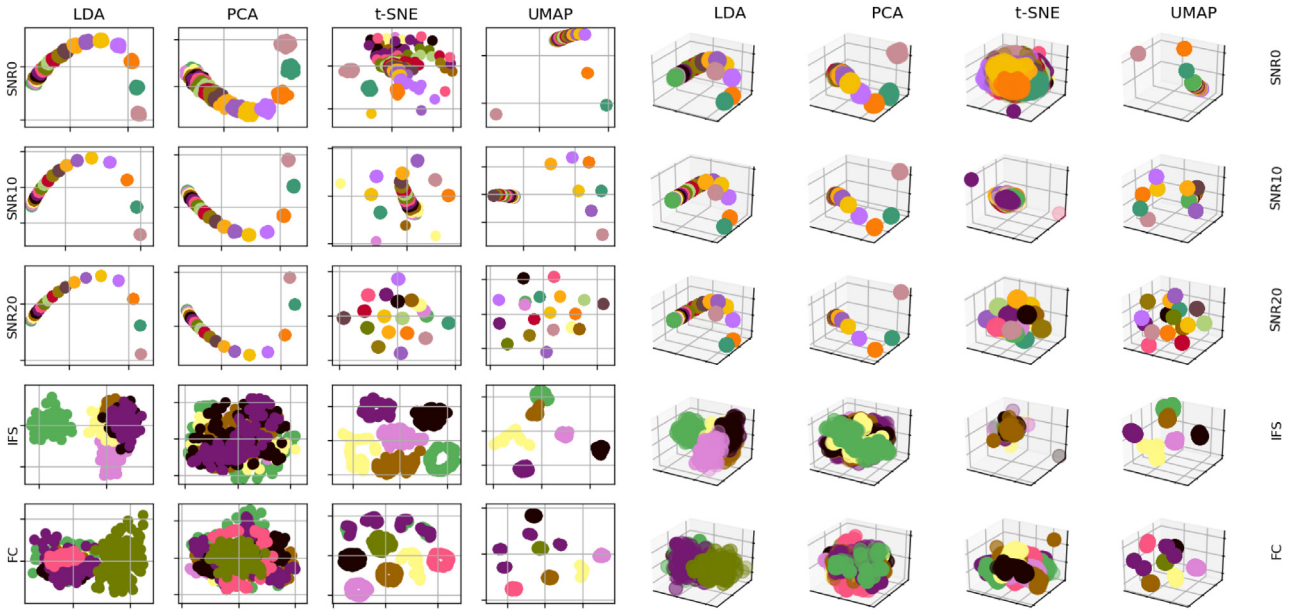
Fig. 11. Result of approach II where the unsupervised DR technique is used as both a damage indexes extractor and a dimension reduction tool fed directly by raw signals X_{RAW} that finally projects data toward $\mathbb{R}^2$ *[Left]* or $\mathbb{R}^3$ *[Right]* maps (see Fig. 3) for all the considered datasets.
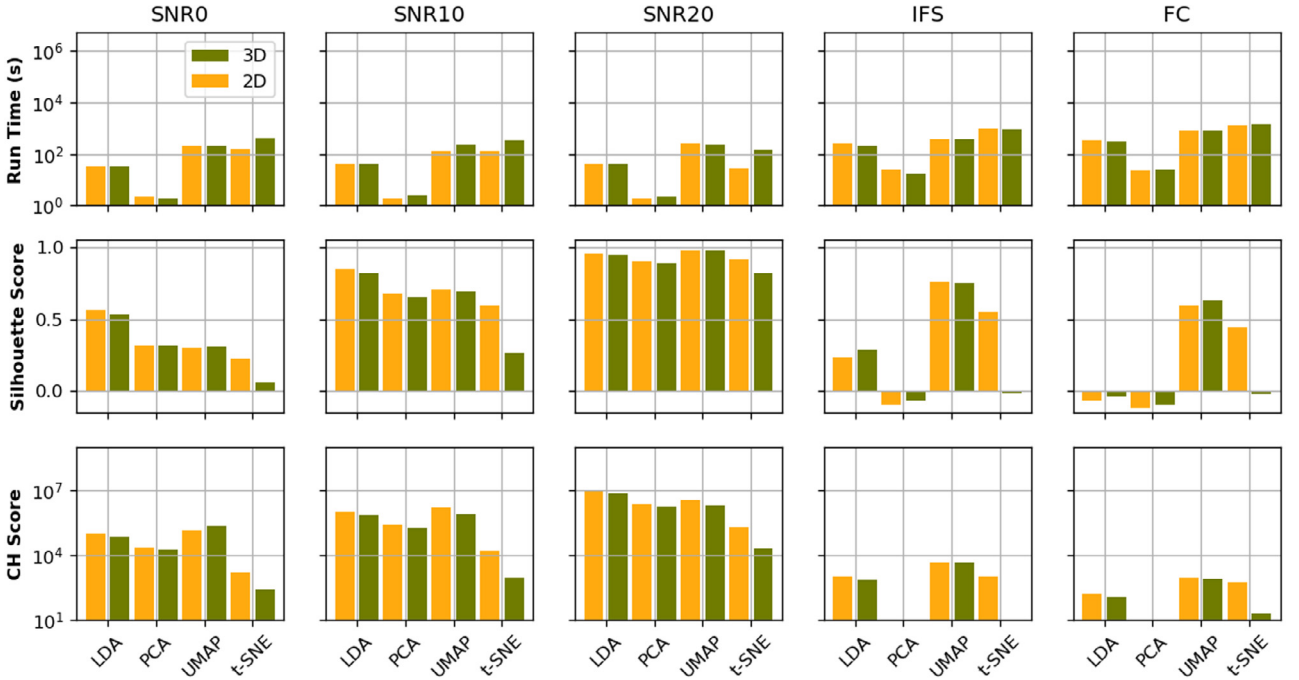

Fig. 12. Scores of approach II where the unsupervised DR technique is used as both a damage indexes extractor and a dimension reduction tool fed directly by raw signals X_{RAW} that finally projects data toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps (see Fig. 3) for all the considered datasets.

computation times, it can be observed that on the simulated datasets (SNR0, SNR10 and SNR20), UMAP and t-SNE needs more resources than PCA or LDA. Again, this gap in computation time between algorithms decreases by one or two orders of magnitudes when applied to the experimental datasets FC and IFS. It can be furthermore observed that the level of noise (SNR) has still a large effect on silhouette score with an increase for all algorithms with the SNR. By analyzing the silhouette scores, the total failure of PCA and LDA in clustering experimental data coming from the FC and IFS datasets can be observed. Among the benchmarked algorithms, UMAP generally provides a higher value for silhouette score and even outperforms LDA for the SNR20, IFs and FC datasets. Regarding CH score, the performances of UMAP and

t-SNE are stable over the various datasets even if a decrease when moving from simulated to experimental data can be noticed. Obtained silhouette scores for UMAP and t-SNE are in general comparable to PCA and LDA performances for the simulated cases and higher to the experimental ones. Finally, when comparing the computation time and score obtained for the $\mathbb{R}^2$ or $\mathbb{R}^3$ cases, no significant effect of the embedding dimension can be observed except that t-SNE performs very badly in both experimental cases when projecting data toward $\mathbb{R}^3$.

Finally, the scores obtained by means of approach I and II can be compared to conclude with respect to the usefulness of damage indexes computation. As the embedding dimension seems to have a minor effect on the scores, the choice has been made to compare results of both
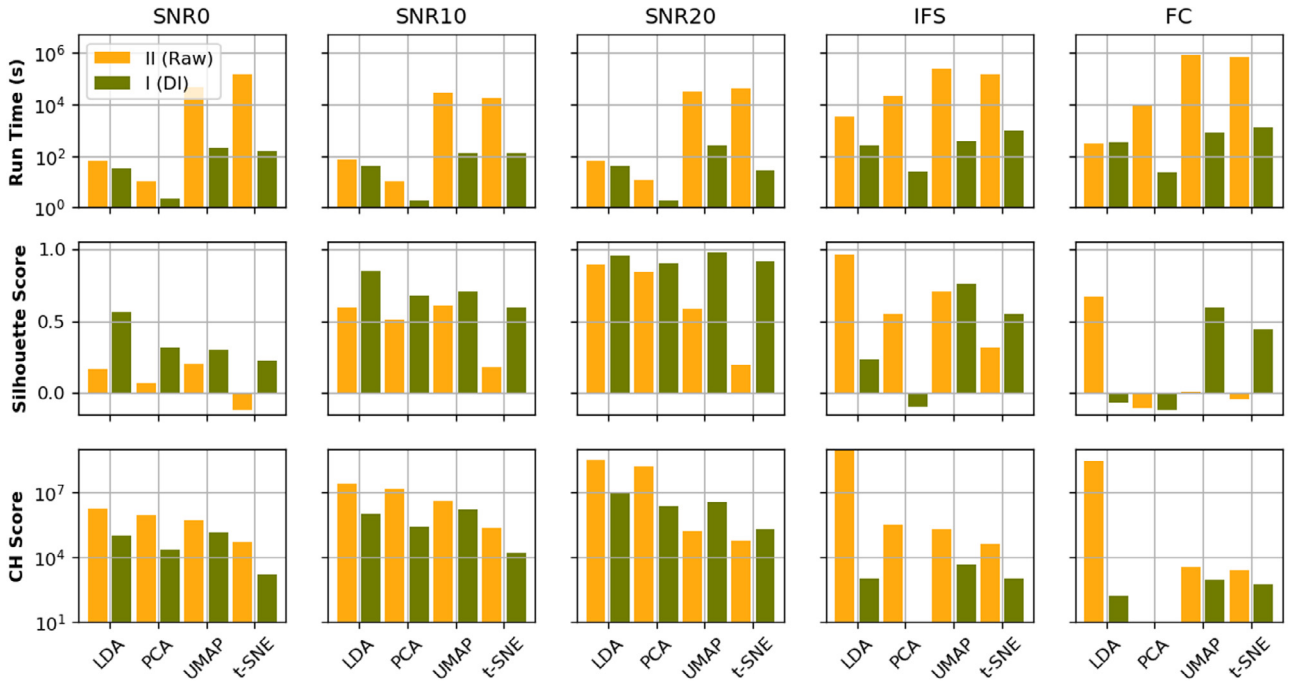
Fig. 13. Scores of approaches I and II for projections of data toward $\mathbb{R}^2$ maps for all the considered datasets.

approaches when data are projected to $\mathbb{R}^2$. The scores of approaches I and II for projections of data toward $\mathbb{R}^2$ maps for all the considered datasets are presented in Fig. 13. It can be seen that computation time is larger for approach I than for approach II. This is due to the fact that in approach I, repetitions signals of each case are compared with repetitions signals of the reference case leading to a much higher number of samples to cluster than for approach II. However, this should be counterbalanced by the fact that samples in approach II contain many more features than the ones of approach I. If computation time associated with damage index computation is also taken into account, it can be concluded that approach II is very interesting in comparison with approach I. Regarding the silhouette score, it can be stated that whatever the dataset or the algorithm considered, approach II leads to better scores than the approach I. This means that approach II is also interesting with respect to that criterion. Finally, the CH scores associated with approach I are in general larger than the ones obtained using approach II, the difference between both approaches being relatively small. However care should be taken when looking at CH score as, when considered alone, it is not really indicative of clustering performances. For example, UMAP obtains a correct CH score for the FC case using approach I whereas its silhouette score is close to 0 meaning that it is in fact unable to cluster correctly. To conclude, it can be said that this criterion is globally equivalent between both approaches but that it is less important than the silhouette score.

5. Extended dimension reduction algorithms for SHM applications

5.1. General methodology to extend t-SNE and UMAP for SHM

The last contribution addressed here is related to the extension of t-SNE and UMAP in order to learn the projection and to solve the out-sample issue. Concerning SHM applications, the main drawback of UMAP and t-SNE is their inability to cluster any incoming and unknown sample for t-SNE and the hidden behavior of projection for t-SNE and UMAP. This is because what is learned by UMAP or t-SNE is a low dimensional map that has similar topological or statistical properties than the high dimensional map but not the transformations allowing to project one sample x_i living in the high dimensional space to y_i living in the low dimensional one. A learning approach based on a

neural network is proposed here to catch the hidden behavior of t-SNE and UMAP projections and to address this out-sample problem which is particularly relevant with respect to SHM applications. A specific outlier detector for t-SNE or UMAP already exist (Roman-Rangel and Marchand-Maillet, 2019; Espadoto et al., 2019) but as a general solution neural networks (Altan and Karasu, 2019; Altan et al., 2019, 2018; Karasu and Altan, 2019) are proposed to consider all the DR techniques. The underlying neural network is trained on a small subset of data provided by the dimension reduction techniques to learn the projection from high dimensional space to low dimensional space and is then used to process any incoming new sample. Such an approach has already been tested successfully using t-SNE or UMAP for other applications (Roman-Rangel and Marchand-Maillet, 2019; Espadoto et al., 2019) but not for SHM applications. According to Section 4, approach II is more efficient in terms of computational time and of clustering efficiency. This approach is thus used to illustrate the extension procedure of UMAP and t-SNE proposed here.

Practically speaking, a neural network is trained on a training set consisting of 20 % of uniformly and randomly selected samples of X_{RAW} and their related 2D or 3D projections. The methodology used here to train and test the neural networks for each X_{RAW} datasets classically includes three successive steps: (i) selection of the training data as detailed previously, (ii) training of the neural network on this data set, and finally (iii) projection of the remaining 80 % of the dataset by the previously trained neural network for validation purposes. As a first pre-processing step, a principal component analysis is first achieved to reduce the dimension of input data to $\mathbb{R}^{10}$. The PCA solely captures the linear features. The lack of nonlinearity of PCA is compensated by the neural network. Then, a neural network built upon 4 fully connected layers that respectively include 1000, 800, 256 and 2 or 3 neurons is trained. The two first layers are powered by RelU activation functions and the remaining layers by tanh functions. An overview of the chosen network is provided in Fig. 14. Before training, the weights have been initialized by uniform variance scaling and a small bias value is set (0.0001). The neural network is then optimized using the Adam optimizer. Runs of 200 epochs with 128 batch-size are considered. Furthermore, early stopping criteria with 20 patience is selected.
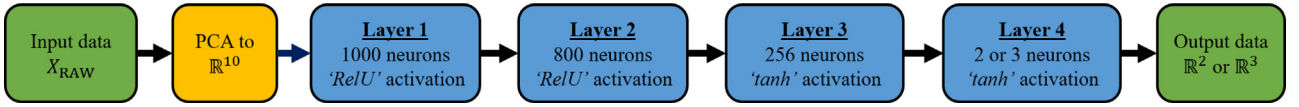
Fig. 14. Proposed neural network architecture to learn the UMAP and t-SNE clustering maps.
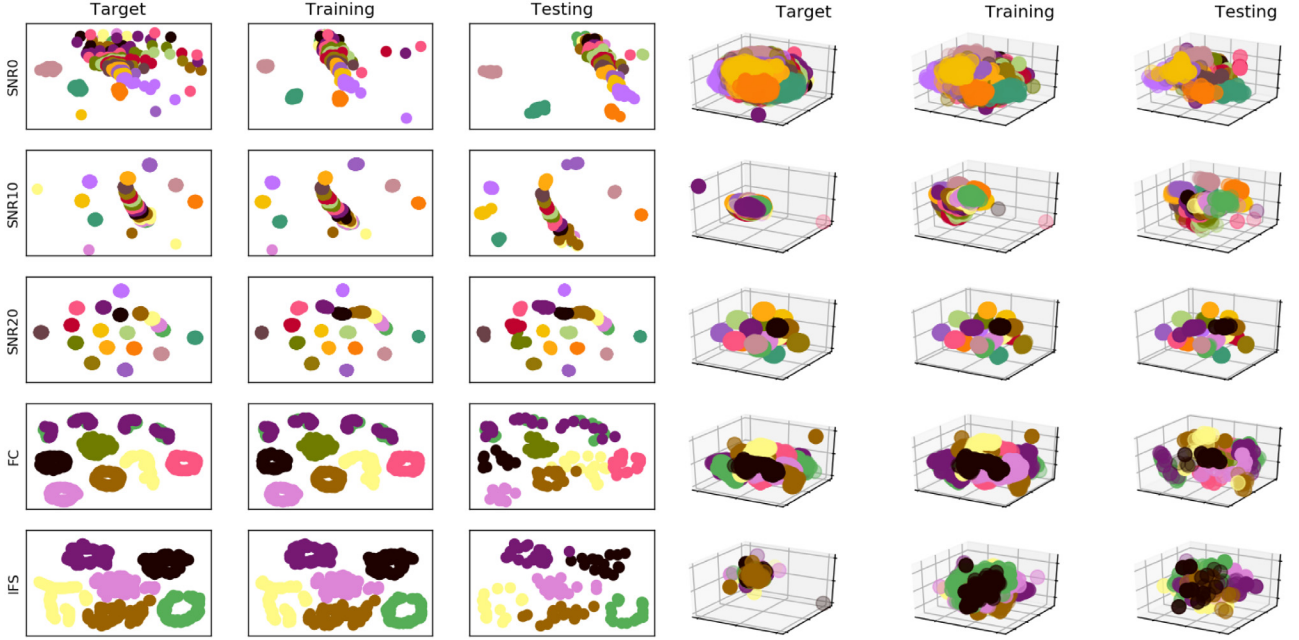

Fig. 15. Projection results obtained by neural networks in comparison with the original t-SNE clustering results (*[Left]*: 2D, *[Right]*: 3D). *"Target"* refers to the original clustering, *"Training"* to the projection obtained after training the neural network on 20 % of the X_{RAW} datasets, and *"Testing"* to the projection obtained after testing the neural network on the remaining 80 % of the X_{RAW} datasets.

5.2. Validation of the neural networks extension of UMAP and t-SNE

Projection results obtained by neural networks in comparison with the original t-SNE results are provided in Fig. 15 for both the $\mathbb{R}^2$ *[Left]* and the $\mathbb{R}^3$ *[Right]* cases. The *"Target"* columns refer to the original clustering results provided by t-SNE, the *"Training"* columns to the projections obtained after training the neural network on 20 % of the X_{RAW} datasets, and the *"Testing"* columns to the projection obtained after testing the neural network on the remaining 80 % of the X_{RAW} datasets. It can be observed that for any given X_{RAW} dataset the clustering results for the *"Target"*, *"Training"* and *"Testing"* cases are equivalent even if not exactly similar. The global patterns discovered by t-SNE are correctly caught by the neural networks during the training phase, and these clusters are still present in the testing phase. As a consequence, it can be said that the projections achieved by t-SNE following approach II can be adequately learned by a neural network and generalized to upcoming unknown samples for the tested datasets.

Projection results obtained by neural networks in comparison with the original UMAP results are furthermore provided in Fig. 16 for both the $\mathbb{R}^2$ *[Left]* and the $\mathbb{R}^3$ *[Right]* cases. As in the previous figure, the *"Target"* columns refer to the original clustering results provided by UMAP, the *"Training"* columns to the projections obtained after training the neural network on 20 % of the X_{RAW} datasets, and the *"Testing"* columns to the projection obtained after testing the neural network on the remaining 80 % of the X_{RAW} datasets. From this figure it can again be observed that for any given X_{RAW} dataset the clustering results for the *"Target"*, *"Training"* and *"Testing"* cases are equivalent even if not exactly similar. The global patterns discovered by UMAP are correctly caught by the neural networks during the training phase, and these clusters are still present in the testing phase. As a

consequence, it can also be said that the projections achieved by UMAP following approach II can be adequately learned by a neural network and generalized to upcoming unknown samples for the tested datasets.

The quality of the projections achieved by the neural networks can also be assessed quantitatively through the use of the previously defined silhouette and CH scores. Fig. 17 presents a comparison of the silhouette and CH scores obtained by the selected dimension reduction methods (t-SNE and UMAP) in comparison with the neural networks (NN) supervised learning versions over the whole X_{RAW} datasets for all the cases under study. The labels *"NN(t-SNE)"* and *"NN(UMAP)"* refer to the t-SNE and UMAP projections learned by the neural networks. It can be seen that even if the learning achieved by the neural networks is not perfect, the silhouette scores slightly increase when using neural networks in place of the original dimension reduction algorithms. More precisely, the silhouette scores increase on simulation data but remain almost equal for the experimental data. Regarding the CH score, the same observation can be made for the SNR0 and SNR10 cases, with an increase of the CH score for the neural networks cases. However for the SNR20 case, a decrease of the CH score is observed. For the experimental datasets, CH scores remain very close. As a consequence, it can be stated that the projection learned by neural networks are equivalent to the original projections, but with a great advantage: they are not prone to the out-sample issue and are able to correctly cluster any new incoming sample, which is mandatory for SHM applications.

6. Conclusion and discussion

Structural Health Monitoring (SHM), *i.e.* the action of monitoring structures in real-time and in an automated manner, is a major challenge in several industrial fields such as aeronautic or civil engineering.
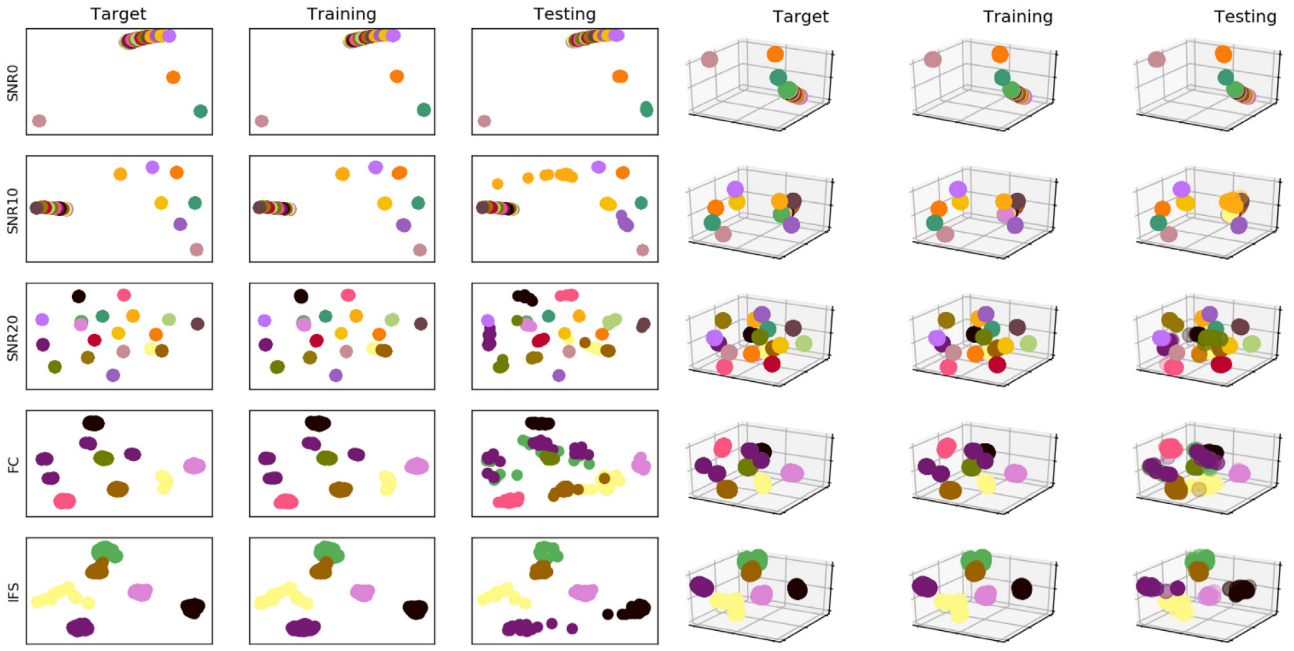
Fig. 16. Projection results obtained by neural networks in comparison with the original UMAP clustering results ([*Left*]: $\mathbb{R}^2$, [*Right*]: $\mathbb{R}^3$). *"Target"* refers to the original clustering, *"Training"* to the projection obtained after training the neural network on 20 % of the X_{RAW} datasets, and *"Testing"* to the projection obtained after testing the neural network on the remaining 80 % of the X_{RAW} datasets.
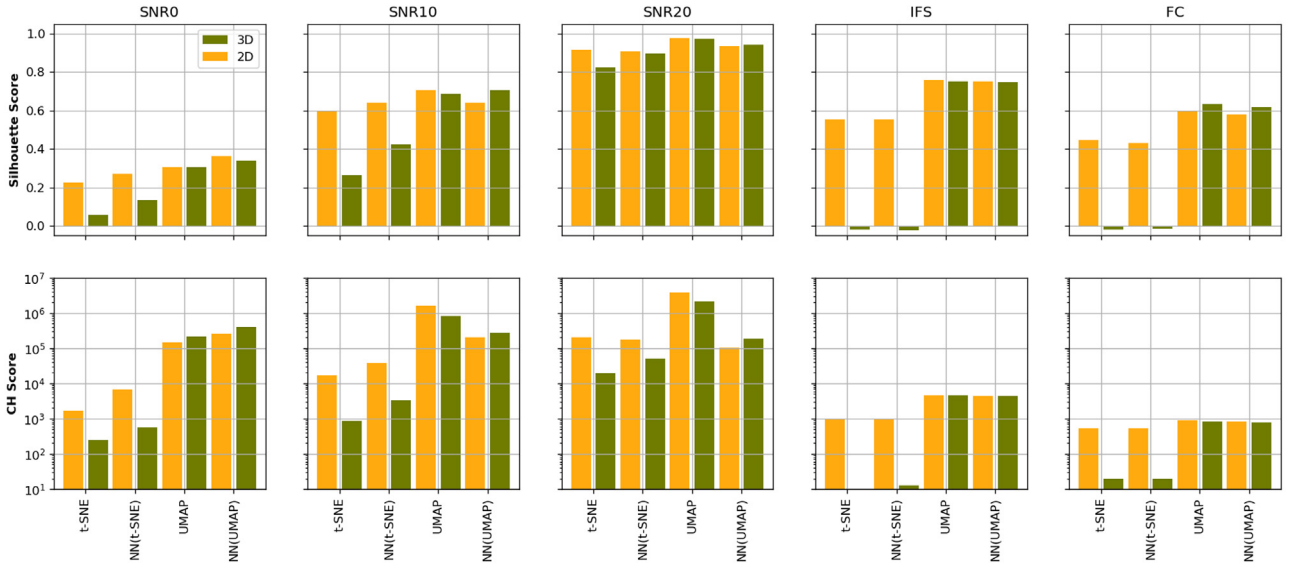

Fig. 17. Comparison of the silhouette and CH scores obtained by the selected dimension reduction methods (t-SNE and UMAP) in comparison with the neural networks (NN) supervised learning versions over the whole X_{RAW} datasets for all the cases under study.

SHM is by nature a very high dimensional data-driven problem that possesses several specificities when addressed as a machine learning problem: data in damaged cases are rare and very costly, any incoming sample should be instantaneously clustered, and handcrafted damage indexes are commonly used as a first dimension reduction step due to large datasets. As a consequence, unsupervised dimensionality reduction (DR) techniques that project very high dimensional data into a two or three-dimensional space (such as t-SNE or UMAP) are very appealing in such a context. However these methods suffer from one major drawback with respect to SHM: they are unable to cluster any unknown incoming sample (out-sampling problem). The high potential impact of t-SNE and UMAP for SHM purposes and an efficient way to circumvent the out-sample issue have been explored through various

simulated and experimental datasets coming for complex aeronautical composite structures monitored by Lamb waves. The contributions presented here can be summarized as follows:

- *Dimension reduction (DR) algorithms benchmarking on SHM datasets*: all the applicable dimension reduction technique from PCA to state of the art algorithms (t-SNE, UMAP) have been benchmarked on an unsupervised damage clustering task in complex aeronautical composite structures monitored by Lamb waves in order to highlight the potential of t-SNE and UMAP for SHM. A solution to efficiently tune the DR algorithms parameters has been proposed. Obtained results demonstrate that ISOMAP, LLE and MDS are not suitable to process such datasets. Furthermore, even if PCA can

Table 2
Implemented damage indexes.

DI name	Comments	Definitions	DI name	Comments	Definitions
CC	FFT based implementation of the maximum of the correlation	$1 - max\left(\dfrac{IFFT[FFT[x_{ij}(t)]FFT[y_{ij}(t)]^*]}{\sqrt{(E_{x_{ij}}E_{y_{ij}})}}\right)$	CCA	MATLAB based implementation of the maximum of the correlation	$1 - max(xcorr[x_{ij}(t), y_{ij}(t)])$
CC0	MATLAB based implementation of the zero-lag correlation	$1 - xcorr[x_{ij}(t), y_{ij}(t)](0)$	CRC	MATLAB based implementation of the correlation coefficient	$1 - corrcoef[x_{ij}(t), y_{ij}(t)]$
NRE	Normalized residual energy	$\dfrac{\int_T^0 (x_{ij}(t) - y_{ij}(t))^2 dt}{2 \times \int_T^0 x_{ij}(t)^2 dt + \int_T^0 y_{ij}(t)^2 dt}$	NRE2	Normalized residual energy 2	$\dfrac{\int_T^0 x_{ij}(t)^2 dt - \int_T^0 y_{ij}(t)^2 dt}{max(\int_T^0 x_{ij}(t)^2 dt, \int_T^0 y_{ij}(t)^2 dt)}$
NRE3	Normalized residual energy 3	$1 - \dfrac{min(\int_T^0 x_{ij}(t)^2 dt, \int_T^0 y_{ij}(t)^2 dt)}{max(\int_T^0 x_{ij}(t)^2 dt, \int_T^0 y_{ij}(t)^2 dt)}$	NRE4	Normalized residual energy 4	$\sqrt{\dfrac{\int_T^0 (x_{ij}(t)^2 dt \times y_{ij}(t)^2 dt)}{max(\int_T^0 x_{ij}(t)^2 dt, \int_T^0 y_{ij}(t)^2 dt)}}$
NRE5	Normalized residual energy 5	$\dfrac{\int_T^0 (x_{ij}(t)/max(x_{ij}) - y_{ij}(t)/max(y_{ij}))^2 dt)^2 dt}{2 \times (\int_T^0 [x_{ij}(t)/max(x_{ij})]^2 dt \int_T^0 [y_{ij}(t)/max(y_{ij})]^2 dt)^2}$	NRE6	Normalized residual energy 6	$(1 - NRE4)\dfrac{min(\int_T^0 x_{ij}(t)^2, \int_T^0 y_{ij}(t)^2)}{max(\int_T^0 x_{ij}(t)^2, \int_T^0 y_{ij}(t)^2)}$
MA	Maximum Amplitude of difference	$max[x_{ij}(t) - y_{ij}(t)]$	MAR	Maximum Amplitude Relative	$max[x_{ij}(t) - y_{ij}(t)]/max(x_{ij}(t))$
FFT	FFT ratio of the difference signal over the sum off signals at f0	$\left(\dfrac{FFT[x_{ij}(t) - y_{ij}(t)]_{f=f_0}}{FFT[x_{ij}(t) + y_{ij}(t)]_{f=f_0}}\right)$	STFT	Short Time Fourier Transform	$\left(\dfrac{FFT[y_{ij}(t)W(t)]_{f=f_0}}{FFT[x_{ij}(t)W(t)]_{f=f_0}}\right)$
Env	Energy of the envelope of the difference	$\sqrt{\dfrac{\sum_{t=T_1}^{T_2} env[y_{ij} - x_{ij}]^2}{\sum_{t=T_1}^{T_2} env[x_{ij}]^2}}$	PHI	Energy of the phase of the difference	$\sqrt{\dfrac{\sum_{t=T_1}^{T_2} \Phi[y_{ij} - x_{ij}]^2}{\sum_{t=T_1}^{T_2} \Phi[x_{ij}]^2}}$
TDM	Time Delay of Max	$\dfrac{argmax(\mathcal{H}(y_{ij})) - argmax(\mathcal{H}(x_{ij}))}{argmax(\mathcal{H}(x_{ij}(t)))}$	TD1	Time Delay of the first wave packet	$TD1 = \dfrac{t_2 - t_1}{t_1}$
SPAR	Signal Amplitude Peak Ratio	$\left|1 - \dfrac{max(|y_{ij}(t)|)}{max(|x_{ij}(t)|)}\right|$	SAPS	Signal Amplitude Peak squared percentage differences	$\left[\dfrac{max(|y_{ij}(t)|) - max(|x_{ij}(t)|)}{max(|x_{ij}(t)|)}\right]^2$
SAHM	Signal Amplitude Hilbert transform Maximum percentage differences	$\left[\dfrac{max(|(\mathcal{H})(y_{ij}(t))|) - max(|(\mathcal{H})(x_{ij}(t))|)}{max(|(\mathcal{H})(y_{ij}(t))|)}\right]^2$	SSSD	Signal Sum of Squared Differences	$\dfrac{\int_0^T (x_{ij}(t) - y_{ij})^2 dt}{\int_0^T x_{ij}(t)^2 dt}$
CCTOF	Cross-correlation-based TOF percentage difference	$\left|\dfrac{argmax(xcorr[x_{ij}, x_{ij}]) - argmax(xcorr[x_{ij}, y_{ij}])}{argmax(xcorr[x_{ij}, x_{ij}])}\right|$	WPSD	Welch-based Power Spectral Density	–
WTF	Welch-based transfer function	–	CCMPD	Cross-correlation maximum percentage difference	–
DWTC	Discrete Wavelet Transform approximation coefficient	–			

in some cases provide acceptable results, t-SNE and particularly UMAP largely outperform qualitatively and quantitatively PCA for SHM clustering purposes on the tested datasets.

- *Usefulness of damage indexes*: Dimension reduction techniques have been tested in two ways in order to assess whether or not the computation of damage indexes is useful for clustering purposes:

    1. Approach I: the unsupervised DR technique is used in order to project the data from the handcrafted damage indexes space toward $\mathbb{R}^2$ or $\mathbb{R}^3$ maps.
    2. Approach II: the unsupervised DR technique is used as both a damage indexes extractor and a dimension reduction tool fed directly by raw signals that projects data toward a $\mathbb{R}^2$ or $\mathbb{R}^3$ map.

Obtained results indicate that the dimension of the output space ($\mathbb{R}^2$ or $\mathbb{R}^3$) is of little importance for SHM applications but that the retained approach is not. Clustering results obtained through approach II are much more convincing than the ones obtained using approach I. More specifically, t-SNE and UMAP exhibit very good clustering abilities using approach II. In that context, computing damages indexes is not anymore needed for clustering purposes when relying on approach II and on t-SNE or UMAP.

- *Extension of t-SNE and UMAP to solve the out-sample issue*: a learning approach based on neural networks has been proposed to catch the hidden behavior of t-SNE and UMAP projections and to address the out-sample problem. The underlying neural networks are trained on a small subset of 20 % of data provided by the DR techniques to learn the projection from high dimensional space to low dimensional space and are then used to process any incoming new sample. Obtained results first indicate that the generic neural network architecture retained here is able to learn both t-SNE and UMAP projections through approach II. Secondly, the tests of the trained neural networks on the remaining 80 % of the datasets provide very good clustering results from both qualitative and quantitative points of view. It can be concluded that this neural network extension of UMAP and t-SNE constitutes an adequate solution to the *out-sample* issue and thus promotes the adoption of t-SNE and UMAP for SHM applications.

The fact that computing handcrafted damage indexes is not necessary can be linked with previous observations done in the field of image analysis and pattern recognition. The classical approach in pattern recognition consists of finding the best set of features able to describe the image being processed. As in SHM, these features were initially handcrafted. With the arrival of machine learning, successful attempts to avoid or limit designing and computing handcrafted features for this specific application have been published (Frintrop et al., 2015; Handa et al., 2016; DeCost et al., 2017). The specificity of SHM application

Table 3
Overview of the simulated database available with SNR equal to 0 dB, 10 dB and 20 dB.

Class	# of paths	# of samples	# of features	# of rep
Healthy	25	1000	26	150
Damage 5%	25	1000	26	150
Damage 10%	25	1000	26	150
Damage 15%	25	1000	26	150
Damage 20%	25	1000	26	150
Damage 25%	25	1000	26	150
Damage 30%	25	1000	26	150
Damage 35%	25	1000	26	150
Damage 40%	25	1000	26	150
Damage 45%	25	1000	26	150
Damage 50%	25	1000	26	150
Damage 55%	25	1000	26	150
Damage 60%	25	1000	26	150
Damage 65%	25	1000	26	150
Damage 70%	25	1000	26	150
Damage 75%	25	1000	26	150
Damage 80%	25	1000	26	150
Damage 85%	25	1000	26	150
Damage 90%	25	1000	26	150

with respect to pattern recognition lies in the fact that SHM datasets possess much more features and that the two-dimensional correlation existing in images is only slightly present in SHM datasets due to the sparsity of piezoelectric sensors networks. Despite these differences, the same conclusion is drawn.

A second point to discuss with respect to the results presented here is related to the physical interpretation of the low dimensional representation provided by t-SNE or UMAP and then learned by neural networks. With respect to SHM applications, the input data are physically meaningful and the expected outputs are also physically meaningful. Indeed, the end users expect to know damage localization, size, type as an output of the SHM process. As the projections achieved here are low dimensional, an interesting perspective would be to be able to link the dimensions of the low dimensional space with physical dimensions (damage location coordinates and size for example). One approach to do so could be to find a way to keep the physical interpretation of the processed Lamb wave signals and previous knowledge of the inspected structure during the machine learning steps (in particular time information and piezoelectric element physical positions). Another approach could be to use some samples in a supervised manner in order to physically *scale* the resulting low dimensional maps.

Finally, what is shown here is that a t-SNE or UMAP projection associated with a neural network has to be learned for each aeronautical structure under study. This is a limiting factor in practice as one would like to be able to build one clustering model for a set of similar structures and not to need to learn a model for each new structure. Thus, transfer learning, *i.e.* the ability of one clustering model learned on a given structure to cluster data coming from another structure should be assessed in order to make this approach really relevant for SHM purposes (Bull et al., 2020; Gosliga et al., 2020; Gardner et al., 2020).

Appendix A. Damage indexes computation

Let x_{ij} and y_{ij} be the reference and unknown signals for the path between actuator i and sensor j. The 26 damage indexes being computed are defined in Table 2.

Appendix B. Databases details

For the simulated database, the sampling frequency is $f_s = 3.33$ MHz. The damage is represented by a local reduction in material properties in the damaged area. The rigidity is reduced step by steps from the healthy state toward a totally damaged state resulting in one healthy reference case and 18 damaged cases. After the simulation, a white Gaussian noise is added to the simulation results. A central frequency of $f_0 = 120$ kHz is used with SNR values equal to 0 dB, 10 dB and 20 dB. The X_{RAW} database contains 2850 samples, each one associated with 25 000 features and divided into 19 different classes. The X_{DI} database contains 416 175 samples, each one associated with 650 features and divided into 19 different classes.

The first geometrically complex aeronautics structure under study consists here is the fan cowl part of a nacelle of an Airbus A380 equipped with 30 piezoelectric elements. The excitation signal sent to the PZT element is a 5 cycles burst with an excitation frequency of $f_0 = 200$ kHz and with an amplitude of 10 V. Resulting signals are then simultaneously recorded by the other piezoelectric elements and consist of 1000 data points sampled at 1 MHz. Signals were acquired 100 times in both the healthy (reference) and damaged (unknown) states (except the last damaged state for which 150 repetitions have been achieved). One healthy case and 7 damage cases have been considered. The X_{RAW} database contains 800 samples, each one associated with 900 000 features and divided into 8 different classes. The X_{DI} database contains 123 675 samples, each one associated with 23 400 features and divided into 8 different classes.

The other geometrically complex aeronautics structure under study is the inner fixed structure part of a nacelle of an Airbus A380 equipped

Table 4
Overview of the experimental database "Fan Cowl".

Class	# of paths	# of samples	# of features	# of rep.
Healthy	900	1000	26	100
Magnets (P5:35 mm & P1:40 mm)	900	1000	26	100
Magnets (P2:35 mm)	900	1000	26	100
Magnets (P3:35 mm)	900	1000	26	100
Magnets (P4, a:35 mm, b:40 mm & c:45mm)	900	1000	26	100
Magnets (P1:45 mm + 40 mm)	900	1000	26	100
Magnets (P1:45 mm)	900	1000	26	100
Hole (P1:6 mm)	900	1000	26	150

Table 5
Overview of experimental database "Inner Fixed Structure".

Class	# of paths	# of samples	# of features	# of rep.
Healthy	841	1000	26	100
6 mm hole, inner skin, T1	841	1000	26	100
6 mm hole, both skins, T1	841	1000	26	100
6 mm holes, T1 & T2	841	1000	26	100
6 mm holes, T1, T2 & T3	841	1000	26	100
6 mm holes, T1, T2, T3 & T4	841	1000	26	100

Table 6
Overview of the tuning parameters selected for UMAP and t-SNE for the various datasets under study and for the approaches DI and Raw.

Data set	Approach	Algorithm	Distance metric	Perplexity or number of neighbors
FC	I	t-SNE	–	–
		UMAP	Manhattan	60
	II	t-SNE	Manhattan	20
		UMAP	Manhattan	20
IFS	I	t-SNE	Manhattan	55
		UMAP	Manhattan	65
	II	t-SNE	Manhattan	10
		UMAP	Manhattan	10
SNR0	I	t-SNE	Euclidean	65
		UMAP	Manhattan	65
	II	t-SNE	Correlation	65
		UMAP	Manhattan	65
SNR10	I	t-SNE	Chebyshev	65
		UMAP	Euclidean	60
	II	t-SNE	Correlation	70
		UMAP	Manhattan	60
SNR20	I	t-SNE	Manhattan	65
		UMAP	Manhattan	65
	II	t-SNE	Cosine	65
		UMAP	Manhattan	65

with 29 piezoelectric elements. The excitation signal sent to the PZT element is a 5 cycles burst with an excitation frequency of $f_0 = 100$ kHz and with an amplitude of 10 V. Resulting signals are then simultaneously recorded by the other piezoelectric elements and consist of 1000 data points sampled at 1 MHz. Signals were acquired 100 times in both the healthy (reference) and damaged (unknown) states. One healthy case and 5 damage cases have been considered. The X_{RAW} database contains 600 samples, each one associated with 841 000 features and divided into 6 different classes. The X_{DI} database contains 54 950 samples, each one associated with 21 866 features and divided into 6 different classes.

Appendix C. Hyperparameters selection

The hyperparameters chosen for all the datasets, for t-SNE and UMAP algorithms, and for the approaches I and II are summarized in Table 6. It can be observed from this table that choosing the Manhattan distance metric with a perplexity or a number of neighbors between 50 and 60 leads to good clustering results most of the time.

References

Altan, A., Aslan, O., Hacıoğlu, R., 2018. Real-time control based on NARX neural network of hexarotor UAV with load transporting system for path tracking. In: 2018 6th International Conference on Control Engineering & Information Technology. CEIT, IEEE, pp. 1–6.

Altan, A., Karasu, S., 2019. The effect of kernel values in support vector machine to forecasting performance of financial time series. J. Cogn. Syst. 4 (1), 17–21.

Altan, A., Karasu, S., Bekiros, S., 2019. Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. Chaos Solitons Fractals 126, 325–336.

Balmes, E., 2020. Structural dynamics toolbox (for use with MATLAB). www.sdtools.com.

Bhuiyan, M.Z.A., Wang, G., Wu, J., Cao, J., Liu, X., Wang, T., 2015. Dependable structural health monitoring using wireless sensor networks. IEEE Trans. Dependable Secure Comput. 14 (4), 363–376.

Borate, P., Wang, G., Wang, Y., 2020. Data-driven structural health monitoring approach using guided Lamb wave responses. J. Aerosp. Eng. 33 (4), 04020033.

Bouzenad, A.E., El Mountassir, M., Yaacoubi, S., Dahmene, F., Koabaz, M., Buchheit, L., Ke, W., 2019. A semi-supervised based K-Means algorithm for optimal guided waves structural health monitoring: A case study. Inventions 4 (1), 17.

Bull, L.A., Gardner, P.A., Gosliga, J., Rogers, T.J., Dervilis, N., Cross, E.J., Papatheou, E., Maguire, A.E., Campos, C., Worden, K., 2020. Foundations of population-based SHM, part I: homogeneous populations and forms. Mech. Syst. Signal Process. 148, 107141.

Cai, J., Luo, J., Wang, S., Yang, S., 2018. Feature selection in machine learning: A new perspective. Neurocomputing 300, 70–79.

Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stat. 3 (1), 1–27.

Chandrashekar, G., Sahin, F., 2014. A survey on feature selection methods. Comput. Electr. Eng. 40 (1), 16–28.

DeCost, B.L., Jain, H., Rollett, A.D., Holm, E.A., 2017. Computer vision and machine learning for autonomous characterization of am powder feedstocks. JOM 69 (3), 456–465.

Di Sante, R., 2015. Fibre optic sensors for structural health monitoring of aircraft composite structures: Recent advances and applications. Sensors 15 (8), 18666–18713.

El Mountassir, M., Yaacoubi, S., Ragot, J., Mourot, G., Maquin, D., 2016. Feature selection techniques for identifying the most relevant damage indices in SHM using guided waves. In: 8th European Workshop on Structural Health Monitoring. EWSHM 2016.

Espadoto, M., Hirata, N.S.T., Telea, A.C., 2019. Deep learning multidimensional projections. arXiv preprint arXiv:1902.07958.

Farrar, C.R., Worden, K., 2012. Structural Health Monitoring: A Machine Learning Perspective. John Wiley & Sons.

Fendzi, C., Mechbal, N., Rébillat, M., Guskov, M., Coffignal, G., 2016a. A general Bayesian framework for ellipse-based and hyperbola-based damage localization in anisotropic composite plates. J. Intell. Mater. Syst. Struct. 27 (3), 350–374.

Fendzi, C., Rebillat, M., Mechbal, N., Guskov, M., Coffignal, G., 2016b. A data-driven temperature compensation approach for structural health monitoring using Lamb waves. Struct. Health Monit. 15 (5), 525–540.

Frintrop, S., Werner, T., Martin, G.G., 2015. Traditional saliency reloaded: A good old model in new shape. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 82–90.

Gardner, P., Bull, L.A., Gosliga, J., Dervilis, N., Worden, K., 2020. Foundations of population-based SHM, part III: heterogeneous populations-mapping and transfer. Mech. Syst. Signal Process. 149, 107142.

Gharibnezhad, F., Mujica, L.E., Rodellar, J., 2015. Applying robust variant of principal component analysis as a damage detector in the presence of outliers. Mech. Syst. Signal Process. 50, 467–479.

Gosliga, P.A., Gardner, P.A., Bull, L.A., Dervilis, N., Worden, K., 2020. Foundations of population-based SHM, part II: heterogeneous populations-graphs, networks, and communities. Mech. Syst. Signal Process. 148, 107144.

Handa, A., Bloesch, M., Pătrăucean, V., Stent, S., McCormac, J., Davison, A., 2016. GVNN: Neural network library for geometric computer vision. In: European Conference on Computer Vision. Springer, pp. 67–82.

Harvey, D.Y., Todd, M.D., 2014. Structural health monitoring feature design by genetic programming. Smart Mater. Struct. 23 (9), 095002.

Hinders, M.K., Miller, C.A., 2020. Intelligent structural health monitoring with ultrasonic Lamb waves. In: Intelligent Feature Selection for Machine Learning using the Dynamic Wavelet Fingerprint. Springer, pp. 45–114.

Hinton, G.E., Roweis, S.T., 2003. Stochastic neighbor embedding. In: Advances in Neural Information Processing Systems. pp. 857–864.

Karasu, S., Altan, A., 2019. Recognition model for solar radiation time series based on random forest with feature selection approach. In: 2019 11th International Conference on Electrical and Electronics Engineering. ELECO, IEEE, pp. 8–11.

Lizé, E., 2018. Détection d'endommagement sans état de référence et estimation de la température pour le contrôle santé intégré de structures composites par ondes guidées (Ph.D. thesis). ENSAM, Paris.

Manson, G., Worden, K., Allman, D., 2003. Experimental validation of a structural health monitoring methodology: Part II. Novelty detection on a GNAT aircraft. J. Sound Vib. 259 (2), 345–363.

McInnes, L., Healy, J., Melville, J., 2018. UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

Miller, C.A., Hinders, M.K., 2014. Multiclass feature selection using computational homology for Lamb wave–based damage characterization. J. Intell. Mater. Syst. Struct. 25 (12), 1511–1527.

Mitra, M., Gopalakrishnan, S., 2016. Guided wave based structural health monitoring: A review. Smart Mater. Struct. 25 (5), 053001.

Papaelias, M., Cheng, L., Kogia, M., Mohimi, A., Kappatos, V., Selcuk, C., Constantinou, L., Muñoz, C.Q.G., Marquez, F.P.G., Gan, T., 2016. Inspection and structural health monitoring techniques for concentrated solar power plants. Renew. Energy 85, 1178–1191.

Park, C., Tang, J., Ding, Y., 2010. Aggressive data reduction for damage detection in structural health monitoring. Struct. Health Monit. 9 (1), 59–74.

Pavlopoulou, S., Worden, K., Soutis, C., 2016. Novelty detection and dimension reduction via guided ultrasonic waves: Damage monitoring of scarf repairs in composite laminates. J. Intell. Mater. Syst. Struct. 27 (4), 549–566.

Pearson, K., 1901. LIII. On lines and planes of closest fit to systems of points in space. Lond. Edinb. Dublin Phil. Mag. J. Sci. 2 (11), 559–572.

Pezzotti, N., Lelieveldt, B.P.F., van der Maaten, L., Höllt, T., Eisemann, E., Vilanova, A., 2016. Approximated and user steerable t-SNE for progressive visual analytics. IEEE Trans. Vis. Comput. Graph. 23 (7), 1739–1752.

Pozo, F., Arruga, I., Mujica, L.E., Ruiz, M., Podivilova, E., 2016. Detection of structural changes through principal component analysis and multivariate statistical inference. Struct. Health Monit. 15 (2), 127–142.

Raghavan, A., Cesnik, C.E.S., 2007. Review of guided-wave structural health monitoring. Shock Vib. Dig. 39 (2), 91–116.

Roman-Rangel, E., Marchand-Maillet, S., 2019. Inductive t-SNE via deep learning to visualize multi-label images. Eng. Appl. Artif. Intell. 81, 336–345.

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.

Roweis, S.T., Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290 (5500), 2323–2326.

Staszewski, W.J., Robertson, A.N., 2006. Time-frequency and time-scale analyses for structural health monitoring. Phil. Trans. R. Soc. A 365 (1851), 449–477.

Su, Z., Ye, L., 2009. Identification of Damage Using Lamb Waves: From Fundamentals to Applications. In: Lecture Notes in Applied and Computational Mechanics, Springer-Verlag, London.

Su, Z., Ye, L., Lu, Y., 2006. Guided Lamb waves for identification of damage in composite structures: A review. J. Sound Vib. 295 (3), 753–780.

Tenenbaum, J.B., 1998. Mapping a manifold of perceptual observations. In: Advances in Neural Information Processing Systems. pp. 682–688.

Tenenbaum, J.B., De Silva, V., Langford, J.C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science 290 (5500), 2319–2323.

Tharwat, A., Gaber, T., Ibrahim, A., Hassanien, A.E., 2017. Linear discriminant analysis: A detailed tutorial. AI Commun. 30 (2), 169–190.

Tibaduiza, D.A., Mujica, L.E., Rodellar, J., 2013. Damage classification in structural health monitoring using principal component analysis and self-organizing maps. Struct. Control Health Monit. 20 (10), 1303–1316.

Tibaduiza, D.A., Torres-Arredondo, M.A., Mujica, L.E., Rodellar, J., Fritzen, C.P., 2013. A study of two unsupervised data driven statistical methodologies for detecting and classifying damages in structural health monitoring. Mech. Syst. Signal Process. 41 (1–2), 467–484.

Tibaduiza-Burgos, D.A., Torres-Arredondo, M.A., 2015. Investigation of an expert health monitoring system for aeronautical structures based on pattern recognition and acousto-ultrasonics. Smart Mater. Struct. 24 (8), 085020.

Torgerson, W.S., 1952. Multidimensional scaling: I. Theory and method. Psychometrika 17 (4), 401–419.

Torres-Arredondo, M.A., Tibaduiza, D.A., Mujica, L.E., Rodellar, J., Fritzen, C.P., 2014. Data-driven multivariate algorithms for damage detection and identification: evaluation and comparison. Struct. Health Monit. 13 (1), 19–32.

Van Der Maaten, L., 2014. Accelerating t-SNE using tree-based algorithms. J. Mach. Learn. Res. 15 (1), 3221–3245.

van der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. J. Mach. Learn. Res. 9 (Nov), 2579–2605.

Worden, K., Farrar, C.R., Manson, G., Park, G., 2007. The fundamental axioms of structural health monitoring. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 463 (2082), 1639–1664.