Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 44 (2023) 1546–1553

Structural Integrity Procedia

www.elsevier.com/locate/procedia

XIX ANIDIS Conference, Seismic Engineering in Italy

Impact of the model error on the neural network-based damage detection

Federico Ponsi[a], Giorgia Ghirelli[b], Elisa Bassoli[b], Loris Vincenzi[b,*]

[a]*University of Bologna, Department of Civil, Chemical, Environmental, and Materials Engineering, Viale Risorgimento 2, Bologna 40126, Italy*
[b]*University of Modena and Reggio Emilia, Department of Engineering Enzo Ferrari, Via P. Vivarelli 10, Modena 41125, Italy*

Abstract

Machine Learning (ML) techniques applied to vibration-based damage detection of structures showed promising results in identifying damage for their capability of feature discrimination even in presence of noise-corrupted data. In this context, the paper presents the application of a damage detection procedure based on neural networks to a railway bridge and proposes a procedure to take into account for the model error in the training phase. The output of the network, that depends on the dynamic features given as input, allows to classify the structure state as undamaged, lightly damaged or severely damaged. The procedure employs only simulated data but includes a series of expedients to approach a real situation, like the stochastic modelling of measurement errors and the use of two different models to account for the model error. The performances of the networks are analyzed with respect to datasets generated by the two different models. The results show the relevance of accounting for model error in the calibration of the network to obtain a robust damage identification.

Keywords: Damage detection; Data-driven methods; Artificial Neural networks; Principal Component Analysis; Model error.

1. Introduction

Numerous structures and infrastructures are in a state of damage or degradation caused by several factors, like seismic events, inappropriate human actions or the protracted exposure to adverse environmental conditions. Vibration-based damage detection of structures is increasingly widespread thanks to its ability to point out possible

* Corresponding author. Tel.: +39-059-2056213; fax: +39-059-2056126.
E-mail address: loris.vincenzi@unimore.it

2452-3216
10.1016/j.prostr.2023.01.198

anomalies with minimal invasiveness (Doebling et al., 1996). Modal properties are often used as representative features of the structural health state, since they provide information about both the global (for instance natural frequencies) and the local (for instance mode shapes or modal curvatures) behavior of the structure (Doebling et al. 1996, Sohn et al. 2002, Comanducci et al. 2016).

Machine Learning (ML) techniques are suitable and promising approaches to this problem since they are capable of working with uncertain and noise-corrupted data (Khan and Yairi, 2018). The application of these techniques to the field of damage detection is quite recent, but a vast amount of works has been produced. Review works that aim at the classification of ML-based damage detection techniques are those of Avci et al. (2021) and Hou and Xia (2021). Numerical models are not directly involved in the identification of a damaged state if data-driven methods are used, but they may be employed in preliminary phases, to evaluate effects of structural damage that cannot be reproduced in a real situation.

The objective of the present work is to define a complete procedure for damage detection that exploits ML techniques, in particular Artificial Neural Networks (ANN). The procedure has been developed with only simulated data but includes a series of expedients to approach a real situation, like the stochastic modelling of measurement errors and the use of two different models to account for the model error. The latter is the discrepancy between the response obtained from the numerical model and the experimental results. In the paper, the response of an accurate model (called the "reference" model) is considered in place of the experimental data while a simpler model (called the support model) is used to train the network. In this way, the authors want to take into consideration the fact that the simulated data never exactly reproduce the reference results even if the numerical model is well calibrated with respect to the experimental data. Moreover, different networks are studied, each of one takes as input modal properties that have been elaborated in different ways. The procedure is applied to the case study of a railway bridge and the performances of the networks are analyzed with respect to datasets generated by two different models, to assess the effective applicability of the presented procedure.

2. Multi-layer perceptron for classification

Multi-layer perceptron (MLP) is the most popular kind of ANNs that applies supervised training. It is composed of neurons arranged into layers. Each neuron in a given layer is connected to all the neurons of the following layer. The connections between neurons do not form cycles, therefore the information elaborated by the system moves only in the forward direction, from the input layer to the output one (Haykin, 1999). In general terms, the connection among the outputs of the neurons belonging to the j-th layer and the output of the i-th neuron belonging to the j+1-th layer is composed of a weighted sum and the contribution of the so-called transfer function that introduces non-linearity in the process. Since the network is employed for classification, a common choice for the transfer function of the output layer is the soft-max function, while the transfer function of the hidden layers is generally chosen among the well-known sigmoid logistic function, hyperbolic tangent function and rectified linear unit (ReLU) function. The definition of these functions can be found in Bishop (2006).

The key aspect for a MLP is its training, that is the process where the network coefficients are tuned in order to increase the ability of the network to make correct predictions on the basis of the available data. The performance of a network in classification problems is quantified by the average cross-entropy loss function, that measures the discrepancy between the prediction vectors $\mathbf{s}_n$ and the corresponding targets $\mathbf{t}_n$ related to the training set (Bishop 2006). The optimization of the network coefficients can be performed with several algorithms, usually gradient-based methods. In this work, the authors have chosen the scaled conjugate gradient back-propagation algorithm proposed by Moeller (1993), known for its efficiency in problems with a large number of neurons. Finally, the data over-fitting is avoided adopting a sufficiently large number of data and a limited number of layers and neurons (Ying 2019).

3. The proposed damage detection procedure

The proposed procedure involves the use of ANNs for the identification of a possible damaged state of a structure. The localization and the quantification of a damaged state are not investigated. On the basis of modal-based features, the trained ANN is able to classify the structure in an undamaged condition, in a lightly damaged condition or in a severely damaged condition. The role played by data in this procedure is crucial for its success. In particular, the collection of data related to different damaged conditions of the structure has to be carried out. Because the execution of damage tests over a full-scale bridge involves feasibility problems and economical drawbacks, in this work the damage scenarios of the bridge are simulated through a Finite Element (FE) model of the bridge (in the following called support model or model S). The support model must reproduce as better as possible the experimental data that are acquired by the monitoring system on the structure. It is worth noting that the simulated data never exactly reproduce the structural behavior even if the numerical model is well calibrated with respect to the experimental data. The difference between the experimental response and the numerical prediction is the model error. In this work, to calibrate the network and to take into account to the model error, the experimental data are replaced by the outcomes of an accurate numerical model (in the following the "reference" model or model R). On the basis of the modal properties of the reference model, the support model is calibrated and used to extract the modal features to train the neural network. Finally, it is in the interest of the authors to assess the performances of the ANN in presence of noise-corrupted data and noise-filtered data. For this reason, a white noise process is added to the structural response of both models. In summary, to calibrate the network and to test the effectiveness of the Neural Network for damage detection, the fundamental steps of the proposed procedure are listed in the following:

- select the case study, built the reference model and the support model;
- simulate the structural response of the reference model for the undamaged condition and add the noise to the structural response;
- calibrate the support model to obtain modal properties as close as possible to those of the reference model;
- generate the datasets for network training and test from both models;
- apply a noise filtering of the model response with Principal Component Analysis (PCA);
- perform the network training and optimization;
- test the network with data simulated by the support model and by the reference model to assess the effect of model and measurements errors.

All the previous aspects are treated in more detail in the following.

3.1. Case study and numerical models

The case study is a railway girder bridge 40 m long and 4.3 m wide. Two simply supported steel girders support a concrete slab 0.34 m thick. The slab is connected with the girder top flanges through pegs to prevent the slip between the steel girders and the slab. The two steel girders are connected to each other by a three-dimensional bracing system. A detailed FE model (model R, see Fig. 1) is developed using the FE software MIDAS CIVIL and its response replaces the experimental data. It will be employed for the test of the networks in the last phase of the procedure. Flanges and webs of the steel girders and the concrete slab are modeled with shell elements, while beam and truss elements are used for the bracing system. A simpler model (model S) has been developed for the generation of the network training dataset. The model S is a simply supported beam with 100 finite elements characterized by an equivalent rectangular cross section. Each beam element has both flexural and shear deformability. The properties of the cross section are defined through the calibration procedure described below.

Natural frequencies and mode shapes of both FE models are obtained performing modal analysis and modifying the exact values of the structural response by adding noise with the aim to reproduce measurement errors and uncertainties characterizing the modal identification procedure. A dynamic monitoring system is assumed to be installed on the bridge. The measurement equipment is supposed to be composed of five accelerometers connected to the structure and placed along the bridge length at the locations indicated in Table 1. Therefore, the mode shape components are available only for the points corresponding to the sensor locations. A Gaussian noise with a coefficient

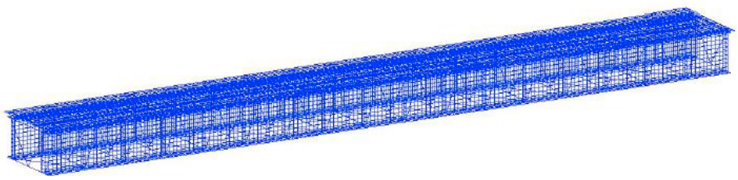

Fig. 1. Isometric view of the reference model (model R).

Table 1. Name and location of the sensors of the hypothetic monitoring system.

Name	A1	A2	A3	A4	A5
Location [m]	6	13	20	27	34

Table 2. Comparison between natural frequencies of model S and R after calibration.

	1st fr. [Hz]	2nd fr. [Hz]	3rd fr. [Hz]	4th fr. [Hz]
Model S	1.680	6.102	12.079	18.713
Model R	1.697	5.859	12.958	18.193
Difference	0.017	-0.243	0.879	-0.519

of variation depending on the property type is added to the exact values computed by the models. The coefficient of variation is set to 0.01 for frequencies and 0.05 for mode shapes. Moreover, to introduce non-negligible uncertainties also for near-zero components of mode shapes, a further specific source of error is introduced by adding a noise extracted from a uniform distribution defined in the interval $[-4\cdot10^{-3}, +4\cdot10^{-3}]$.

The values of elastic modulus and shear modulus of model S are calibrated with respect to the response of model R. In particular, these parameters are chosen with the aim to minimize the difference between the exact natural frequencies of models S and R. Only the first four vertical (flexural) frequencies are considered. Mode shapes have not been included in the calibration since they are not sensitive to the modification of the chosen structural parameters. Natural frequencies of the model S and R are listed in Table 2. It is worth noting that even if the reference model (model R) is not really complex, a residual discrepancy remains also after the calibration due to the different modeling strategies adopted. The same can be stated for mode shapes, as shown in Fig. 2, where the comparison for mode 2 and 3 is presented.

3.2. Damage scenarios

Damage is simulated in both models through the reduction of the elastic modulus of one or more finite elements. Considering the elastic modulus E_u of an undamaged element and the reduced elastic modulus E_d of a damaged element, it is possible to define the damage severity r as:

$$r = 100\left(E_u - E_d\right)/E_u \tag{1}$$

As concerns model S, several damage scenarios are created by varying damage severity and location of a single damaged element, that may assume different values. In detail, damage location varies with a step-size of 5% of the bridge length over the whole structure, while its severity varies from 0 to 40% with a step-size of 2.5%. The structure condition is considered as lightly damaged if the severity of damage is lower than 15%, otherwise it is considered as severely damaged. Considering the stochastic modeling of measurement errors, 200 samples compose the dataset for

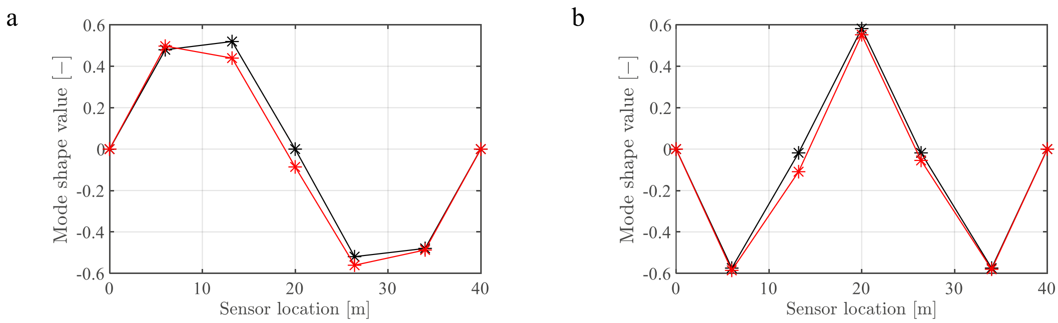

Fig. 2. Comparison between mode shapes computed by model S (Black asterisks) and model R (red asterisks); (a) mode no. 2 and (b) mode no.4.

Table 3. Description of the damage scenarios of model R.

Code	Damaged part	Location along the bridge length	Extension
S1	Undamaged condition	-	-
S2	Steel flange	One fourth	Element strip
S3	Steel flange	Middle	Element strip
S4	Steel flange	Three fourth	Element strip
S5	Concrete slab	One fourth	Element strip
S6	Concrete slab	Middle	Element strip
S7	Concrete slab	One third	Nearly semi-circular area

a specific scenario. The single sample is obtained by adding the randomly extracted values of noise to the exact value of the modal response. Finally, 7 additional scenarios are created for model R. The first represents the undamaged state of the model, while the remaining six are damage scenarios (see Table 3). Damage in the steel beam (with reference to scenarios S2, S3 and S4) is introduced by decreasing the elastic modulus of an element row of the bottom flange to a quasi-zero value, aiming at simulating material discontinuity. For scenarios involving damage of the slab, i.e. S5, S6 and S7, the elastic modulus of concrete is reduced by 50 % simulating a cracked condition.

3.3. Network definition and optimization

This section describes the networks employed in the procedure, with reference to the construction of the input vector of the MLP. The networks and their corresponding identifiers are listed in Table 4 together with a brief description of the input features. All these networks have an input vector composed of the natural frequencies followed by the mode shape components. Only the components corresponding to the five sensors of the monitoring system (A1-A5 in Table 1) are considered. Network N1 takes as input noise-corrupted modal properties, as described in section 3.1. For network N2 and N3, a noise filtering is performed through the PCA.

The PCA is a well-known technique of multivariate statistics that is usually employed for data compression or noise filtering. PCA involves the eigenvalue analysis of the covariance matrix of the normalized original data in order to identify the principal directions of data variance (also known as principal components, collected in the eigenvector matrix $\mathbf{L}$) and the amount of variance associated to each direction, represented by the eigenvalues. At this step, it is possible to discard the components of lower significance on the basis of the relative variance. These components can represent the noise of the data. The reduced eigenvector matrix $\mathbf{L}_p$, obtained discarding the previous mentioned components, is called loading matrix and it allows the representation of data in a new coordinate system or the reconstruction of the data in the original system after denoising. The denoised data matrix $\mathbf{Z}_p$, that is of interest for the following application, is evaluated as follows:

$$\mathbf{Z}_p = \left(\mathbf{L}_p \mathbf{L}_p^T\right)\mathbf{Z} \tag{2}$$

where $\mathbf{Z}$ is the matrix containing the normalized original data.

A baseline set of frequencies and mode shapes is created considering only data referred to the undamaged state. Then, PCA is separately applied to the frequency and mode shape sets in order to compute the loading matrices $\mathbf{L}_{p,f}$ and $\mathbf{L}_{p,\varphi}$. At the end, data are reconstructed in the original coordinate system using Eq. (2).

The core of the problem lies in the choice of the number of principal components to retain. As for the frequency set, each of the four principal component describes about 25% of the total variance. Consequently, all the four principal components are retained and the original data are not filtered. As concerns mode shapes, the cumulative percentage of variance described by retaining 14, 15 or 16 principal components is 91.5%, 96.0% and 99.9%, respectively. Since the value of 99.9% is high and is presumable that it includes the variability due to noise, only the cases of 14 or 15 principal components are selected. In the first case (14 components) the reconstructed data form the input vector of network N2, while 15 components are selected for the network N3.

Table 4. Studied networks and their optimal architecture.

Network identifier	Damage features	Transfer function	Layer size (number of neurons for each layer)
N1	Noised modal properties	Hyperbolic tangent	[24, 20]
N2	Denoised modal properties (14 comp.)	Hyperbolic tangent	[10, 10]
N3	Denoised modal properties (15 comp.)	Hyperbolic tangent	[10, 10]

The optimization of the network architecture is performed with the aim to tune a series of hyper-parameters so that the performance of a network is improved as much as possible. The hyper-parameters considered are the number of hidden layers, their size and the type of transfer function used. The hidden layers can be one or two, their size may vary between the dimension of the output vector (in the present case 3) and the dimension of the input vector (equal to 24). The possible choices for the transfer function have been introduced in section 2. They are sigmoid logistic function, hyperbolic tangent function and rectified linear unit (ReLU) function. In this work, the network optimization has been conducted with a trial-and-error strategy saving the architectures that resulted in better performances at each trial. With the adopted strategy, the obtained optimal number of hidden layers is two for all the network (N1, N2 and N3) but the optimal number of neurons depends on the presence of noise on data. The number of neurons of each hidden layer is summed up in Table 4.

4. Results

4.1. Training and test with the dataset of model S

Table 5 presents the performances of the networks in terms of accuracy and percentage of uncertain predictions. The accuracy is a measure of the errors that a network commits and it derives from the comparison between network predictions and the associated target class. The accuracy of each network is calculated with reference to two different criteria. In the first case, the exact correspondence between the predicted output and the target is considered. It is denoted in Table 5 as the "strict accuracy". In the second criterion, a tolerance of 5% to the damage severity boundaries reported in section 3.2 for the three classes is applied. If the classification output is not correct but the damage severity is included in the tolerance range, the mistake made is considered slight. Therefore, the "soft accuracy" reported in Table 5 takes into account only errors that are out of the tolerance range. The percentage of uncertain predictions is computed considering the probability expressed by the output layer of the MLP. Let $y_{\mathrm{ord}} = [p_1, p_2, p_3]$ be the output vector of the network with probabilities ordered in descending order (i.e. $p_1 > p_2 > p_3$) associated to each class.
If a small gap is found between the probability p_1 and p_2, the damage class is defined with higher uncertainty with respect to the case $p_1 >> p_2$. For this reason, a threshold value η equal to 0.33 is defined, which discriminates certain predictions from uncertain ones. When $p_1 - p_2 < \eta$ the result is considered uncertain.

Results reported in Table 5 show very small differences between the same quality indices (strict accuracy, soft accuracy or percentage of uncertain predictions) related to the training and test dataset for the same network. This excludes over-fitting of data for all the networks. The performance of network N1 is the worst, both in terms of accuracy and uncertain predictions. This is due to the fact that the network N1 takes as input noise-corrupted modal properties and the effect of noise covers the modification of modal properties caused by damage.

The application of PCA is successful to filter the noise. For network N2 and N3, strict and soft accuracy are larger than 90 % and the percentage of uncertain predictions is lower than 10 %. The choice between 14 or 15 principal components for the construction of the loading matrix does not produce substantial differences in this phase.

4.2. Test with the dataset of model R

Test with model R is aimed to investigate the effect of the model error on the accuracy of networks N1, N2 and N3. First, network predictions using the exact values of modal properties computed by model R are considered (thus without introducing measurement errors). Table 6 contains the prediction results of the 7 scenarios listed in Table 3. All networks identify all scenarios as damaged, including the scenario S1 that corresponds to the undamaged state of the detailed model R. The reason of these bad results has clearly to be attributed to the model error. Although model S has been calibrated based on the response of the model R, residual errors still remain both for natural frequencies (Table 2) and mode shapes, as highlighted in Fig. 2 for mode 2 and 3.

To avoid that the network identifies as damaged any structural condition due to the model error, a proposal is here presented. The modal properties computed by the reference model (model R or, in real case studies, the experimental modal properties) must be adjusted by adding the residual error obtained at the end of the calibration. The residual error has to be applied for all the structural conditions, both undamaged and damaged, despite it was computed referring only to the undamaged state. There is no guarantee that the residual error does not change when the structure is damaged, but its computation for different damage scenarios is not feasible in almost all real applications. Moreover, the proposal to adjust the network input data by adding the residual term aims to cut down the effect of model error for the undamaged condition. In the authors' opinion, it is of less relevance if it slightly alters the prediction of modal properties in the damaged condition.

Predictions of networks obtained by adding the residual term to the reference modal properties are shown in Table 7. As concerns N1 and N2, the first four scenarios are correctly identified. Scenarios involving damage in the concrete slab (S5, S6 and S7) are correctly identified by the network N1 except for S6, while the network N2 identifies all of them as undamaged. This result confirms that it is not trivial to identify local damage on the concrete slab. Network N3 presents the worst performances. Although it identifies damage for the scenarios S5, S6 and S7, the undamaged condition (scenario S1) is classified as lightly damaged and scenario S4 is mistakenly recognized as undamaged.

The behavior of all networks is analyzed also by adding the measurement noise. In particular, the exact values of modal properties computed by model R are corrupted by a Gaussian noise. The frequency coefficient of variation (CV) is fixed to 1%, while CV in the range [1%, 10%] is selected to modify the mode shape components.
For each value of the coefficient of variation, 100 samples of pseudo-experimental data are extracted and the predictions of networks N1, N2 and N3 are computed. Results reported here refer only to the scenario S1 (undamaged condition of model R). Fig. 3 shows the trend of the accuracy with the coefficient of variation of the mode shape components for the three networks. As expected, network N3 is not able to identify the undamaged condition regardless of the added noise. The accuracy of network N1 rapidly decreases if the level of noise increases while N2

Table 5. Performances of networks N1, N2 and N3 for the dataset of model S.

Performance	N1		N2		N3	
	Train	Test	Train	Test	Train	Test
Strict accuracy [%]	67	67	94	93	93	93
Soft accuracy [%]	81	81	97	97	96	96
Uncertain predictions [%]	54	54	7	8	8	9

Table 6. Results of the test with model R. U: undamaged; LD: light damage; SD: severe damage.

Network identifier	Scenario						
	S1 (Undam.)	S2 (Dam.)	S3 (Dam.)	S4 (Dam.)	S5 (Dam.)	S6 (Dam.)	S7 (Dam.)
N1	SD	SD	LD	SD	SD	SD	SD
N2	SD	SD	SD	SD	SD	SD	SD
N3	SD	SD	SD	SD	SD	SD	SD

Table 7. Results of the test with model R when model error is considered. U: undamaged condition; SD: severe damage; LD: light damage.

Network identifier	Scenario						
	S1 (Undam.)	S2 (Dam.)	S3 (Dam.)	S4 (Dam.)	S5 (Dam.)	S6 (Dam.)	S7 (Dam.)
N1	U	SD	LD	LD	LD	U	LD
N2	U	SD	SD	LD	U	U	U
N3	LD	SD	SD	U	LD	SD	LD

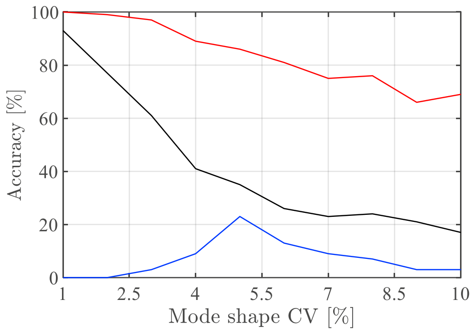

Fig. 3. Accuracy trend of networks N1 (black line), N2 (red line) and N3 (blue line) with mode shape CV for noise-corrupted data when model error is considered.

keeps good performances, with accuracy never lower than 65 %.

5. Conclusions

In this paper, a damage detection procedure using artificial neural networks has been presented. The procedure has been applied to the case study of a railway bridge. Modal properties have been chosen as damage sensitive features. Only simulated data have been adopted, but model error has been taken into account by using two different FE models of the same structure. Data generated by the models have been corrupted with Gaussian noise in order to simulate measurement uncertainties. Damage detection has been performed by a network using noise-corrupted modal properties (N1) and by two networks using noise-filtered modal properties (N2 and N3). The noise filtering has been carried out by means of the PCA technique. As concerns the training and test phase with model S, N1 has shown the worst performances, while N2 and N3 have presented large values of accuracy and low percentages of uncertain predictions. The test phase with model R has highlighted the need to account for the model error. Working with the exact data, all networks have not been able to identify the undamaged condition. Conversely, if the input data are corrected with the residual error obtained after the model calibration, networks N1 and N2 correctly identify the heathy state and the presence of damage on the steel beam. Damage on the concrete slab has been more difficultly identified. Finally, results obtained by also adding measurement noise have shown for N2 a limited reduction of the accuracy with an increasing level of noise.

References

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., Inman, D. J., 2021. A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. Mechanical Systems and Signal Processing, 147, 107077.

Bishop, C. M., 2006. Pattern Recognition and Machine Learning. Springer.

Comanducci G., Magalhães F., Ubertini F., Cunha Á, 2016. On vibration-based damage detection by multivariate statistical techniques: Application to a long-span arch bridge. Structural Health Monitoring, 15 (5), 505- 524.

Doebling, S. W., Farrar, C. R., Prime, M. B., Schevitz, D. W., 1996. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review. Technical Report LA-13070-MS, Los Alamos National Laboratory.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. Prentice Hall.

Hou, R., Xia, Y., 2021. Review on the new development of vibration-based damage identification for civil engineering structures: 2010–2019. Journal of Sound and Vibration, 491, 115741.

Khan, S., Yairi, T. 2018. A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing, 107, 241-265.

Møller, M., 1993. A Scaled Conjugate Gradient Algorithm For Fast Supervised Learning. Neural Networks, 6, 525-533.

Sohn, H., Farrar, C. R., Hemez, F. M., Czarnecki, J. J., 2002. A Review of Structural Health Monitoring Literature 1996-2001. Report number: LA-13976-MS, Los Alamos National Laboratory.

Ying, X., 2019. An Overview of Overfitting a.nd its Solutions. Journal of Physics: Conference Series, 1168, 022022.