AI based bridge health assessment

Original
AI based bridge health assessment / Marasco, G., Chiaia, B., Ventura, G.. - (2021), pp. 350-364. (9TH INTERNATIONAL WORKSHOP ON RELIABLE ENGINEERING COMPUTING. Risk and Uncertainty in Engineering Computations online May 17-20, 2021).

Availability:
This version is available at: 11583/2900534 since: 2021-05-14T12:41:57Z

Publisher:
Non applicabile

Published
DOI:

AI based bridge health assessment

G. Marasco[1)], B. Chiaia[1)], and G. Ventura[1)]
1)Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Torino 10129, Italy, giulia.marasco@polito.it, bernardino.chiaia@polito.it, giulio.ventura@polito.it

Abstract: Starting from the data extracted from a long-term monitoring system installed on a steel bridge, it has been possible to outline the undamaged behaviour of the structure. The structure under monitoring is a steel suspended arch bridge of long span that has been instrumented with several types of sensors, e.g. triaxial accelerometers, load cells and environmental sensors. The records of the measurements during the first period of structural life and the lack of construction problems ensure the good respect of the structural nominal conditions.

The accelerometric data stored during this period have been used to extrapolate the dynamic characteristics of the bridge: natural frequencies, damping ratios and modal shapes. The use of a specific stochastic subspace technique (SSI-UPCX), allowed to obtain not only the modal parameters but also their uncertainty. In this way, the range of variation of modal parameters, e.g. affected by environmental factors, has been calculated and a minimum and maximum threshold for each parameter has been determined. Consequently, the assessment and control of structural health is updated and linked to these ranges of variation.

In addition, a promising modern approach to tackle the problem is the use of machine learning techniques within the broad field of AI. After the selection/reduction of the parameters that better represent the data, signal detection has been used and the obtained outcomes compared. In the light of both the above approaches, albeit in a different way, it is possible to create a model of the normal operating condition of the structure and consider the deviations from the pattern as an anomaly.

The work represents a first step and a benchmark for the wider damage and ageing identification problem to figure out which method is the most appropriate and effective for this specific case of structural assessment, in terms of effort and accuracy.

Keywords: Structural health monitoring; Modal parameters; Machine learning; Damage detection.

1. Introduction

One of the most crucial problems in the civil engineering field is the identification of structural damage. Accurate results, reduction of false and missing alarms, and detection time (DT) are fundamental aspects of this problem. They have a great weight both on economic and on safety point of view.

Similar to the human body, that during the individual life can be subjected to diseases, the facilities are complex systems that during their useful life are liable to degradation. In both cases, to tackle in a successful way the problem, an interdisciplinary approach is required.

In the infrastructural field, this issue is particular evident for large scale structure subjected to increased loads and different environmental conditions, as the bridges. The presence of some phenomena, e.g. ageing, fatigue and corrosion, implies a change of material properties, structural characteristics, boundary conditions, and element connectivity. This entails the need of checking the structure and developing strategies to optimize the maintenance operations by allocating correctly the available resources. Indeed, early warnings of anomalies are essential for facilitating effective remedying actions. Strategies within this aim belong to structural health monitoring (*SHM*) field and foster a correct development of prognostics and health management (*PHM*).

SHM includes three fundamental approaches. The first consists in continuous monitoring, that allows for verifying alteration of the global structural response. The second includes non- destructive periodic testing (*NDT*), like acoustic emission, ultrasonic, magnetic fields, and radiography that enable a structural local analysis. The last is realized by spot tests. The common target of the three approaches is to highlight early-stage damage and to reduce its spread to avoid the achievement of a critical damage level leading to failure.
Operational modal analysis techniques are one of the most used techniques to address this problem. They belong to the vibration-based methods and allow the extraction of modal structural parameters (natural frequencies, damping ratios and mode shapes) in an inexpensive way and without interfering with the service structural operations. Changes in modal parameters are attributed to damage structural state. For example, a formation of a crack would imply a geometry change with consequent stiffness shifting that could be easily grasped by frequency change. Scour of a bridge pier and loosing of a bolted connection are other types of damage that can be captured by modal parameters alterations.

It is worth pointing out that the data processed to extract the features characterizing the structure, are often related to its normal behaviour, namely to the undamaged structural state. Indeed, the cases in which experimental data are available in damaged cases are very rare (Farrar and Worden, 2012; Diez at al., 2016). On the other hand, the data produced through models (e.g. F.E.M. model), due to high complexity of the systems, to intrinsic inaccuracy of a model, and to the difficulty in defining effective damage cases, are not always reliable for representing the effective structural status. For these reasons, damage diagnosis is often relied on a developed normal condition model.

Unsupervised learning techniques allow to distinguish the damaged from the undamaged structural status avoiding the major problem encountered by machine learning techniques in the field of the *SHM*, consisting in the need of damaged data as supervisor in learning process. These data, in general, are missing or difficult to gather. These techniques are often defined, in the machine learning area, with the term *novelty detection* and their success is linked to the *representativeness* of training dataset (Michau at al., 2019) corresponding to the normal structural conditions. In fact, a deficiency of representativeness will imply a missing distinction between faults and operating conditions (Michau at al., 2018) producing false alarm when new operating conditions, that has not been observed in training period, will classify as faults (Wang at al., 2019). Some approaches are developed to face the low representativeness of many training data sets, due for example to the short observation periods. They are based on the concept of transferring information between systems (units of a fleet) with similar characteristics but that have been under different conditions (Michau and Fink, 2019; Michau at al., 2018). On the other hand, promising approaches have been elaborated in learning the relevant features (Michau at al., 2020). This aspect is essential in the applications where the features *dimensionality*, i.e. the number of parameters monitored and used for damage detection, is very high.
Once used the reference dataset (healthy dataset) to train the chosen learning algorithm, without knowledge about the damaged system conditions (Michau at al., 2020), and designed the decision boundary, an *health indicator (HI)* is frequently used for the damage detection. It consists in the distance between the testing data and the reference data.

One of the criteria to differentiate the novelty detection techniques is the shape of the features distribution in the nominal condition (Farrar and Worden, 2012). For Gaussian distributed normal condition, techniques that involve the discordance calculation between the test data and the normal condition with subsequent comparison between the calculated value and an alarm threshold are often used. In the multivariate case the discordance is calculated by means of Mahalanobis squared-distance and the threshold with the Monte Carlo method. For non-Gaussian normal condition, auto-associative neural network (AANN) can be a valid method to address the problem. Research in this field has produced a lot of improvement. Among these, (Hu at al., 2017) proposed a new approach defined by the terms “Auto-Associative Extreme Learning Machines”. This last looks promising due to its marked learning ability and its low computational cost.

Further approaches involve the use of clustering techniques as the k-means algorithm (Mehdinia at a., 2017; Bounzenad at al., 2019). They, by means of the creation of clusters based on the similarities between the features (Tryon, 1939), are able to distinguish the healthy from the damaged structural states. The distance between the feature vector of each data and the cluster centroid of nominal condition highlights the variation of the structural behaviour in case a damage event occurs. Larger is this distance, more the damage is extended.

This paper presents an AI application in which the k-means algorithm, within the proposed damage identification flowchart, is used for a steel arch highway bridge.

2. Structure under monitoring and sensor network

The structure under monitoring is a highway steel arch bridge in the northern of Italy. Its span is 250 meters. Vertical steel cables have variable number of strands and link the arch to the inferior way. The arch has a trapezoidal section, while the inferior way is composed by a chain hexagonal beam and transverse cantilevers, located along the bridge axis every 8 meters. A view of the bridge is reported in Figure 1 and more details are reported in (Chiaia at al., 2020a).

Figure 1. View of the suspended arch steel viaduct

The bridge is equipped with a customised monitoring network system that, for its capacity in real-time structural assessment, has been defined as “Active Monitoring System”. In a previous paper (Chiaia at al., 2020b) an in-depth illustration of the system has been presented.

Different types of sensors form the monitoring network. There are high-resolution servoinclinometers, steel surface temperature probes, air temperature and humidity sensors, triaxial accelerometers, differential wind pressure transducers, strain gauges at runway cantilevers, load cells at each suspension cable.

3. Feature extraction: Operational Modal Analysis

The signals must be first preprocessed in order to extract their characterizing features.
To this end, accelerometric signals, covering unevenly a period of about one and a half years, have been analysed. As frequently happens in civil structures, output-only techniques exploiting the environmental excitation have been used to avoid the interruption of the infrastructure service. By means the algorithm DD-SSI-UPCX (Data Driven Stochastic Subspace Identification Extended Unweighted Principal Component), present in the ARTeMIS software (Andersen, 2010), an operational modal analysis has been done to extract the features of interest: natural frequencies, damping ratios, and modal shapes. This particular algorithm utilizes a parametric model to fit the raw data. The fundamental assumptions of the SSI methods are infinite amount of data, linear system, and white noise excitation. The algorithm has good performance also in presence of not very large datasets. The validation of the identified structural modes is done by means of the Stabilisation diagram that through more or less restrictive stabilization criteria is able to avoid different types of errors.

The pick of this specific algorithm is supported by several reasons. Among them there are: its capacity to correctly analyse flexible structures (i.e. with low vibration frequencies), its robustness, its calculation speed, and its ability to provide also the uncertainties for modal parameters. This last characteristic allows to comprehend how reliable the obtained results are or not. As known from the literature (Rainieri and Fabbrocino, 2014), the uncertainties related to damping ratios estimation are greater than those related to the natural frequencies. Consequently, for identifying damage, the natural frequencies will be the utilized characteristics.

4. Detection of anomalous behavior

4.1 Traditional approach: thresholds for scattering range

Having available the structural response in time in terms of modal parameters, it has been possible to define a range that encompasses the environmental variability. The first six natural frequencies have been reported in Figure 2. In Table I on page 5, the range of variability of each of them has been defined.
The values, maximum and minimum for each range, represent nominal conditions boundaries.
A departure from these values represents an anomalous behaviour.

The simplicity of this approach hides an important limit. In presence of very small damages, namely in the early stage of the development of an anomalous phenomenon, the extracted features could appear within the nominal defined range due to factors that are different by damage effects (Li at al., 2010). Among them, there are environmental factors (temperature, solar radiation, wind velocity, and humidity), operational factors (intensity of traffic flow and potential traffic jam), and errors due to poor data set and limits of processing techniques. Indeed, the probability density function of the damaged and undamaged state could be overlapped. For the investigated bridge, the variation due to EOVs (environmental and/or operational variations) turns out smaller than that shown for other highway bridges case studies (Ko and Ni, 2005) but

still not negligible. To improve the effectiveness of this approach, techniques of elimination/mitigation of EOVs, like regression modelling and machine learning approaches, should be implemented but it is demonstrated that successful results can be obtained only with a robust data normalization. This last is feasible merely in presence of a large volume (Magalhães at al. ,2012) and a high accuracy of experimental data, which are missing in this case study, so it is not possible to reach valid results following these procedures.
The followed strategy to attenuate the variability of the features has been to widen the observation window of the signals to extract the modal parameters. In detail, the observation window considered includes 31 signals. In this way, a more accurate estimate of modal parameters is possible, and many highly uncertain values located near the boundary of the range are deleted. Moreover, very long records already include a temperature variability that allows to smooth the structural behaviour in the examined temperature range.

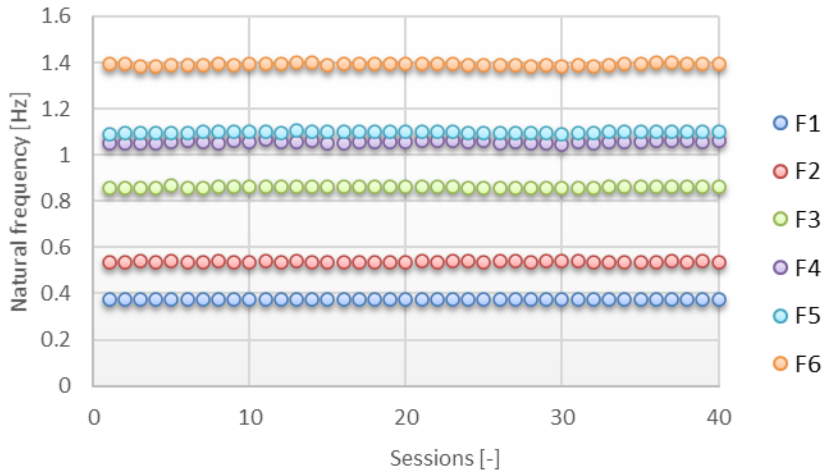

Figure 2. Natural frequency for definition of nominal condition

Table I. Nominal conditions: range of frequency scattering						
VIBRATION MODE	**MODE1:** BENDING OUT OF PLANE	**MODE2:** DOUBLE VERTICAL INFLECTION	**MODE3:** SINGLE VERTICAL INFLECTION	**MODE4:** OUT OF PLANE DOUBLE INFLECTION	**MODE5:** TORSIONAL	**MODE6:** TRIPLE VERTICAL INFLECTION
FREQUENCY	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**
F_MIN [HZ]	0.373	0.532	0.856	1.044	1.091	1.383
F_MAX [HZ]	0.3772	0.540	0.866	1.065	1.105	1.400
VAR REL [%]	1.069	1.398	1.167	1.979	1.239	1.238

The percentage of signals not leading to an alarm, for very small damage (weak signals), can be extremely high. As will be shown in the following, more than half of the damaged signals will not exceed the nominal condition boundaries. This means that 50% of the damaged signals would not be considered as such and the process of damage identification would be significantly less sensitive. Analysing signals, belonging to undamaged and damaged structural states, the percentages of false and missing alarm has been calculated. The False (FA) and Missing Alarm (MA) have been expressed, respectively, as:

$$FA = \frac{N_d}{N_{test\_ud}} \cdot 100$$

$$MA = \frac{N_{ud}}{N_{test\_d}} \cdot 100$$

Where N_d and N_{ud} are the number of data identified as damaged and undamaged, respectively. On the other hand, N_{test_ud} and N_{test_d} are the number of tested undamaged and damaged data. Three potential damaged states have been simulated imposing variations in the acceleration signals.

Many of the most dangerous phenomena for the structural integrity like cracks, foundation settlements, malfunction of a bearing devices, and losses of connections imply change in structural stiffness. A variation in this structural characteristic, occurring locally or in a widespread way, causes a change in the signal frequency. In particular, a stiffness reduction occurs with an ensuing increase of the signal period. Thus, to simulate damaged signals, a delay in the structural response (stretching of time axis) has been imposed.
Analyses carried out on other real bridges (Dilena and Morassi, 2011; Chang, 2016; Magalhães at al., 2012), both with real and simulated damages, have been taken into account to establish the magnitude order of a realistic frequency shift produced by a potential structural damage. Notches, specifically, that can simulate slight damage stemmed from impacts of objects, corrosion or overload, lead to a natural frequency shift of the order on average of 0.1-0.5%.

Initially, the analysed response in natural frequency is that related to the Mode1, Mode2, Mode3 and Mode6, see Table II. These four modes have been considered the most relevant for implementing damage identification strategies.

Table II. Errors for each natural frequency.

DAMAGE LEVELS	DEGREE OF SEVERITY	SIGNAL VARIATION	FREQUENCY	F1	F2	F3	F6
			FA[%]	0.0	12.5	0.0	0.0
LD7	LOW	~0.33%	MA[%]	75.0	97.9	62.5	70.8
LD8	MEDIUM-LOW	~0.66%	MA[%]	41.6	75.0	27.1	39.6
LD9	MEDIUM	~1.0%	MA[%]	2.1	35.4	2.1	8.3

A combination criterion, based on the alarm trigger when one of the features is outside the limits of the ranges defined in Table I on page 5, has been used to improve the effectiveness in damage detection.
As can be noted from Table II, the contribution of the frequency of second mode is the worst both in term of false and missing alarm. The second mode gives a contribution to decrease the missing alarm if a combination of the four frequencies is considered, see Table III.

Table III. Errors for combination F1-F2-F3-F6.

Combination F1-F2-F3-F6	MA [%]	FA [%]	TE [%]
LD7	56.3	12.5	69.0
LD8	14.6	12.5	27.1
LD9	0.0	12.5	12.5

In Table IV, the results in terms of missing, false, and total errors have been reported for a combination of F1, F3, and F6. In this case the total errors decrease, due to the lack of false alarms. Nevertheless, the missing error, in particular way for LD8, increases of about thirty per cent.

Table IV. Errors for combination F1-F3-F6.

Combination F1-F3-F6	**MA [%]**	**FA [%]**	**TE [%]**
LD7	58.3	0.0	58.3
LD8	18.8	0.0	18.8
LD9	0.0	0.0	0.0

For the smallest damage level (LD7), a large part of the monitored frequencies slips back into the nominal condition range. Of course, as can be deduced from the comparison between Table II, III, and IV, the missing alarm percentages decrease if the alarm is emitted when only one of the features is out of the limit values. The reduction of missing alarm percentages is more marked for the damage level LD8.
Concluding, for this structure and for the specific type of damage, the classical approach is able to identify discreetly anomalies produced by a damage level equal and greater than LD8, namely a medium-low damage.

4.2 UNSUPERVISED MACHINE LEARNING: CLUSTERING TECHNIQUE

The traditional approach results deficient for very low damage levels. Nevertheless, it is a complement tool to visual inspection and non-destructive test because it can assess the structural condition even when the damage location is not known a priori and the deteriorated part is not accessible. A specific clustering technique has been used to meet the need to improve the potential structural damages identification. A cluster analysis is applicable when the data have not assigned labels and it consists in researching groups of data, namely clusters, such that the data in a group will be similar (or related) to one another and different from (or unrelated to) the data in other groups. There are three fundamental types of clustering algorithms: K-means and its variants, hierarchical clustering, and density-based clustering. For the problem at hand, K-means algorithm has been chosen and exploited.

4.2.1. *Selection of the features*

Damage-sensitive features are essential to reach a good damage identification. Thus, it appears reasonable to exclude the features that are more sensible to the environmental/operational factors. The contribution of the second, the fourth, and the fifth natural frequency has been considered not relevant for damage identification purpose, as mentioned in subsection 4.1. due to different reasons: the fourth frequency, as underlined by the value in Table I on page 5, has a high relative variation already in the nominal conditions; the fourth and the fifth modes are much closed space modes and therefore are easily confused; finally, Table II on page 6 shows that the second frequency is the least damage-sensitive.

4.2.2. *K-means algorithm: theoretical aspects and SHM perspective*

The flowchart, displayed in Figure 3, describes the steps of the clustering process, that allows to attribute a structural “health/damaged” state to the signals, based on the k-means algorithm.

K-means can be classified as an unsupervised learning algorithm and is a partitional clustering approach. It allows data, represented by a vector of features, to be divided into non-overlapping subsets (clusters). This algorithm makes it possible to visualise similar data in clusters based on a specific metric. Its aim is to create groups of data in the light of the feature vector distances. High homogeneity among the data belonging to the same group and low homogeneity among those belonging to different groups are the criteria to build clusters. Once defined the number of clusters, k centers must be found so to minimize the intra-cluster distance. This algorithm requires considerable calculation effort (Bouzehand, 2019). Several iterative processes have been proposed to converge rapidly to an optimum. One of them (Lloyd, 1982) is composed by three fundamental following steps:

1. Choice of k point centers, namely k centroids.

2. Assign each data/signal, characterize by its features vector, to the closest cluster centroid.

3. Calculate the new k point centers.

The last two steps are repeated as long as convergence is achieved. At each iteration, the k-means algorithm minimizes the intra-cluster distance (Bouzenad at al., 2019). To assess the performance of this algorithm, a matching matrix can be used. There are three factors that can negatively affect the performance. Indeed, K-means performs not well when clusters are of differing sizes, densities, and non-globular shapes. Often, to get over these limitations many clusters are used. Another drawback of this approach is that a different choice of initial centers can imply various created clusters. To overcome this problem, it is good practice to repeat the algorithm several times. Major details related to the k-means theory can be found in (MacKay and Ma Kay, 2003).

From *SHM* point of view, the classical k-means approach allows to analyse collected data/signals and to distinguish between two clusters. One will correspond to healthy and one to damaged state (in the hypothesis of a single damage source). This can provide a "photograph" of the structural state but prevents from tracking structural changing in time and possible evolution of damage from its early stages before damage reaches a critical size (Bouzenad at al., 2019). To intercept initial damage signs, the outlined process in the flowchart in Figure 3 can be considered. The starting point is the creation of an undamaged state, with the Td training data, represented by one cluster (k=1). The centroid (C) and the maximum distance (Dmax) from it are the benchmarks characterizing this nominal state. Then, new signals are analysed one at a time. For each new signal, the counter c is updated and a distance (d) between the new signal and the nominal centroid is calculated. If this quantity is greater than Dmax, the counter (count) is incremented, by revealing an anomaly. When this value (count) is equal to a persistence number (N), the number of cluster k is imposed equal to 2. The features of the undamaged pattern and the persistence number are used to reduce false alarms. Indeed, if the k-means algorithm with a number of cluster greater than 1 were used every time that a new signal was recorded or when a very small number of signals presented d > Dmax, the number of false alarms would be very high, and the data would most likely appear separated in two groups that would not linked to healthy/damaged state rather to warm/cool state. To track the evolution of the damage level, a check of the optimal number of clusters is done when critical time units are reached. The number of signals contained in a critical time unit is defined by Ccritical. The critical time unit is therefore a control unit that, through a discretized calculation of the kopt, allows a speeding up of the analysis process. On the other hand, it corresponds to a time in which initially at most one level damage can be developed. Indeed, in this process, the hypothesis made at the beginning is that in a first critical time unit, the level of potential damage is one and there is no check of the ability of the algorithm to build a model of the underlying structure in the data with a different value of k that results optimal for the problem at hand. The concept that maximum one level of damage can be present in the first critical unit and that then a verification is required (e.g. for c=2*Ccritical, c=3*Ccritical etc) is due to the fact that for the first critical units the appraisal of some factors like materials properties, environmental impact and load turns out simpler and the hypothesis made results realistic. In conclusion in the developed process, the value of k is equal to:

- 1 if count < N
- 2 if (count >= N and Kopt <2)
- Kopt calculated for n*Ccritical if (count >= N, n*Ccritical < c < n+1 Ccritical, and Kopt >2), with n natural number starting from 2

4.2.3. *Application aspects*
The algorithm implementation has been realised in Matlab. The choice of the k initial cluster centroids, corresponding to the point 1 of the list on page 8, is done by means of k-means++ algorithm, that optimizes the running time and the final solution (Arthur and Vassilvitskii, 2007). The points 2 and 3 of the just mentioned list correspond to so-called *batch updates phase* aims at minimizing the sum of point-to-centroid distances. This phase could not converge to the correct solution and it is sometimes followed by so-called *online updates phase*. For the problem at hand, the *batch updates phase* turns out sufficient and thus it is the only performed phase. Thereby, the process is very fast. Having the possibility to choose among several distance metric, it should be noted that Squared Euclidean distance has been used. Besides, as maximum number of iterations to reach the convergence has been used the default value in Matlab (100). As regard the empty clusters, they have been removed but the algorithm has been set to keep track of their presence. As underlined in the previous subsubsection, it is good practice to repeat the algorithm several times with different initial centers so to maximize the inter-cluster distance. In this case, the argument 'Replicates' has been set equal to 10. Another important aspect that must not be forgotten, is the data normalization (Mohamad and Usman, 2013). This is an important pre-processing task, that scaling data in a specific range is able to ensure the same weight to every attribute. For this specific problem, a Min-Max normalization has been applied. In addition to the aspects just analysed, k-means clustering is also very sensitive to outliers because their presence can skew the right positions of the k point centers (Alamdari at al., 2017). In the present case, the removal of the outliers does not lead to benefits.

4.2.4.1. *Learning of undamaged pattern.* The number of clusters is imposed equal to 1 in the nominal condition. The localization of the normalized centroid and the maximum normalized distance (Dmax) are reported in Table V.

Table V. Normalised centroid and maximum distance from it for nominal conditions

M1	M3	M6	Dmax
0.518	0.420	0.470	0.588

4.2.4.2. *Damaged pattern.* The number of cluster k is updated to the value of two when a number N of signals has a distance from the normal condition centroid higher than the defined threshold (Dmax).
Furthermore, some criteria may be considered to evaluate if there are structural changes, which in case of early-stage damage recognition can indicate grow of negative phenomena in time. They allow to evaluate the optimal number of clusters (Kopt), verifying thus the existence of a potential development of the damage. If it is growing, an increase in the damage level can be deduced. Among the criteria that can be found in the literature there are the Calinski-Harabasz clustering criterion (Calinski and Harabasz, 1974), the Silhouette index (Rousseeuw, 1987; Kaufman and Rousseeuw, 1990), the Davies-Bouldin criterion (Davies, 1979) and the gap statistic criterion (Tibshirani at al., 2001). The Calinski-Harabasz relies upon the sums of squared Euclidean distance between the feature vectors and the centroids of the predicted clusters and the optimal value of k corresponds to its maximum value. The Silhouette index evaluates the difference between inter-cluster distances and intra-clusters distance and the optimum value of k must maximize this index. The

Davies-Bouldin criterion utilizes the intra and the inter-clusters distance as well and, for this criterion, the optimal value of k corresponds to its minimum. Finally, the gap criterion calculates the logarithmic mean of the pairwise distance. In this case the optimum value of k is correlated to the maximization of this criterion. In this way, there will not be the need to choose thresholds that will define damage levels. The algorithm, on the basis of available data, will point out by means of calculated Kopt, a potential worsening of the situation.

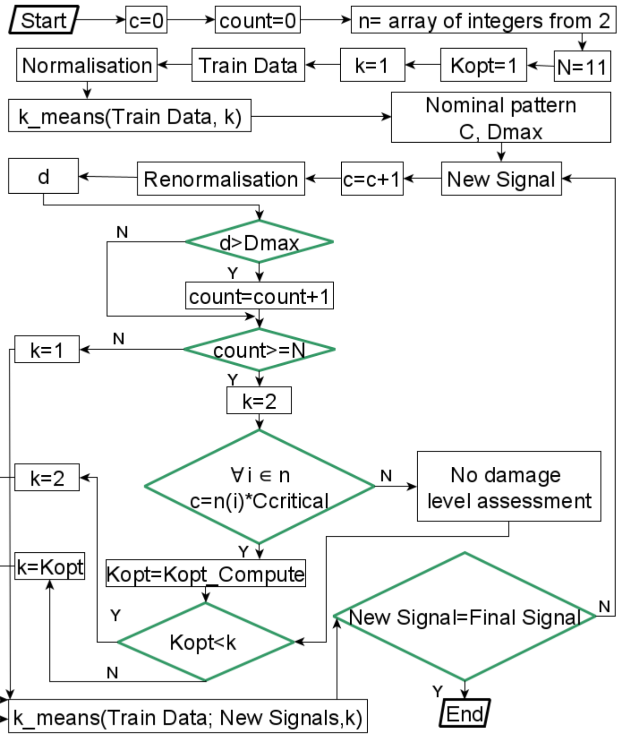

Figure 3. Proposed flowchart for damage identification based on k-means algorithm.

4.2.4.3. *Evaluation of the accuracy: the matching matrix.* The matching matrix has been used to assess the accuracy of the clustering process and to understand if there are improvements with respect to the traditional approach. The comparison between real and predicted clustered signals gives an idea about the capacity of the algorithm in correctly identifying the structural condition. The rows and the column of the matching matrix correspond to the real and to the predicted classes of the signal, respectively.
A combination of test undamaged dataset (Test_UD) and datasets corresponding to the different levels of damage, described in subsection 4.1, are used. The total error (TE) has been defined as:

$$TE\ [\%] = MA[\%] + FA[\%]$$

Where MA and FA correspond to missing and false alarms defined in the aforementioned subsection.

For all the three cases, the number of signals is less than the hypothesized Critical threshold, so the number of used clusters is equal to 2. The following figures 4, 5, and 6 display the division of the data (signals) for each analysed Dataset. Cluster 1 includes the signals belonging to the undamaged condition while Cluster 2 contains the ones of the damaged condition. Of course, both are associated with a centroid.

– Dataset 1: Test_UD and LD7

Table VI. Matching matrix Datasets 1			
Real Predicted	Healthy	Damaged	TE
Healthy	68.8%	31.2% (FA)	60.3%
Damage	29.1% (MA)	70.1%	

As can be seen in Table VI, the total error is similar (variation of about 3%) to the one obtained by the traditional approach. The error induced by missed alarm is reduced by half at the expense of the error of false alarm. Thus, although the error level is still high, the clustering technique results on the safe side.

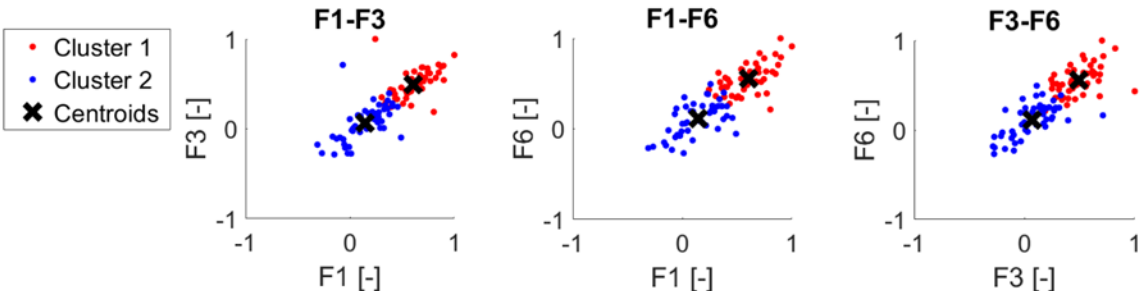

Figure 4. Clustered data (Dataset 1)

– Dataset 2: Test_UD and LD8

Table VII. Matching matrix Datasets 2			
Real Predicted	Healthy	Damaged	TE
Healthy	77.1%	22.9% (FA)	27.0%
Damage	4.1% (MA)	95.9%	

From the comparison between Table VI and Table VII, it is clear that the total error is drastically decreased in the passage from Dataset 1 to Dataset 2. Compared to the traditional approach, the total error is increased (about 50%) but it is worth to stress that the error due to missed alarm has decreased more than four times. In this case, this approach results more expensive than the traditional approach, due to the fact that about 1/5 of the undamaged signals have been indicated as damaged. On the other hand, the drastic reduction of missing alarm makes this technique on the safe side.

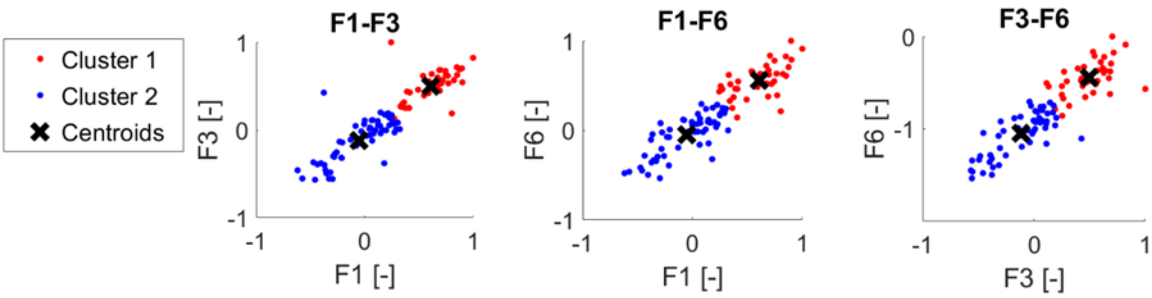

Figure 5 Clustered data (Dataset 2)

– Dataset 3: Test_UD and LD9

Table VIII. Matching matrix Datasets 3			
Real Predicted	Healthy	Damaged	TE
Healthy	100.0%	0.0% (FA)	2.1%
Damage	2.1% (MA)	97.9%	

In this case, the cluster approach shows a very good performance but slightly lower than that of the traditional approach.

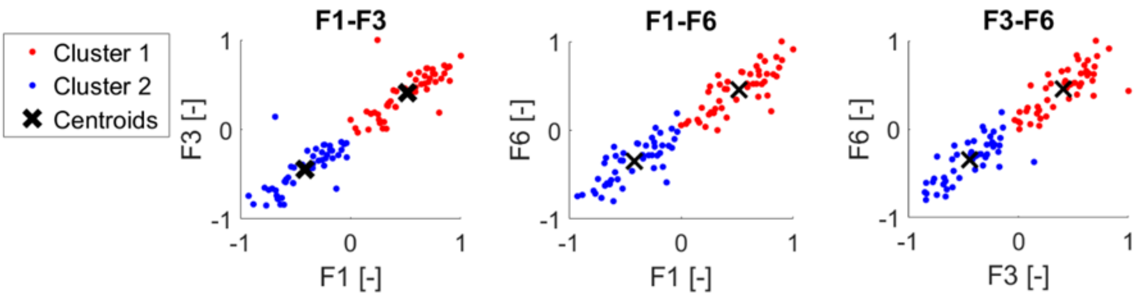

Figure 6 Clustered data (Dataset 3)

Better results in terms of missing alarms have been achieved for all the three damage levels if as number of clusters the maximum value, among the optimal value calculated by means of the three criteria, is used. The smallest damage level LD7 shows the most marked absolute improvement. Indeed, for LD7 the percentage of errors for missing alarms goes down to the value of 6.25%. There is a reduction of about four times compared to the previous case. The obtained total error is less than the previous, showing a value of 52.05%. Of course, the error for false alarms increases. However, this growth is much lower compared to the decrease of the missing alarms errors. From these considerations it can be deduced that, if the structure has been just built and so the first hypothesized damage level is very small (LD7 - LD8), the calculation and the use of kopt would be appropriate also if there are a number of signals less than 2*Ccritical. It could be greater than 2 dues to the presence of environmental effects.

Indeed, a very strong similarity could be present between the undamaged states measured at high temperatures and damaged states measured at low temperatures. On the other hand, in cases in which a greater level of damage is immediately present (LD9) the use of the maximum value of k implies worst performance of the algorithm due to the fact that a small gain in terms of missing alarms and a huge increase of false alarms occurs. Concluding, the k-means algorithm shows promising results in particular for very small damage levels. Its utilization implies a marked decrease of the error due to missing alarm. It produces acceptable outcomes from the safety point of view already for the damage level LD7. It avoids several missing alarms, paying the price of increasing the false alarms. For the early-stage damage levels it is preferable to the traditional approach as it ensures higher safety.

4.2.4.4. *Recognition of increasing damage.* When the critical unit of time is reached, the optimal value of k is calculated. A higher value of k, in the succession of control units, could point out an increase in the level of structural damage. By utilizing the three criteria for the calculation of the optimal value of k, in the analysis of the data covering all the three damage levels (LD7-LD8-LD9), it is possible to verify the capacity of recognition of an increasing damage. The results for the first three time units are displayed in Figure 7, 8 and 9, respectively. A change in the optimum value of k is displayed for all the three units by the Gap criterion and for the last two units by the Calinski-Harabasz criterion.

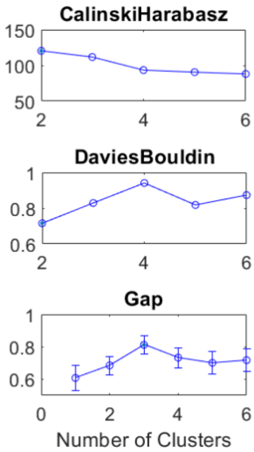

Figure 7. First unit

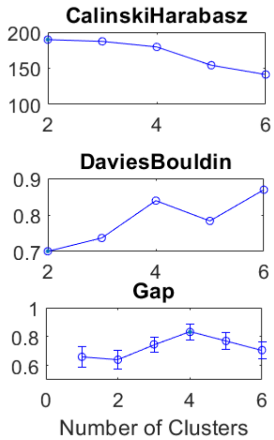

Figure 8. Second unit

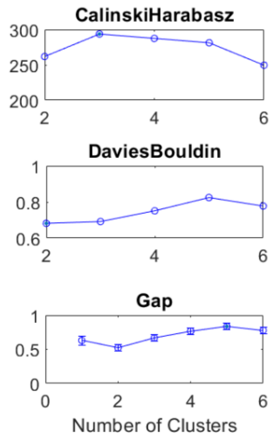

Figure 9. Third unit

5. Conclusions

This paper has illustrated two different approaches to address the bridge health assessment. Both the methods start from the elaboration of the acceleration data by means of the SSI-UPCX method that provides not only modal parameters but also their uncertainty. This additional information allows an increased awareness of the reliability of the modal parameters. The natural frequencies, due to their very low uncertainty, have been selected for the damage identification goal. The analysis aims to compare the performance of the two approaches using a test sample incorporating variability induced by environmental effects.

The main findings, observations, and conclusions stemmed from this survey are summarised as follows:

- The traditional approach shows relevant errors for very small damage levels. Such errors are solely due to the worst part of the total error, namely the missing alarms.
- For the small damage levels analysed, the clustering technique has a better performance. It greatly reduces the error due to missing alarms but increases the one due to false alarms. From the engineers' point of view, this swap turns out undoubtedly positive. It is safer to correct a false positive by means of inspections than not to grasp the presence of damage.
- For very small damage level, an increased computational effort in the calculation of the optimum value of numbers of clusters in the application of the k-means algorithm produces a further improvement as regards missing alarm errors. Thus, this clustering approach is preferable if a high level of safety must be ensured. Another positive aspect of this method is the possibility to observe the evolution of the optimal number of clusters for assessing the increase of damage. Indeed, it is possible to deduce a changing in the data and a potential growth of damage severity from a growing value of optimal k over time.
- A greater uniformity, completeness, and quality of data would imply an improvement of the already promising performance of the clustering algorithm. The role of the persistence number (N) and the optimal number of clusters (Kopt) is decisive in mitigating the expected increase in error, in terms of FA and MA, in critical scenarios for the k-means algorithm (clusters with sizes and densities markedly different).
- In the application of the clustering algorithm, the uncertainty in the choice of the optimum value of clusters due to the not-perfect correspondence among all the values indicated by various criterions is an aspect that must be improved. Certainly this, and the consequent not-perfect correspondence between each optimal value and the real number of data groups, suggest the research of more performant features as a perspective of this work. More effective features would imply a betterment for both the investigated approaches. In particular, with higher quality of the signals, the Spectral Moments (SMs) could be exploited as damage features. In (Alamdari at al., 2017) the high ability of these features, covering the whole frequency range, in detecting subtle differences between normal and distorted signals have been evidenced. Moreover, the use of indicators for modal shapes like MAC and COMAC, could be accounted due to the fact that they are the less sensitive to other factors. Then again, the combination of several parameters could lead to better damage-sensitive features.

References

Alamdari, M. M., T. Rakotoarivelo, and N. L. D. Khoa. A spectral-based clustering for structural health monitoring of the Sydney Harbour Bridge. Mechanical Systems and Signal Processing, 87: 384-400, 2017.

Andersen, P. ARTeMIS Extractor Online Help. Structural Vibration Solutions A/S, 2010.

Arthur, D. and S. Vassilvitskii. K-means++: The Advantages of Careful Seeding. In SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2007.

Bouzenad, A. E., M. El Mountassir, S. Yaacoubi, F. Dahmene, M. Koabaz, L. Buchheit, and W. Ke. A Semi-Supervised Based K-Means Algorithm for Optimal Guided Waves Structural Health Monitoring: A Case Study. Inventions, 4(1): 17, 2019.

Caliński, T. and J. Harabasz. A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1): 1-27, 1974.

Chang, K. C. and C.W Kim. Modal-parameter identification and vibration-based damage detection of a damaged steel truss bridge. Engineering Structures, 122: 156-173, 2016.

Chiaia, B., G. Marasco, G. Ventura, and C. Zannini Quirini. Customised active monitoring system for structural control and maintenance optimisation. J Civil Struct Health Monit, 10: 267- 282, 2020a.

Chiaia, B. G. Ventura, C. Zannini Quirini, and G. Marasco. Bridge active monitoring for maintenance and structural safety. In Arêde A, Costa C (eds) Proceedings of ARCH 2019. ARCH 2019. Structural integrity, vol 11. Springer, Cham, 866–873, 2020b.

Davies, D.L. and D.W. Bouldin. A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, 2: 224-227, 1979.

Diez, A., N. L. D. Khoa, M.M. Alamdari, Y. Wang, F. Chen, and P. Runcie. A clustering approach for structural health monitoring on bridges. Journal of Civil Structural Health Monitoring, 6(3): 429-445, 2016.

Dilena, M. and A. Morassi. Dynamic testing of a damaged bridge. Mechanical Systems and Signal Processing, 25(5): 1485-1507, 2011.

Farrar, C. R. and K. Worden. Structural health monitoring: a machine learning perspective. John Wiley & Sons, 2012.

Hu, Y., T. Palmé, and O. Fink. Fault detection based on signal reconstruction with Auto-Associative Extreme Learning Machines. Engineering applications of artificial intelligence, 57:105-117, 2017.

Kaufman, L. and P.J. Rousseeuw. Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, 2009.

Ko, J. M. and Y. Q. Ni. Technology developments in structural health monitoring of large-scale bridges. Engineering structures, 27(12): 1715-1725, 2005.

Li, H., S.Li, J. Ou, and H. Li. Modal identification of bridges under varying environmental conditions: temperature and wind effects. Structural Control and Health Monitoring 17(5): 495-512, 2010.

Lloyd, S. Least squares quantization in PCM. IEEE Transactions on information theory, 28(2): 129-137,1982.

Mackay, D.J. and D.J. Mac Kay. Information theory, inference and learning algorithms. Cambridge university press, 2003.

Magalhães, F., A. Cunha, and E. Caetano. Vibration based structural health monitoring of an arch bridge: from automated OMA to damage detection. Mechanical Systems and Signal Processing, 28: 212-228, 2012.

Mehdinia, S., Y. Hamishebaharb, M. Bitarafc, A. Kayvan, and M.R. Ghorbani-Tanhad. Structural damage detection using K-means Clus-tering algorithm under Unknown Environmental and Operational Conditions.

Michau, G., T. Palmé, and O. Fink. Fleet PHM for critical systems: bi-level deep learning approach for fault detection. In Proceedings of the European Conference of the PHM Society, 2018.

Michau, G. and O. Fink. Unsupervised Fault Detection in Varying Operating Conditions. In 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), 1-10, IEEE,2019.

Michau, G., Y. Hu, T. Palmé, and O. Fink. Feature learning for fault detection in high-dimensional condition monitoring signals. In Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2020.

Mohamad, I.B. and D. Usman. Standardization and its effects on K-means clustering algorithm. Research Journal of Applied Sciences, Engineering and Technology, 6(17): 3299-3303, 2013.

Rainieri, C. and G. Fabbrocino, Influence of model order and number of block rows on accuracy and precision of modal parameter estimates in stochastic subspace identification. Int J Lifecycle Perform Eng 1(4):317-334, 2014.

Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20: 53-65, 1987.

Tibshirani, R., G. WALTHER, and T. Hastie. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2): 411-423, 2001.

Tryon, R.C. Cluster analysis: correlation profile and orthometric analysis for the isolation of unities in mind and personality. Ann Arbor: Edward Brothers, 1939.

Wang, Q., G. Michau, and O. Fink Domain adaptive transfer learning for fault diagnosis. In 2019 Prognostics and System Health Management Conference (PHM-Paris). IEEE, 279-285, 2019.