Available online at www.sciencedirect.com

ScienceDirect

Procedia Structural Integrity 84 (2026) 1231–1238

Structural Integrity Procedia

www.elsevier.com/locate/procedia

III Fabre Conference: Existing Bridges, Viaducts, and Tunnels: Research, Innovation, and Applications

Detect and classify damage in bridges through MLP neural networks trained on SHM data

Augusto Montisci[a], Francesca Pibi[b], Maria Cristina Porcu[b,]*

[a]*Dept. of Electrical and Electronic Engineering, University of Cagliari, Via Marengo 2, 09123 Cagliari, Italy*
[b]*Dept. of Civil, Environmental Engineering and Architecture, University of Cagliari, Via Marengo 2, 09123 Cagliari, Italy*

Abstract

The paper presents an intelligent procedure to detect the presence, classify the type, and identify different levels of damage in existing bridges. Based on Multi-Layer Perceptron (MLP) neural networks, the procedure is tested referring to the Z24 bridge benchmark, where different damage scenarios were progressively induced. The proposed methodology trains MLP neural networks on data from both Forced Vibration Tests (FVT) and Ambient Vibration Tests (AVT), specifically trained for each sensor setup using frequency-domain selected features. A majority voting criterion is then applied to aggregate independent diagnoses into a robust final classification. Results demonstrate that this computationally efficient framework accurately distinguishes complex structural pathologies, offering a scalable solution for automated, real-time structural health monitoring. Future application of the proposed procedure involves training the MLP network on simulated damage scenarios of well-identified numerical models of real structures, which can lead to significant advancements in automated structural health monitoring.

Keywords: intelligent structural health monitoring; artificial neural networks; multi-layer perceptron; damage detection and classification; concrete bridges.

* Corresponding author. Tel.: +39-070-675-5414.
E-mail address: mcporcu@unica.it

10.1016/j.prostr.2026.06.158

1. Introduction

Ensuring structural safety is a concern that covers the entire service lifespan of bridges. These structures may be affected by the progressive degradation of material properties, peak operational loading and the effects of fatigue due to cyclic loads, as well as extreme loading caused by environmental phenomena such as earthquakes, strong winds and flooding. Each of these scenarios may impact safety, stress and deformation levels, as well as integrity requirements. Continuous Structural Health Monitoring (SHM) of external loading and structural responses may enable real-time evaluation of structural safety and integrity. Based on SHM data, various techniques can be employed to detect damage in bridges and assess their safety, cf. e.g. Seo et al. (2016); Porcu et al. (2025); Caredda et al. (2022); Porcu et al. (2019). Artificial Intelligence (AI) has been recently employed to help diagnose the structural integrity of bridges based on SHM sensor network measurements, see Zinno et al. (2022).

This paper proposes an intelligent diagnosis method that can identify, classify and distinguish different levels of damage in bridges through artificial neural networks (ANNs) trained SHM datasets. This method was tested using data from the Z24 highway post-tensioned concrete bridge, which was deliberately damaged in various controlled scenarios prior to its demolition in 1998. Valuable data was obtained by recording the time histories of accelerations at different structural points under both harmonic Forced Vibration Tests (FVTs) and Ambient Vibration Tests (AVTs) for each damage scenario. Maeck and De Roeck (2003) provided details of the experimental campaign that was carried out. The databases of the Z24 benchmark case were used to assess several machine-learning approaches for damage detection in bridges, cf. e.g. Sony et al. (2022), Dabbous et al. (2024), Hoang et al. (2024), Giglioni et al. (2023), Nguyen-Tran et al. (2023).

Unlike other intelligent methods used to detect structural damage, the method proposed in this paper analyses SHM data using Multi-Layer Perceptron (MLP) ANNs, which are the simplest neural structures that can be used to classify non-linearly separable sets. One of the main advantages of MLP ANNs is that they are more computationally efficient than other neural network architectures that could be used for damage classification (Foddis et al. (2015); Secci et al. (2015); Foddis et al. (2019)).

The effectiveness of the proposed MLP-ANN approach is assessed by considering the Z24 bridge databases relevant to FVTs and AVTs. Acceleration signals acquired at different bridge's locations, under various damage scenarios, are used to train simple MLP ANN classifiers, and to detect and identify different types and levels of damage. A preliminary investigation done by Montisci et al. (2025) showed a good performance of the proposed method when the FVT data are considered. Here, the performance of the method is assessed with both FVT and AVT data. A comparison of the results provides useful insights into the practical application of the present intelligent approach.

Nomenclature

AVT	Ambient Vibration Test
FVT	Forced Vibration Test
PD	Progressive Damage
$\boldsymbol{X_{jik}}$	MLP Input vector
$\boldsymbol{Y_{jik}}$	MLP Output vector
$\boldsymbol{H_{jik}}$	MLP Hidden Layer vector
$\boldsymbol{\overline{Y}_{ji}}$	Average Output vector
$\boldsymbol{\overline{G}_{j}}$	Average Output Gap vector

2. Z24 bridge benchmark databases

The case of the Z24 bridge (Fig. 1a) is a well-known benchmark in structural engineering. It was a three-span post-tensioned concrete box-girder bridge (Figg. 1b and 1c) located in Switzerland, which entered service in 1963 and was demolished in 1998 due to the need for a new bridge (Maeck and De Roeck (2003)). Before its demolition, the Z24 bridge was tested under full-scale different damage scenarios to provide an experimental basis for evaluating the feasibility of using vibration-based SHM methods to identify damage. To this purpose, the bridge was extensively instrumented (De Roeck (1999)), and it was subjected to 17 short-term controlled progressive damage (PD) scenarios (Krämer et al. (1999); Roeck (2003)). There are three reference scenarios PD-1, PD-2 and PD-8 relevant respectively to the undamaged condition (PD-1), the installation of the settlement system in Pier 2 (PD-2) and restoring the pier condition (PD-8). Five reversible settlement scenarios (PD-3, PD-4, PD-5, PD-6) are then considered, where Pier 2 was lowered in turn by 20 mm, 40 mm, 80 mm and 95 mm and tilted (PD-7). Finally, ten irreversible scenarios are considered: concrete spalling (PD-9, PD-10), landslide (PD-11), failure of concrete hinges (PD-12), failure of tendon anchor heads (PD-13, PD-14), and rupture of tendons (PD-15, PD-16, PD-17).

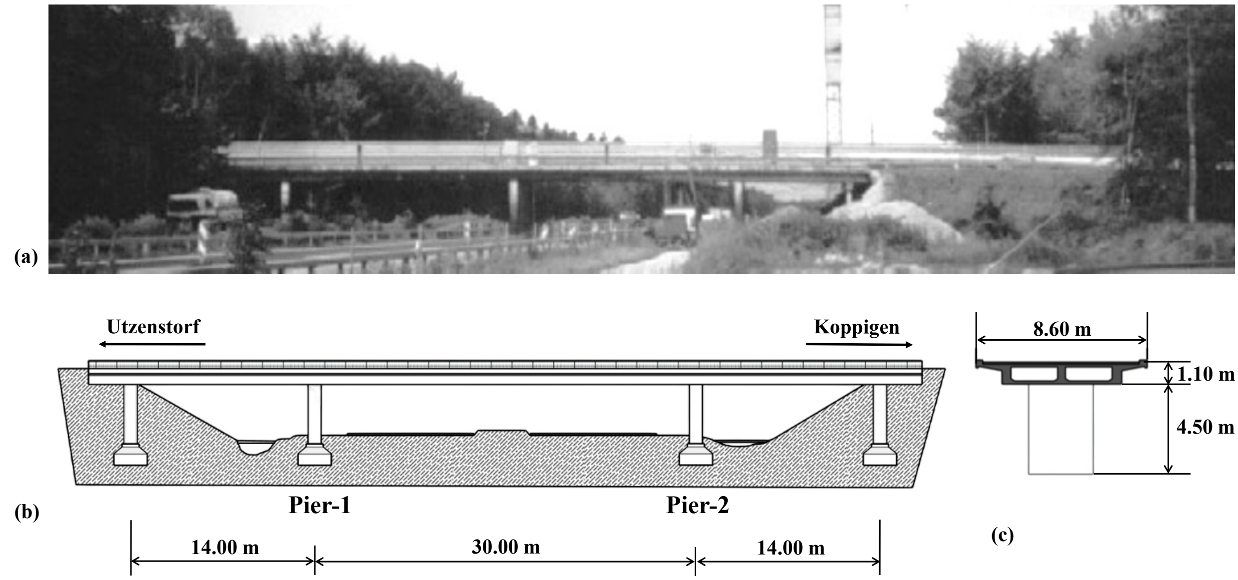

Fig. 1. (a) Archive picture of the Z24 bridge, after Maeck and De Roeck, 2003; (b) bridge's longitudinal view and (c) cross section.

Ambient vibration tests (AVTs) and forced vibration tests (FVTs) were carried out for each PD scenario. In FVTs, harmonic forces were applied through two hydraulic shakers placed on the deck (Krämer et al. (1999)). The response was acquired by roving a set of 15 three-axial accelerometers along the deck paired with a couple of three-axial accelerometers roved on Pier 1 or Pier 2, depending on the setup. In total, 9 different setups, referred to as S01, S02, …, S09, were considered. Three accelerometers were also kept fixed at three reference positions on the deck. Synchronized signals were recorded from 34 channels (33 accelerometer channels and one shaker channel), see Krämer et al. (1999) and Peeters (2000). Each acquisition lasted 10.9 minutes and was sampled simultaneously at 100 Hz. Each signal consisted of 65,536 time samples.

3. MLP damage classification method

The method proposed in the paper is based on training MLP neural networks on SHM acceleration data to identify the different damage scenarios. The effectiveness of the procedure was assessed by using signals acquired in the Z24 bridge from the 9 setups and under the 17 PD scenarios. A classification problem was thus defined, where each damage scenario represented a class. Two datasets were created: one relevant to the FVTs and the other one to the AVTs. The two datasets collected the raw data acquired at the 9 setups in the 17 PD scenarios under forced and ambient vibrations, respectively. Since signals belonging to different setups were not synchronized with each other, a specific MLP was trained for each setup. Therefore, 9 different diagnostic systems were developed. The set of diagnoses obtained from the different MLPs was finally combined with the majority voting criterion. AVT and FVT datasets are analyzed separately.

Fig. 2 shows the procedure followed both in training and recall phases for each setup. In the training phase the outputs are compared to the target values and the difference is used to train the network. In the recall phase, the outputs represent the diagnosis. The procedure consists of three main steps: data preprocessing, frequency component selection and MLP recall. The steps were performed using MATLAB (MATLAB version: 24.2 (R2024b), (2024)).

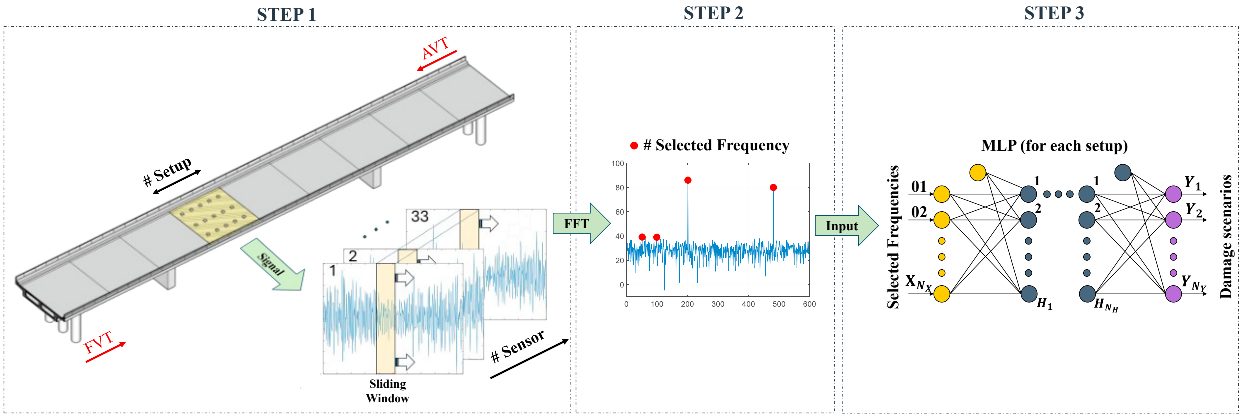

Fig. 2. Procedure followed both in training and recall phases for each setup. In the training phase, the outputs are compared to the target values, and the difference is used to train the network. In the recall phase, the outputs represent the diagnosis.

3.1. Preprocessing of data

The first step involves preprocessing the raw data obtained from SHM. The preprocessing procedure is applied to the two Z24 bridge's datasets. For each setup and PD, the signals acquired by each channel are divided into sliding windows of 4,000 samples. The windows are shifted by 500 samples, resulting in 123 windows per sequence. Synchronized time windows are collected from all the setups to create a 4,000 × 9 matrix representing an instance of the dataset for each PD. A total of 2,091 instance matrixes is eventually considered, as obtained from 123 windows multiplied by 17 PDs. The 2-dimensional fast Fourier transform (FFT) is applied to each instance matrix. Only the FFT amplitude was considered, as the phase of the frequency components was subject to uncertainty due to signals that had not been triggered.

3.2. Frequency component selection

A selection of a proper number N_f of frequency components is made. To make training easier, the number of components should be kept to a minimum. To this end, a quality index, given by the ratio between the range of each class (PD) and the range of the entire dataset, is calculated for each frequency component. The smaller the quality index the more suitable the frequency component to distinguish a class from the others. A ranking of quality indexes is defined for each class and for each setup. The first N_f frequency components of the ranking are selected for each setup. The union of the frequencies (taken once), selected in the N_Y classes (PDs), gives the input pattern. The components of the input patterns are normalized to the interval $[-1,1]$ by considering the range of values within the training set. When the FVT dataset is concerned, it is chosen $N_f = 7$, while $N_f = 50$ is set for the AVT dataset. Owing to the higher number of frequency components, in AVT dataset a correlation-based criterion is also adopted to avoid redundancy of information provided by the selected frequencies. The set of frequencies is built in stages. A frequency is added only if the correlation threshold, with all previously selected frequencies, is not exceeded.

3.3. MLPs training and recall

An MLP neural network with variable structure is used as classifier for each setup (see **Errore. L'origine riferimento non è stata trovata.**). The network associates the input vector $\boldsymbol{X}$ with the output vector $\boldsymbol{Y} = f(\boldsymbol{X})$, where $f(.)$ denotes the neural network function. In the training phase, output $\boldsymbol{Y}$ is compared to the target vector $\boldsymbol{T}$ that represents the class which the current instance belongs to. The difference between $\boldsymbol{Y}$ and $\boldsymbol{T}$ drives the training process. In the recall phase, output $\boldsymbol{Y}$ is the diagnosis of the system. The input vector $\boldsymbol{X}$ collects N_X selected frequencies. A suitable number N_H of hidden layers is considered, each of them consisting of H_h (h=1, …, N_H) neurons. The output vector $\boldsymbol{Y}$ has N_Y neurons, where N_Y is the number of classes (PDs). The output layer activation is a hyperbolic tangent function. For the two Z24 bridge datasets, the training phase is carried out as described below.

- *Network structure for the FVT dataset*. An MLP $N_X - H_1 - N_Y$ structure is used, with only one hidden layer ($N_H = 1$). Due to the variable dimension of the input pattern, the number N_X of input neurons may change from one setup to the other. The first seven frequencies ($N_f = 50$) in the ranking are sufficient to create the neural network's input pattern. The union of the selected frequencies (taken once) is used as an input pattern. The components of the input patterns are normalized to the interval $[-1,1]$ by considering the range of values within the training set. A hidden layer made of 10 neurons ($H_1 = 10$) is considered for all the setups. The output layer consists of as many neurons as the number of classes (damage scenarios), that is $N_Y = 17$. The current class is assigned the value 1 and all the others are assigned the value -1.

- *Network structure for the AVT dataset*. An $N_X - H_1 - H_2 - N_Y$ MLP structure is used, with two hidden layers made of respectively 20 and 10 neurons ($H_1 = 20$ and $H_2 = 10$) for all setups. The first fifty frequency components in the ranking are considered to create the input vector. Again, the union of the selected frequencies is used as an input pattern, and the components of the input patterns are normalized to the interval $[-1,1]$. The output layer consists of 17 neurons ($N_Y = 17$) and the current class is assigned the value 1 and all the others are assigned the value -1.

In all cases, the number of training epochs is set to 1,000, although training stops before reaching this number. Three independent random subsets are considered for both the FVT and AVT datasets: the training set (75% of samples); the validation set (15% of samples); the test sets (15% of samples). Several runs are performed by selecting different subsets for validation and testing. The performance trend on both the validation and test sets is found to be like that on the training set. This indicates that all three subsets are representative of the whole dataset, meaning that the distribution in the input space is well represented.

4. Results and discussion

The effectiveness of the proposed procedure is assessed by evaluating the number of right diagnoses. For this purpose, the average output gap and the confusion matrix are defined below.

For each setup i ($i = 1, 2, \dots, N_S$, where $N_S = 9$), and each class j ($j = 1, 2, \dots, N_Y$, where $N_Y = 17$), let's consider the k −th window ($k = 1, 2, \dots, N_w$, where $N_w = 123$). The input vector of the instance is denoted by $\boldsymbol{X_{ijk}}$, and the output vector $\boldsymbol{Y_{ijk}} = f(\boldsymbol{X_{ijk}})$, where $f(\cdot)$ is the neural network function. The diagnosis is based on the average output vector $\overline{\boldsymbol{Y}}_{\boldsymbol{ijk}}$, which is calculated as:

$$\overline{\boldsymbol{Y}}_{\boldsymbol{ij}} = \frac{1}{N_w} \sum_{k=1}^{N_w} f(\boldsymbol{X_{ijk}}) \tag{1}$$

The highest-value component of $\overline{\boldsymbol{Y}}_{\boldsymbol{ij}}$ is assumed as the diagnosis provided by the MLP_i, based on the $i - th$ setup. Let's define the output gap $\boldsymbol{G_{ji}}$ as the difference between the first and second highest values of $\overline{\boldsymbol{Y}}_{\boldsymbol{ij}}$. The average output gap $\overline{\boldsymbol{G}}_{\boldsymbol{j}}$ over the N_S setups is defined as:

$$\overline{\boldsymbol{G}}_{\boldsymbol{j}} = \frac{1}{N_S} \sum_{i=1}^{N_S} \boldsymbol{G_{ji}} \tag{2}$$

It is worth noting that, due to normalization, the average output gap $\overline{\boldsymbol{G}}_{\boldsymbol{j}}$ ranges between 2 and 0, where $\overline{\boldsymbol{G}}_{\boldsymbol{j}} = 2$ indicates full classification confidence, while $\overline{\boldsymbol{G}}_{\boldsymbol{j}} = 0$ is obtained when classification is not univocal. Fig. 3a and Fig. 4a provide the gaps of the diagnoses for FVT and AVT datasets, respectively. As can be seen, the FVT dataset allows to saturate the gap in all classes, whereas the AVT dataset produces smaller gaps with different values for each class. The robustness of the diagnosis is evaluated in Figg. 3b and 4b by means of the Confusion Matrix. For each damage class, the diagnoses provided by the nine MLPs are reported in a matrix row. In the ideal case, the confusion matrix would be diagonal. If there are values outside the diagonal, some of the MLPs misclassify. The confusion matrix shows which pairs of classes tend to be confused with each other. From Fig. 3b we can deduce that the diagnostic system based on the FVT dataset is robust, as no MLP misclassifies anything. Conversely, the diagnostic system based on the AVT dataset (Fig. 4b) misclassifies some classes of damage. Nevertheless, the value on the diagonal is by far the largest in the row, implying that the system can make the correct diagnosis in 100% of cases.

The AVT-based diagnostic system is much less robust than the FVT-based one. However, FVTs are more difficult to carry out as they require the temporary suspension of traffic on the bridge. For this reason, the two types of diagnostic system can be used in a complementary way: the AVT-based system can be used for continuous monitoring of in-service bridges, and the FVT-based system can be used when an alarm is triggered by the AVT-based system.

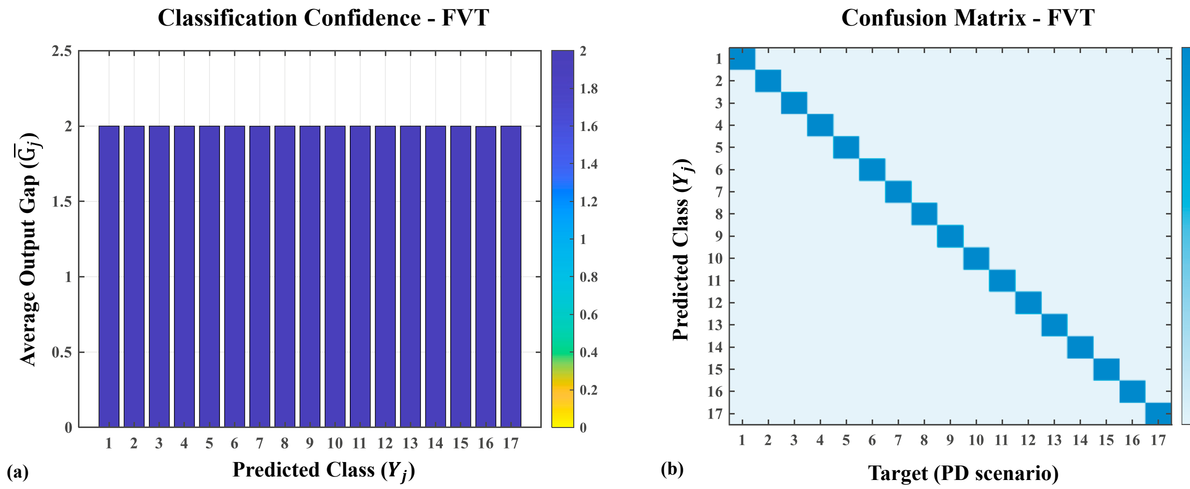

Fig. 3. FVT dataset. (a) Classification confidence and (b) Confusion Matrix.

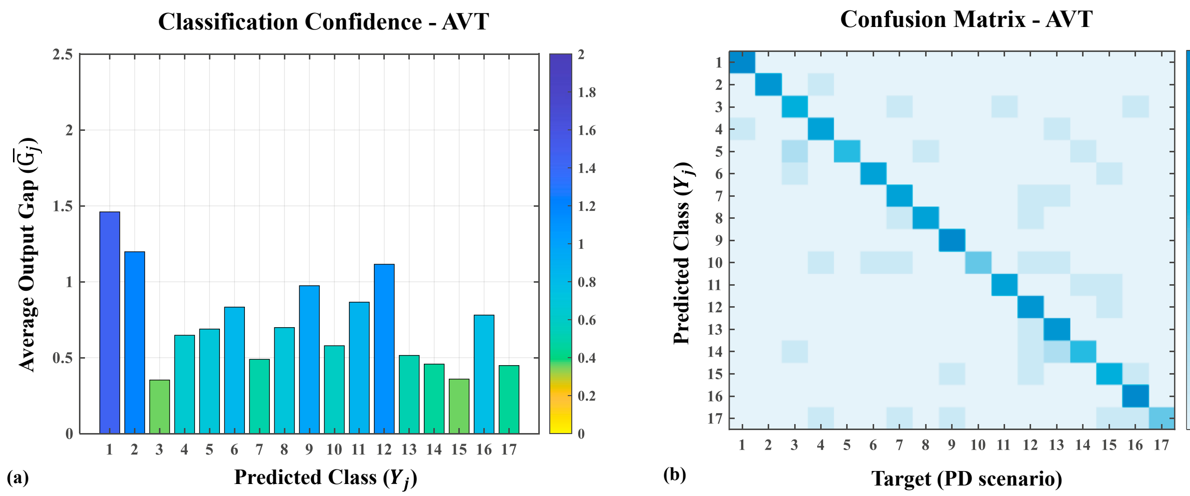

Fig. 4. AVT dataset. (a) Classification confidence and (b) Confusion Matrix.

5. Conclusions

The paper presents an intelligent method to detect and classify damage on bridges. It is based on Structural Health Monitoring (SHM) data and Multi-Layer Perceptron (MLP) Artificial Neural Networks (ANNs). The effectiveness of the method is assessed with reference to the benchmark Z24 bridge. The variety of acquisition setups and sensors, as well as the significant amount of damage scenarios, makes the Z24 bridge datasets particularly suitable for validating the diagnostic ability of the method.

The signals acquired on the Z24 bridge under ambient vibration tests (AVTs) and forced vibration tests (FVTs) from nine accelerometer setups, in 17 damage scenarios, make up the two AVT and FVT raw datasets used to assess the method. A classification problem is defined, where each damage scenario represents a class. The average output gap and the confusion matrix are used to evaluate the level of confidence of the diagnostic system.

The findings demonstrate the effectiveness of the MLP-based method in detecting and classifying damage scenarios when using both AVT and FVT datasets. However, the FVT-based diagnostic system is much more robust than the AVT-based system. Nevertheless, forced oscillation testing on a functioning bridge is more expensive and impractical than ambient testing because it requires the bridge to be closed to traffic temporarily. Conversely, ambient vibrations offer a cost-effective means of testing in-service bridges, albeit with much lower signal-to-noise ratios than FVT, which introduces greater uncertainty in the results. This suggests that the two different kinds of test can be used for an integrated damage diagnosis method. The AVT-based diagnostic system could be used for continuous monitoring, while the FVT-based diagnostic system could be implemented when an alarm is triggered.

Future studies will investigate the implementation of well-identified numerical finite element models, where damage scenarios can be simulated to train the MLPs and make the network suitable for early detection and classification of damage in the real structure.

Acknowledgements

The authors would like to thank the Structural Mechanics Section at KU Leuven for granting them access to the Z24 Bridge Benchmark Database. They also gratefully acknowledge Professor Emeritus Guido De Roeck for allowing them to use the Z24 bridge photography.

It is acknowledged the financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Comp. 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union –NextGenerationEU–CUP F53D23009640001- Assignment Decree n. P20227CSJ5.

References

Caredda, G., Porcu, M.C., Buitrago, M., Bertolesi, E., Adam, J.M., 2022. Analysing local failure scenarios to assess the robustness of steel truss-type bridges. Eng. Struct. 262, 114341.

Dabbous, A., Berta, R., Fresta, M., Ballout, H., Lazzaroni, L., Bellotti, F., 2024. Bringing intelligence to the edge for structural health monitoring: the case study of the Z24 bridge. IEEE Open J. Ind. Electron. Soc. 5, 781–794.

De Roeck, G., 1999. SIMCES Synthesis Report. BRITEEuram BE96-3157 Kathol. Univ. Leuven.

Foddis, M.L., Ackerer, P., Montisci, A., Uras, G., 2015. ANN-based approach for the estimation of aquifer pollutant source behaviour. Water Supply 15, 1285–1294.

Foddis, M.L., Montisci, A., Trabelsi, F., Uras, G., 2019. An MLP-ANN-based approach for assessing nitrate contamination. Water Supply 19, 1911–1917.

Giglioni, V., Venanzi, I., Baia, A.E., Poggioni, V., Milani, A., Ubertini, F., 2023. Deep autoencoders for unsupervised damage detection with application to the Z24 benchmark bridge. In: Rizzo, P., Milazzo, A. (Eds.), European Workshop on Structural Health Monitoring, Lecture Notes in Civil Engineering. Springer International Publishing, Cham, pp. 1048–1057.

Hoang, V.H., Nguyen, N.L., Bui, T.T., Tran, N.H., 2024. A two-stage method for damage detection in Z24 bridge based on K-nearest neighbor and Artificial Neural Network. Period. Polytech. Civ. Eng. 68, 892–902.

Krämer, C., De Smet, C.A.M., De Roeck, G., 1999. Z24 bridge damage detection tests. In: IMAC 17, the International Modal Analysis Conference. Society of Photo-optical Instrumentation Engineers, Kissimee, Fl, USA, pp. 1023–1029.

Maeck, J., De Roeck, G., 2003. Description of Z24 benchmark. Mech. Syst. Signal Process. 17, 127–131.

MATLAB version: 24.2 (R2024b), 2024.

Montisci, A., Pibi, F., Porcu, M.C., 2025. MLP neural networks to identify damage in bridges from SHM data. In: The Seventh International Conference on Artificial Intelligence, Soft Computing, Machine Learning and Optimization in Engineering, Paper 2.2. Civil-Comp Press, Edinburgh, UK, pp. 1–14.

Nguyen-Tran, H., Bui-Ngoc, D., Ngoc-Nguyen, L., Tran, H., Bui-Tien, T., De Roeck, G., Wahab, M.A., 2023. The application of a hybrid Autoregressive and Artificial Neural Networks to structural damage detection in Z24 bridge. Recent Adv. Struct. Health Monit. Eng. Struct., Lecture Notes in Mechanical Engineering 417–425.

Peeters, B., 2000. System identification and damage detection in civil engineering (PhD thesis). Katholieke Universiteit te Leuven.

Porcu, M.C., Buitrago, M., Calderón, P.A., Garau, M., Cocco, M.F., Adam, J.M., 2025. Robustness-based assessment and monitoring of steel truss railway bridges to prevent progressive collapse. J. Constr. Steel Res. 226, 109200.

Porcu, M.C., Patteri, D.M., Melis, S., Aymerich, F., 2019. Effectiveness of the FRF curvature technique for structural health monitoring. Constr. Build. Mater. 226, 173–187.

Roeck, G.D., 2003. The state-of-the-art of damage detection by vibration monitoring: the SIMCES experience. J. Struct. Control 10, 127–134.

Secci, R., Laura Foddis, M., Mazzella, A., Montisci, A., Uras, G., 2015. Artificial Neural Networks and Kriging method for slope geomechanical characterization. Eng. Geol. Soc. Territ. - Vol. 2 1357–1361.

Seo, J., Hu, J.W., Lee, J., 2016. Summary review of Structural Health Monitoring applications for highway bridges. J. Perform. Constr. Facil. 30, 04015072.

Sony, S., Gamage, S., Sadhu, A., Samarabandu, J., 2022. Multiclass damage identification in a full-scale bridge using optimally tuned one-dimensional Convolutional Neural Network. J. Comput. Civ. Eng. 36, 04021035.

Zinno, R., Haghshenas, S.S., Guido, G., Vitale, A., 2022. Artificial Intelligence and Structural Health Monitoring of bridges: a review of the State-of-the-Art. IEEE Access 10, 88058–88078.