Finite Element–Based Machine-Learning Approach to Detect Damage in Bridges under Operational and Environmental Variations

Eloi Figueiredo[1]; Ionut Moldovan[2]; Adam Santos[3]; Pedro Campos[4]; and João C. W. A. Costa[5]

Abstract: In the last decades, the long-term structural health monitoring of civil structures has been mainly performed using two approaches: model based and data based. The former approach tries to identify damage by relating the monitoring data to the prediction of numerical (e.g., finite-element) models of the structure. The latter approach is data driven, where measured data from a given state condition are compared to the baseline or reference condition. A challenge in both approaches is to make the distinction between the changes of the structural response caused by damage and environmental or operational variability. This issue was tackled here using a hybrid technique that integrates model- and data-based approaches into structural health monitoring. Data recorded in situ under normal conditions were combined with data obtained from finite-element simulations of more extreme environmental and operational scenarios and input into the training process of machine-learning algorithms for damage detection. The addition of simulated data enabled a sharper classification of damage by avoiding false positives induced by wide environmental and operational variability. The procedure was applied to the Z-24 Bridge, for which 1 year of continuous monitoring data were available. **DOI: 10.1061/(ASCE)BE.1943-5592.0001432.**

Author keywords: Structural health monitoring; Machine learning; Finite-element modeling; Damage detection; Damage identification.

Introduction

Structural health monitoring (SHM) is defined as the process of implementing a damage-identification strategy for civil, aerospace, and mechanical engineering structures (Charles et al. 2001). The damage identification is defined using a five-level hierarchy: detection, localization, type, extent, and prognosis (Figueiredo et al. 2013). For more than two decades, SHM has been performed mainly based on two independent approaches: model based and data based. The former is typically rooted in finite-element (FE) model-updating techniques, and the latter in the machine-learning field, where machine-learning algorithms (MLAs) have played an important role (Figueiras et al. 2011).

Model-updating techniques aim at identifying structural damage by comparing the measured response of the structure with a baseline FE model, tailored for that specific structure and validated against its undamaged behavior. The FE method is currently the most widely used tool for creating computational models of complex engineering structures (Zienkiewicz et al. 2013). It divides the structure under analysis into small parts (the finite elements), whose nodal displacements are the main unknowns of the problem. Sometimes, these displacements can be compared directly with measured data and the differences used to detect structural damage, which is an approach known as direct model updating. However, due to the necessity of actually measuring displacement data at the FE nodes, direct model-updating procedures are limited to very simple structures, so iterative updating methods are typically used for complex structural systems. Iterative methods exploit the discretized nature of the FE method. The basic idea is to use the recorded structural response to update some calibration parameters (e.g., stiffness properties, boundary conditions), defined at the FE level, to minimize an objective function, defined as the difference between the computed and the measured structural responses (Mirzaee et al. 2015). When structural damage is suspected, the variation of the calibration parameters can be used to identify and classify it.

Unfortunately, considerable uncertainty is inherent to FE model updating. It includes both the random occurrences associated with the experimental readings of the structural response and the epistemic uncertainties associated with the lack of knowledge regarding the exact properties of the materials, cross sections of the structural elements, and FE discretization. Moreover, operational and environmental variations (e.g., traffic, temperature, humidity, and wind speed) often arise as unwanted effects in the structural responses, affecting damage-sensitive features and masking changes caused by damage (Sohn 2007). The inclusion of such uncertainties in the

[1]Associate Professor, Faculty of Engineering, Univ. Lusófona de Humanidades e Tecnologias, Campo Grande 376, Lisbon 1749-024, Portugal; Integrated Member, CONSTRUCT, Institute of R&D in Structures and Construction, R. Dr. Roberto Frias s/n, Porto 4200-465, Portugal (corresponding author). ORCID: https://orcid.org/0000-0002-9168-6903. Email: eloi.figueiredo@ulusofona.pt

[2]Assistant Professor, Faculty of Engineering, Univ. Lusófona de Humanidades e Tecnologias, Campo Grande 376, Lisbon 1749-024, Portugal; Integrated Member, CERIS, Instituto Superior Técnico, Univ. de Lisboa, Av Rovisco Pais, Lisbon 1049-001, Portugal. ORCID: https://orcid.org/0000-0003-3085-0770. Email: dragos.moldovan@ulusofona.pt

[3]Assistant Professor, Faculty of Computing and Electrical Engineering, Univ. Federal do Sul e Sudeste do Pará, F. 17, Q. 4, L. E., Marabá, Pará 68505-080, Brazil. ORCID: https://orcid.org/0000-0002-5940-4961. Email: adamdreyton@unifesspa.edu.br

[4]Graduate Student, Faculty of Engineering, Univ. Lusófona de Humanidades e Tecnologias, Campo Grande 376, Lisbon 1749-024, Portugal. Email: pmfc26@hotmail.com

[5]Full Professor, Applied Electromagnetism Laboratory, Univ. Federal do Pará, R. Augusto Corrêa, Guamá 1, Belém, Pará 66075-110, Brazil. Email: jweyl@ufpa.br

Note. This manuscript was submitted on July 25, 2018; approved on February 4, 2019; published online on May 1, 2019. Discussion period open until October 1, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Bridge Engineering***, © ASCE, ISSN 1084-0702.

FE model shifts model-based SHM from deterministic to probabilistic (Smith and Saitta 2008). A parent FE model is constructed using the nominal values of the geometric and material properties, and offspring models are subsequently generated based on the plausible range of the uncertain parameters and possibly on their probabilistic distribution. Distribution-based approaches directly output the probabilistic distribution of the structural responses but are limited to relatively simple FE models because the offspring sample size is typically on the order of tens of thousands (Catbas et al. 2013). Alternatively, the number of offspring can be considerably reduced by the inclusion of uncertainties using the Latin hypercube method (Barthorpe 2010). However, this approach is less able to reproduce the probabilistic distribution of the structural responses with good resolution. In general, model updating is a time-consuming process that requires significant intervention from users, especially when highly nonlinear effects of severe structural damage are taken into account.

Conversely, approaches based on data models are independent from the physics and complexity of the structure. Several MLAs with different working principles have been proposed in the last two decades to separate changes in the damage-sensitive features caused by structural damage from those caused by varying operational and environmental conditions (Figueiredo et al. 2011; Santos et al. 2016; Figueiredo and Cross 2013; Wursten and Roeck 2014; Santos et al. 2015). In SHM, machine learning is used to get computers and algorithms to model the reality without knowing the physical laws of structures. Thus, MLAs are used to learn the structural behavior from the experience, following the same principle as the human brain. They can be trained in unsupervised or supervised learning modes. In the SHM field, unsupervised learning refers to the case in which training data are only available from the undamaged condition. Supervised learning refers to the case in which data from undamaged and damaged conditions are available to train the algorithms. Most high-capital-expenditure civil engineering structures, such as bridges, are systems dictated by the size and physical environment in which they are built, which challenges the existence of data from all operational and environmental conditions. In addition, due to the one-of-a-kind nature of such structures, it is more difficult to incorporate lessons learned from other nominally *similar* systems to define anticipated damage or even normal response patterns throughout their service lifetime. Therefore, in civil engineering, these algorithms are often output only and unsupervised because they are trained only with damage-sensitive features from the response data related to the undamaged condition, without any measurement from the operational and environmental parameters. They have proved to be suitable to address the first two levels of the damage-identification hierarchy (i.e., detection and localization), but their performance is highly dependent on the amount and variability of the data used during the training process (Figueiredo and Cross 2013).

In the recent years, a hybrid approach aimed at combining the best features of model- and data-based SHM has emerged. The basic principles of the hybrid approach were stated by Barthorpe (2010). The FE model is seen not only as an instrument for exploring the possible causes of past changes in the structural response (through model updating) but also as a tool for designing an *optimal* SHM sensor network and a proxy for damage induction to enable the supervised training of data-based damage-detection algorithms.

Following the hybrid approach, FE models were used to optimize the location of piezoelectric sensors for crack detection in space structures (Gresil et al. 2011) and of two sensors mounted on a laboratory-scale bridge structure (Malekzadeh et al. 2015). Although the FE models have been shown to hold considerable potential for the optimization of small sensing networks, the number of runs required for a large-scale structure is probably prohibitive (for 20 sensors and 100 potential locations, some 10^{20} FE runs would be required).

Regarding the supervised training of data-based damage-detection algorithms, FE models can be used as a proxy for the invasive, damage-inducing procedures needed to obtain response data from typical damage scenarios. Also, FE models can be used to predict the structural responses not recorded by the monitoring systems in undamaged scenarios (environmental temperature variation and large concentrated loads are typical cases).

The combination of FE modeling and machine learning has recently been tackled by other researchers. Liu and Zhang (2017) constructed stochastic FE models of a structure to serve as a baseline for damage identification under environmental variability. However, the authors stopped short of actually identifying damage and only considered temperature variations to elaborate the undamaged scenarios. Conversely, Neves (2017) actually used data from undamaged and damaged FE models to train an artificial neural network, but no monitoring data were used. Finally, Malekzadeh et al. (2015) simulated the damage-sensitive features of a scaled-down laboratory bridge under four damage scenarios and showed that the acquired information was instrumental for the correct damage classification. However, the scale of the test structure was small, and the damage was particularly sleek because it was mainly induced by taking out screws. Moreover, environmental variability was not considered for the training of the damage-detection algorithms.

This paper presents a hybrid approach to detect damage in bridges under operational and environmental variations. The technique is hybrid in the sense that it integrates machine learning, FE modeling, and real monitoring data into a unique damage-detection approach. The MLAs were trained using different training strategies, namely, using only monitoring data as well as monitoring and FE data, and their performances in detecting damage were compared. The FE model was used as a proxy for environmental (negative temperature) and extreme loading events, and the data it provided were used together with monitoring data to enable the MLAs to learn the normal condition and to distinguish between the variations of the structural response induced by damage and environmental or operational variations. The technique was applied to 1-year monitoring data from a real bridge structure. To the extent of the authors' knowledge, this is the first study to combine monitoring data, FE model data, and MLAs to distinguish between the changes of damage-sensitive features induced by environmental and operational variability and, in contrast, damage on a real bridge structure.

Integrated Approach for Damage Detection

Fig. 1 presents the typical long-term damage-detection process, where trained MLAs continuously analyze structural-monitoring data to detect damage. The main challenge in this approach is to distinguish the changes in the data produced by damage from those caused by operational and environmental variability.

To enable this distinction, Fig. 2 presents the proposed MLA training process, where the visual inspection and the monitoring data are enhanced using data gathered from the FE model of the structure. The data produced by the FE model correspond to undamaged scenarios involving operational and environmental conditions that were not observed during the monitoring of the structure. Such scenarios may involve, for instance, extreme environmental and loading conditions.

To produce meaningful data, the FE model must be accurate, in the sense that it must reflect the bridge design project and the findings from the initial visual inspections of the structure, and must be calibrated to simulate correctly the early data acquired by the structural-monitoring systems.

Fig. 1. Conventional data-based damage-detection approach for bridges.

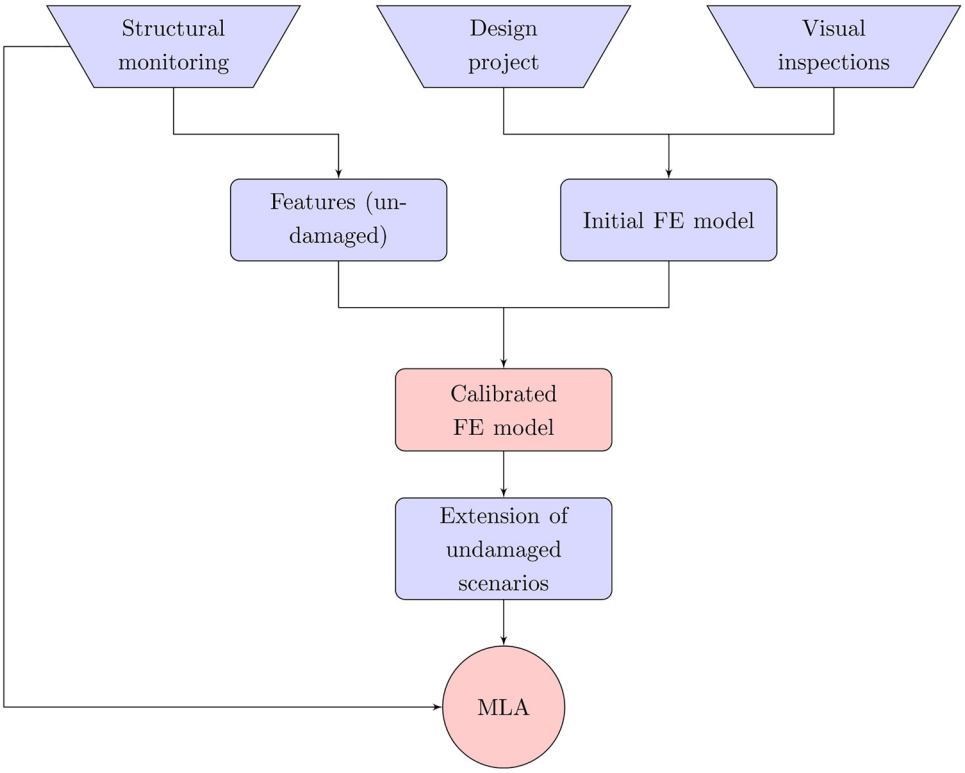


Fig. 2. Flowchart of the hybrid training process.

This approach assumes that the learning process is carried out in a static manner (i.e., the existence of an initial period to collect data for training is assumed) when the structure operates in its undamaged condition. Nevertheless, the authors acknowledge the possibility to implement a dynamic learning process, where the MLA is trained as long as new undamaged data come in.

After the completion of the training, the test process of the MLA involves the assessment of its ability to detect damage under scenarios that were not included in the training. Such scenarios may include data from the monitoring system and/or FE model. In particular, the FE model is useful to produce data from typical damaged scenarios (e.g., settlement of the foundations, spalling of the concrete), because such data normally cannot be retrieved from the monitoring system. The complexity of the FE model may increase when damage is introduced, inducing considerable uncertainty into the simulated data. However, even a simplified, linear FE model should be enough to assess the ability of the MLA to distinguish between the response of the undamaged structure subjected to environmental or operational conditions and that of the damaged structure.

MLA

In this study, the MLA was based on the Gaussian mixture model (GMM). The GMM was used to estimate the main clusters forming the normal structural condition; subsequently, the damage detection was performed on the basis of an outlier formation regarding the chosen clusters. The mathematical formulation of this algorithm was presented by Figueiredo and Cross (2013).

For general purposes, one may assume a training data matrix ($\mathbf{X} \in \mathbb{R}^{n\times d}$) with d-dimensional feature vectors from n different operational and environmental conditions when the structure is undamaged, and a test data matrix ($\mathbf{Z} \in \mathbb{R}^{l\times d}$) where l is the number of feature vectors from the undamaged and/or damaged conditions. The implementation follows a common sequence of steps. First, the algorithm is trained, and its parameters are adjusted using features from $\mathbf{X}$. Second, in the test phase, the algorithms transform each new input feature vector from $\mathbf{Z}$ into a damage indicator (DI).

Basically, a finite-mixture model $[p(\mathbf{X}|\mathbf{\Theta})]$ is the weighted sum of $K > 1$ components $[p(\mathbf{x}|\boldsymbol{\theta}_k)]$ in $\mathbb{R}^d$ (McLachlan and Peel 2000)

$$p(\mathbf{X}|\mathbf{\Theta}) = \sum_{k=1}^{K} \alpha_k p(\mathbf{x}|\boldsymbol{\theta}_k) \tag{1}$$

where $\mathbf{x}$ = d-dimensional data vector; and α_k corresponds to the weight of each component. These weights are constrained as $\alpha_k > 0$ with $\sum_{k=1}^{K} \alpha_k = 1$. For a GMM, each component $[p(\mathbf{x}|\boldsymbol{\theta}_k)]$ is represented as a Gaussian distribution

$$p(\mathbf{x}|\boldsymbol{\theta}_k) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\boldsymbol{\Sigma}_k)}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T \sum_{k}^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right\} \tag{2}$$

where each component is denoted by the parameter $\boldsymbol{\theta}_k = \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$ and is composed of the mean vector ($\boldsymbol{\mu}_k$) and the covariance matrix ($\boldsymbol{\Sigma}_k$). The GMM is completely specified by the set of parameters $\boldsymbol{\Theta} = \{\alpha_1, \alpha_2, \ldots, \alpha_K, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_K\}$.

The expectation-maximization (EM) algorithm is the most common method used to estimate the parameters of the GMMs (Dempster et al. 1977). This approach consists of an expectation step and a maximization step, which are applied until the log likelihood $[\log p(\mathbf{X}|\boldsymbol{\Theta}) = \log \prod_{i=1}^{m} p(\mathbf{x}_i|\boldsymbol{\theta})]$ converges to a local optimum (McLachlan and Peel 2000). The best model of a GMM is selected based on the Bayesian information criterion (BIC), which determines the appropriate number of structural components (Box et al. 2008).

In the damage-detection process and when the best model is selected, for each observation ($\mathbf{z}$), one needs to estimate K DIs. In particular, for each main component (k)

$$\mathrm{DI}_k(\mathbf{z}) = (\mathbf{z} - \boldsymbol{\mu}_k) \sum_{k}^{-1} (\mathbf{z} - \boldsymbol{\mu}_k)^T \tag{3}$$

For each component (k) from the undamaged condition, if a new observation ($\mathbf{z}$) is extracted from the same component, DI_k will be chi-square distributed with d degrees of freedom (χ^2_d). Finally, for each observation, the DI is given by the smallest DI_k estimated for each component

$$\mathrm{DI}(\mathbf{z}) = \min[\mathrm{DI}_1(\mathbf{z}), \ldots, \mathrm{DI}_K(\mathbf{z})] \tag{4}$$

Integration of the FE Model

The data matrix ($\mathbf{X}$) used to train the MLA is typically provided by a SHM system (Figueiras et al. 2011). However, such data matrix may be limited in scope, because it corresponds to the normal and undamaged condition of a structure in a limited time frame.

Therefore, in this paper, a FE model of the structure is proposed to generate data corresponding to undamaged and damaged scenarios that cannot be retrieved from early-stage monitoring and visual inspections. The FE modeling of undamaged scenarios enriches the data available for the training of the MLA by complementing the measured training data ($\mathbf{X}$) with the simulated training data ($\bar{\mathbf{X}}$). The predictions regarding damaged scenarios are used to validate the performance of the MLA, but the accurate modeling of severe damage configurations based on a calibrated FE model using only undamaged experimental data is probably unrealistic. Therefore, the approach suggested here only models *slight* deterioration of the structure, before the development of highly nonlinear effects caused by severe structural damage. Although this procedure is insufficient to provide data under severe damage conditions, it should nevertheless be enough to identify the *onset* of the damage.

The integration of the FE model into the MLA is graphically described in Fig. 2. It consists of three main phases. In the first phase, the FE model is created based on the design data and initial visual inspection of the structure. In the second phase, the FE model is tuned to recover some features ($\mathbf{x}$) selected from those measured by the SHM system under undamaged conditions ($\mathbf{x} \in \mathbf{X}$). In the third phase, the tuned model is used to produce the extended undamaged data ($\bar{\mathbf{X}}$). The three aforementioned stages are described next.

Initial FE Model

This stage consists of the creation of an initial FE model based on the data gathered from the design project of the structure and its initial visual inspection.

The design project is the main source of information for the creation of the FE model. It should provide rigorous information on the geometry of the structure, materials, and boundary conditions.

After the structure is completed, its visual inspection is recommended to identify any differences between the design project and the field reality. The nonstructural elements are of particular interest, because they may considerably influence the dynamic behavior of the structure. If the visual inspection is missing, the dimensions of the structural and nonstructural elements may need to be included in the model calibration process, albeit limited to slight variations around their design values. Conversely, a sound visual inspection eliminates the variability hypothesis of the structural geometry and leaves fewer parameters to be calibrated in the next phase.

As mentioned before, the approach suggested here opts for using simple models and searches for the variation of the structural features in the vicinity of the reference configuration. Therefore, a linear FE model should be sufficient for most applications.

Calibration of the FE Model

In general, the initial FE model does not recover precisely the structural features extracted from the SHM system. This happens because of the inherent imperfections in the construction of the structure; faulty idealization of the supports; and, in the case of the concrete structures, early-stage cracking.

Consequently, for the FE model to be usable for extending the range of training scenarios, it first needs to be calibrated to match the structural features recorded by the monitoring system.

The objective of this calibration is to find *plausible* variations of a set ($\mathbf{c}$) of calibration parameters such that the damage-sensitive features $[\bar{\mathbf{x}}(\mathbf{c})]$ obtained using the FE model *closely* recover some measured features ($\mathbf{x}$) extracted from the training data ($\mathbf{X}$). The calibration parameters may include the stiffness of the soil and structural elements, imperfections of the joints and supports, and dimensions of the structural elements, especially if the execution of the design project was not confirmed by visual inspection.

The input data of the calibration algorithm (Fig. 2) include the target data ($\mathbf{x}$), the initial values ($\mathbf{c}_0$) of the calibration parameters, and the simulated data $[\bar{\mathbf{x}}(\mathbf{c}_0)]$, consistent with the target data ($\mathbf{x}$).

The calibration problem requires canceling of the error function

$$\varepsilon(\mathbf{c}) = \mathbf{x} - \bar{\mathbf{x}}(\mathbf{c}) \tag{5}$$

subjected to a set of constraints ($a_i \le c_i \le b_i$) enforced to preserve the physical plausibility of the calibration parameters ($\mathbf{c}$).

The expansion of the error function in the vicinity of its initial value $[\varepsilon(\mathbf{c}_0)]$ reads

$$\varepsilon(\mathbf{c}) = \varepsilon(\mathbf{c}_0) + \sum_i \frac{\partial \varepsilon(\mathbf{c}_0)}{\partial c_i} \cdot dc_i = 0 \tag{6}$$

Taking into account that $\varepsilon(\mathbf{c}_0) = \mathbf{x} - \bar{\mathbf{x}}(\mathbf{c}_0)$, Eq. (6) generates the linear system

$$\sum_i \frac{\partial \varepsilon(\mathbf{c}_0)}{\partial c_i} \cdot dc_i = \bar{\mathbf{x}}(\mathbf{c}_0) - \mathbf{x} \tag{7}$$

which is solved to find the variations (dc_i) of the calibration parameters.

The elements of the gradient of the error function (ε) are computed by successively applying unit variations to each parameter (c_i) while keeping the other parameters $\left(c_j, j \neq i\right)$ unchanged

$$\frac{\partial \varepsilon(\mathbf{c})}{\partial c_i} = \varepsilon(c_1, c_2, \ldots, c_i + 1, \ldots) - \varepsilon(c_1, c_2, \ldots, c_i, \ldots)$$
$$= -\bar{\mathbf{x}}(c_1, c_2, \ldots, c_i + 1, \ldots) + \bar{\mathbf{x}}(c_1, c_2, \ldots, c_i, \ldots) \quad (8)$$

Eq. (7) is generally overconstrained and solved using a least-squares technique. Multiple solutions are obtained by using different perturbation parameters (w_i) to be applied to each equation of the solving system

$$\left[\mathbf{W} \cdot \frac{\partial \varepsilon(\mathbf{c}_0)}{\partial \mathbf{c}}\right] \cdot \mathbf{dp} = \mathbf{W} \cdot (\bar{\mathbf{x}}(\mathbf{c}_0) - \mathbf{x}) \quad (9)$$

where $\mathbf{W}$ = diagonal matrix that lists the perturbation parameters (w_i).

A set of tuned FE models is obtained in this way. The set can be subsequently filtered when new structural monitoring results become available or by using the training data in $\mathbf{X}$ that were not included in the calibration features ($\mathbf{x}$). Further filtering is also possible based on physical grounds (e.g., nearly symmetric parameter distribution in symmetric structures) or visual inspection (e.g., cracking and deterioration).

Extension of the FE Model

One or more of the tuned FE models are used at this stage to extend the MLA training data beyond the features ($\mathbf{X}$) directly obtained by the structural-monitoring system.

The extensions may include undamaged scenarios that were not directly measured, but may also include a range of key damaged scenarios for the test process of the MLA. Typical undamaged extensions may include exposure to low temperature conditions, which is highly relevant in bridges due to the pronounced stiffening of the asphalt, or the application of special loading conditions, such as traffic.

All extensions should be confined to the vicinity of the model calibrated according to the previous section. In this way, the same FE model can be used to extract the variation of the damage-sensitive features with minimal modification. Although the nonlinear effects of severe damage are clearly not accounted for, this strategy should still enable the recognition of the initial deviation of the damage-sensitive features in the feature space. This information is relevant for a continuous SHM system, enabling it to formulate (typically multiple) damage hypotheses, if its damage-sensitive features change during the lifetime of the structure.

Test Structure, Monitoring Data, and FE Modeling

Test Structure: Z-24 Bridge

The Z-24 Bridge was a post-tensioned concrete box girder bridge composed of a main span of 30 m and two side spans of 14 m, as depicted in Fig. 3(a). The bridge, before complete demolition, was extensively instrumented and tested with the purpose of providing a feasibility benchmark for vibration-based SHM. A long-term monitoring program was carried out, from November 11, 1997, until September 10, 1998, to quantify the operational and environmental variability present on the bridge and to detect damage (pier settlement, foundation tilt, spalling of concrete, landslide at abutment, concrete hinge failure, failure of anchor heads, and rupture of tendons) artificially introduced in the last month of operation. Every hour, eight accelerometers captured the vibrations of the bridge as sequences of 65,536 samples (sampling frequency of 100 Hz), and other sensors measured environmental parameters, such as temperature at several locations (Peeters and de Roeck 2001). Fig. 3(b) presents the first four natural frequencies along with the ambient temperature. The natural frequencies seen in Fig. 3(b) were estimated using a reference-based stochastic subspace-identification method on the vibration measurements from the accelerometers (Peeters and de Roeck 1999); the natural frequencies were split into baseline condition (1–3,470) and damaged condition (DC; 3,471–3,932). The jumps in the natural frequencies were related to the asphalt layer in cold periods, which significantly contributed to the stiffness of the bridge.

FE Modeling

Conforming displacement FEs were used to model the test structure. This is, by far, the most common type of FEs, implemented in a wide range of commercial structural-analysis software. Conforming displacement elements use shape functions that strongly satisfy the compatibility equations in the domain of each element and the displacement continuity conditions at the interface of neighboring elements. However, the equilibrium equations are weakly satisfied in the domain of the elements, and the force balance between adjacent elements is only satisfied at their nodes. This means that the quality of the displacement approximation is generally superior to that of the stress approximation, although techniques to improve the latter in the postprocessing phase have also been developed (Grätsch and Bathe 2005).

The general description of the FE model is followed by a more detailed presentation of the calibration of some of its parameters to better recover the measured natural frequencies of the structure, according to the procedure presented in the "Integrated Approach for Damage Detection" section.

General Description of the FE Model

The Z-24 Bridge was modeled using the commercial FE platform ROBOT. All structural elements except for the six abutment pillars were modeled using three- and four-node (thin) shell FEs. The abutment pillars were modeled with beam elements, with six degrees of freedom per node. The mesh was strictly conforming. An overview of the structural model is given in Fig. 4, and a cross section through the model is presented in Fig. 5.

The foundations of the bridge were modeled as elastic supports using displacement springs at the soil–structure interface (Fig. 5). The stiffness of the springs was taken as prescribed by Masciotta et al. (2016). The abutment pillars and the abutment itself were completely embedded into the soil. The asphalt layer was explicitly included in the model, because its stiffness increment justified the increase of the natural frequencies registered when the environment temperature dropped below 0°C. The asphalt was modeled using shell elements and was connected to the underlaying upper plate of the girder using rigid connectors (Fig. 5).

The mechanical characteristics of the undamaged concrete were also taken from Masciotta et al. (2016) and were $E_0 = 37.5$ GPa for Young's modulus and $\nu = 0.2$ for Poisson's ratio. In the pier settlement damage scenarios, cracking was reported in the bridge girder in the vicinity of the affected pier (Masciotta et al. 2016). The cracking of the concrete was modeled by a reduction of Young's modulus of the concrete. Young's modulus of the asphalt layer is strongly affected by the environment temperature. To enable the modeling of this effect, the (cubic polynomial) stiffness-temperature model proposed by Watson and Rajapakse (2000) was used in this work.

The leading dimension of the FEs was established by analyzing the variation of the natural frequencies of the structure with the

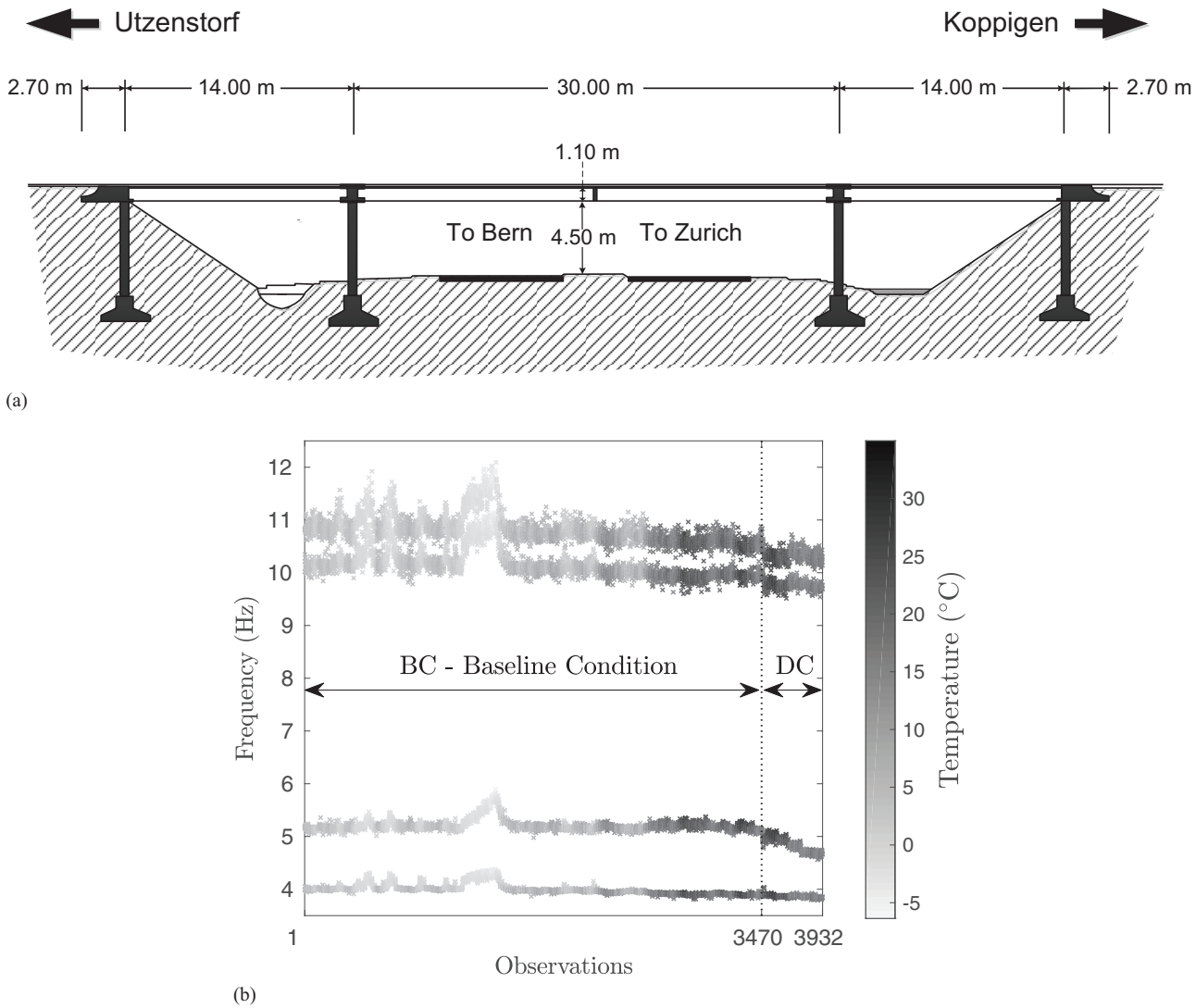


Fig. 3. (a) Longitudinal section from bridge design blueprints; and (b) first four natural frequencies extracted from hourly monitoring data. DC = damage conditions.

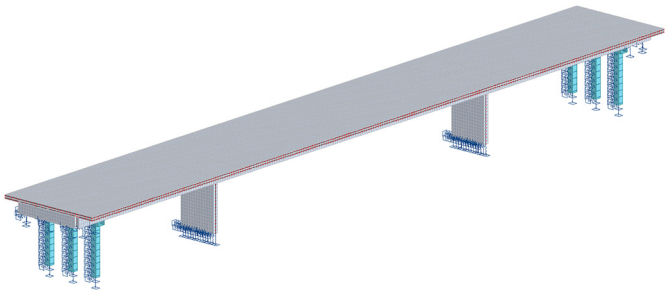


Fig. 4. General perspective of the Z-24 Bridge model.

refinement of the mesh. Elements between 0.2 and 0.5 m were considered, because finer meshes would considerably increase the (spurious) stress concentration effect in the vicinity of the pier-girder and pillar-girder joints. A stable eigenfrequency prediction without excessive stress concentrations is achieved with a mesh of elements with a leading dimension of 0.25 m (Fig. 5), which was therefore adopted in all analyses reported here.

Calibration of the FE Model

The information regarding the geometry of the Z-24 Bridge was compiled using various sources (e.g., Masciotta et al. 2016; Peeters 2000; Peeters and de Roeck 2001). Slight uncertainties remained, however, regarding the thickness of the plates of the girder in the midspan (t_{p1}) and, especially, in the vicinity of the supporting piers (t_{p2}), where they were reported to be thicker (Masciotta et al. 2016). Also, the exact thickness of the asphalt layer (t_a) was unclear, because measurements suggested it was 16 cm, but it was reported to be only half of that in the construction drawings (Peeters 2000). Therefore, instead of calibrating the overall bending stiffness of the girder along its length, as, for instance, was done by Masciotta et al. (2016), these three thicknesses were used as calibrating parameters for the computational model.

In the first phase of the calibration process, we searched for calibration parameters capable of recovering *well* the natural frequencies measured in the undamaged structure, according to the procedure presented in the "Integrated Approach for Damage Detection" section. The target natural frequencies were calculated as averages of the frequency measurements corresponding to positive temperatures. Four plausible combinations of calibration parameters were found, which were further filtered using the negative temperature data

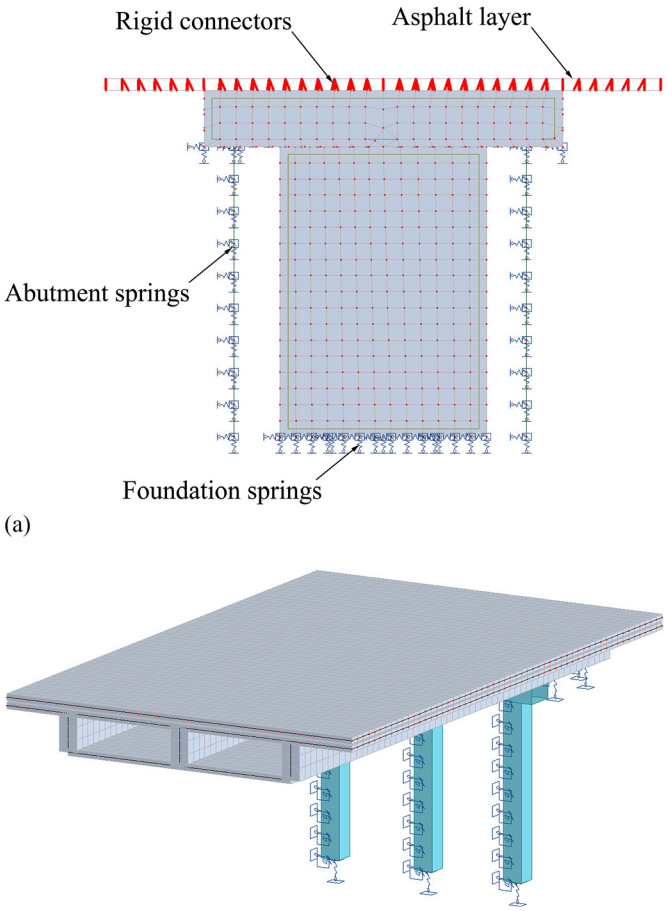


Fig. 5. Cross sections through the Z-24 Bridge model: (a) asphalt layer detail; and (b) deck detail.

recorded over the year-long monitoring period. Upon the completion of this process, the FE model which presented the highest accuracy in modeling the effect of the negative temperatures on the behavior of the structure, was used to assess the deterioration of the concrete stiffness in the girder near the pier where the settlements were applied.

The first phase of the calibration process consisted of tuning the parameters t_{p1}, t_{p2}, and t_a to satisfactorily recover the first four natural frequencies measured on the undamaged structure under positive environment temperatures. The target frequencies were obtained by taking the average of their measurements over the 1-year monitoring period ($f_m = \{3.90, 5.18, 9.97, 10.60\}$ Hz).

The calibration started from the initial estimates of t_{p1} = 15 cm, t_{p2} = 20 cm, and t_a = 16 cm, to which corresponded the following frequency estimates ($f = \{3.90, 5.22, 9.65, 10.50\}$ Hz). The model-updating procedure presented in the "Integrated Approach for Damage Detection" section was applied using various choices of perturbation parameters. The solutions, rounded to integer values, were inserted into the numerical model, and new simulations were ran to compute the natural frequencies.

To assess the accuracy of the natural frequency predictions, an error measure is defined as

$$\epsilon = \sum_{i=1}^{4} w_i \cdot |f_{m_i} - f_i| \tag{10}$$

where w = {0.5, 0.3, 0.15, and 0.05} = weighting parameters designed to reflect the higher importance of the first natural frequencies; and $|f_{m_i} - f_i|$ = absolute value of the difference between the measured and estimated natural frequencies.

Following this strategy, four combinations of the calibration parameters were found to yield error measures [Eq. (10)] inferior to 1.0% and were selected as candidates for the second calibration phase. They are listed, together with their respective frequencies and errors, in Table 1.

The second stage of the calibration process consisted of simulating the effect of the freezing of the asphalt mixture on the natural frequencies of the bridge. The four candidate combinations presented in Table 1 were tested on six temperatures between 0°C and –5°C, with a –1°C increment. The simulated natural frequencies were extracted for each case and compared to those measured during the daily observations presented in Fig. 3. The error measures [Eq. (10)] were calculated for each temperature, and their average over the six tested cases was computed for each combination of calibration parameters. The average error measures are presented in Table 2.

The combination t_{p1} = 18 cm, t_{p2} = 19 cm, and t_a = 17 cm presented a good overall behavior in the first two phases of the calibration process and was therefore selected and used in the next phase of the calibration process and, subsequently, in all simulations reported in the remainder of this paper.

The last phase of the calibration process was aimed at quantifying the deterioration of Young's modulus of the concrete in the girder, near the pier where settlements were applied. The loss of stiffness in that region was inversely proportional to the vertical displacement in the middle of the main span of the bridge. The mid-span displacements caused by the pier settlements were measured and summarized in the long-term monitoring and bridge tests report (EMPA 1999), and were used here to estimate Young's modulus of the cracked zone. It is noted that the extension of this zone was taken from Masciotta et al. (2016).

The calibration process was straightforward, because the degradation of Young's modulus for a given settlement was the only calibration parameter. Table 3 presents the stiffness estimates of the cracked zone of the girder for the tested pier settlements.

Table 1. Calibration parameters selected after the first stage

t_{p1}	t_{p2}	t_a	f_1	f_2	f_3	f_4	ϵ (%)
17	19	17	3.88	5.19	9.74	10.60	0.88
17	20	16	3.89	5.22	9.75	10.61	0.91
18	19	17	3.89	5.18	9.78	10.65	0.63
18	20	16	3.89	5.21	9.80	10.65	0.74
	Target		3.90	5.18	9.97	10.60	—

Table 2. Average error measures under negative temperatures

t_{p1}	t_{p2}	t_a	Average ϵ (%)
17	19	17	3.36
17	20	16	3.05
18	19	17	2.95
18	20	16	3.22

Table 3. Stiffness estimates as functions of initial Young's modulus (E_0)

Pier settlement (cm)	Stiffness estimate (E_0)
2.0	0.85
4.0	0.70
8.0	0.41
9.5	0.30

Results and Discussion

Extraction of Damage-Sensitive Features

Because the third and fourth natural frequencies were highly correlated with the first one (correlation coefficients higher than 0.90), and for feature visualization and interpretation purposes, only the first two natural frequencies (measured daily at 5 a.m.) were used as damage-sensitive features. The natural frequencies extracted from the monitoring data and the ones from the FE model are depicted in Fig. 6, in concatenated format, and summarized in Table 4. Each observation represents a state condition, which changes as a function of the operational and environmental variations (undamaged condition) or by the presence of damage (damaged condition).

Observations 1–221 corresponded to the undamaged condition. In particular, Observations 1–198 corresponded to natural frequencies estimated from the monitoring data under operational and environmental variability; the observed jumps were related to the stiffness increment of the asphalt layer in cold periods, which significantly contributed to the stiffness of the bridge. Observation 199 corresponded to the state condition based on the average natural frequencies for positive ambient temperature. Observations 200–221 corresponded to data obtained from the FE model by adding mass (in the form of forces to represent traffic) and simulating the temperature variations (by changing the stiffness).

In contrast, Observations 222–268 corresponded to the damaged condition. Observations 222–258 were associated with natural frequencies estimated from the monitoring data, when the structure was progressively damaged. Observations 259–268 corresponded to extensions of the FE model in the vicinity of the reference state (see the "Integrated Approach for Damage Detection" section); to challenge the damage-detection process, all of these damage scenarios were modeled along with added mass. All structural scenarios behind these extensions are listed in Table 4. Note that even though the FE model was used to predict damage scenarios related to pier settlement, those observations were only taken into account in the test matrix to test the effectiveness of the learning process.

The damage-detection process was carried out assuming three different training strategies (Training 1, Training 2, and Training 3). Training 1 ($\mathbf{X}^{198\times2}$) only accounted for the observations from the monitoring data (1–198) and when the structure was undamaged. Training 2 ($\mathbf{X}^{216\times2}$) accounted for the monitoring and FE model data (1–216), when the structure was undamaged. The undamaged

Table 4. Undamaged and damaged scenarios from monitoring and FE modeling data

Observations	Condition	Scenario description	Source
1–198	Undamaged	Environmental variability	MON
199	Undamaged	Average positive temperature	MON
200–202	Undamaged	P = 20, 60, 100 kN (1/2 central span)	FE
203–205	Undamaged	P = 20, 60, 100 kN (1/2 lateral span)	FE
206–216	Undamaged	Temperature variations (from –5°C to 0°C)	FE
217, 218	Undamaged	P = 50, 120 kN (1/2 central span)	FE
219, 220	Undamaged	P = 50, 120 kN (1/2 lateral span)	FE
221	Undamaged	Temperature variation (–3.08°C)	FE
222–258	Damaged	Several damage scenarios, including pier settlement	MON
259–262	Damaged	Pier settlement (2, 4, 8, 9.5 cm)	FE
263	Damaged	Scenarios 259 and 202 (100 kN at 1/2 central span)	FE
264	Damaged	Scenarios 260 (4 cm) and 201 (60 kN at 1/2 central span)	FE
265, 266	Damaged	Scenario 261 with scenarios 200 and 205	FE
267, 268	Damaged	Scenario 262 with scenarios 201 and 205	FE

Note: Undamaged and damaged scenarios from monitoring and FE modeling data. MON = monitoring.

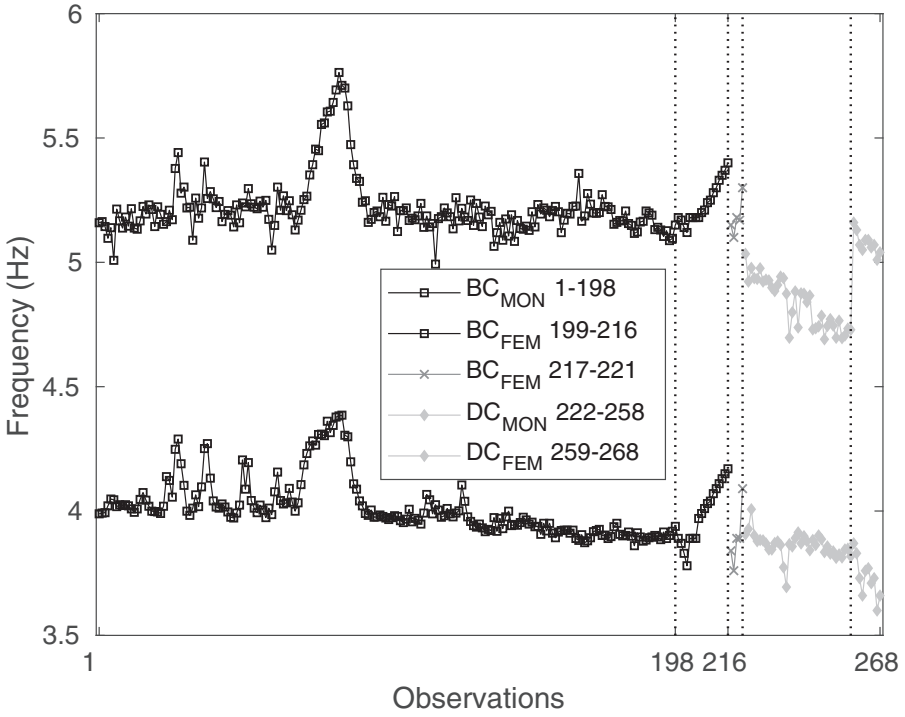
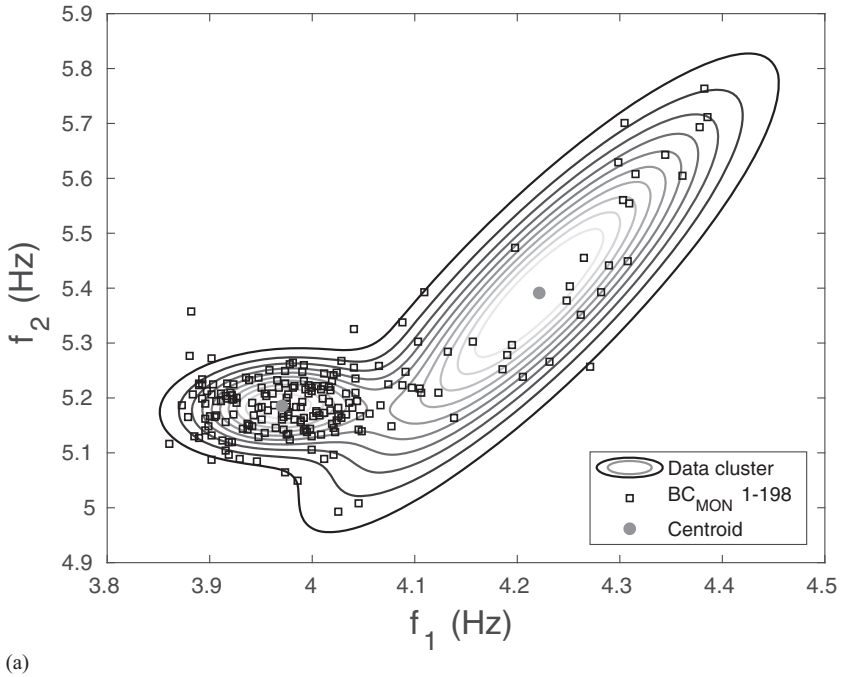


Fig. 6. First two natural frequencies extracted from monitoring system and FE model. BC = baseline condition.

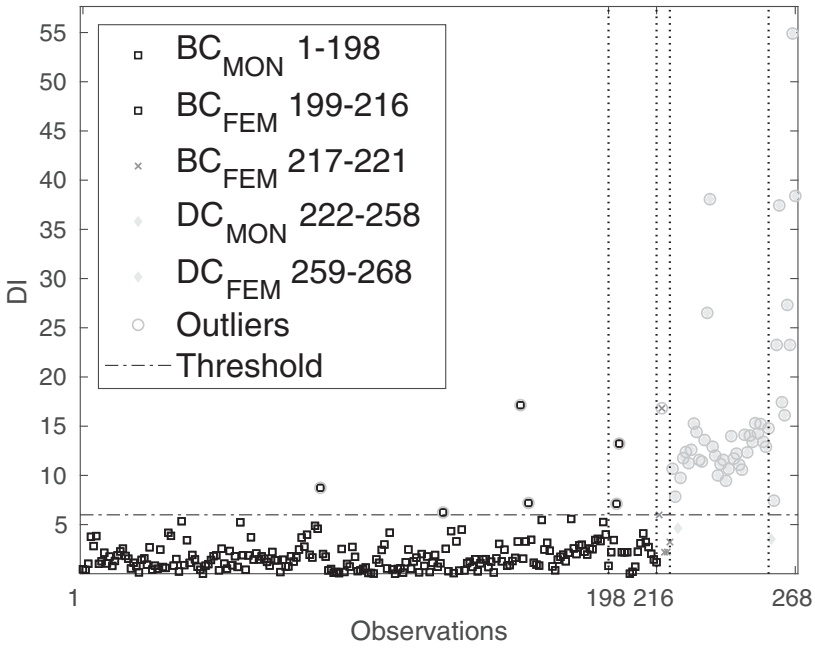
Observations 217–221 were not included in the training process to verify the performance of the MLA in classifying undamaged observations not used in the training process. Finally, Training 3 ($\mathbf{X}^{221\times2}$) accounted for the all of the observations from the undamaged state conditions. All observations were used to test the algorithms for all three strategies ($\mathbf{Z}^{268\times2}$). The damage-detection performances of the GMM-based algorithm, for each training strategy, were compared in terms of Type I/Type II errors. The classification performance was evaluated on the basis of Type I/Type II error trade-offs. In SHM, in the context of damage detection, a Type I error is a false-positive indication of damage, and a Type II error is a false-negative indication of damage. The GMM was set as described by Figueiredo and Cross 2013.

Learning Process Based on Monitoring Data Only: Training 1

The clustering performance of the GMM-based algorithm for the training matrix (composed of all undamaged monitoring data, 1–198) is presented in Fig. 7(a). Two clusters were formed when only the monitoring undamaged data were considered to train the algorithm: the cluster with a higher concentration of observations was related to the baseline condition obtained under relatively small environmental and operational influences, and the other cluster was assigned to the baseline condition under freezing temperature variations, which increased the structural stiffness (Peeters and de Roeck 2001).



(a)



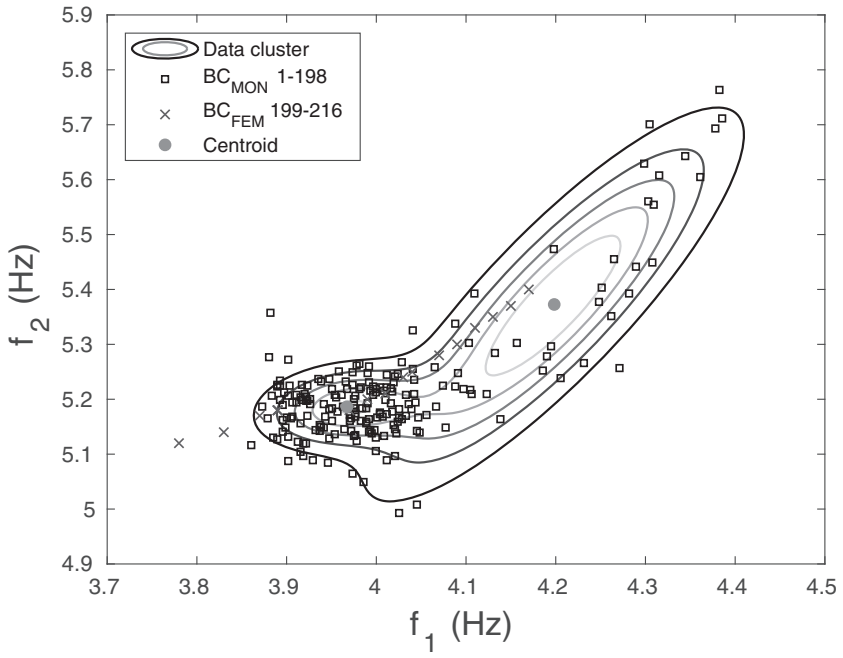
(b)

Fig. 7. Training 1: (a) clustering performance based on undamaged monitoring data; and (b) outlier detection. BC = baseline condition.

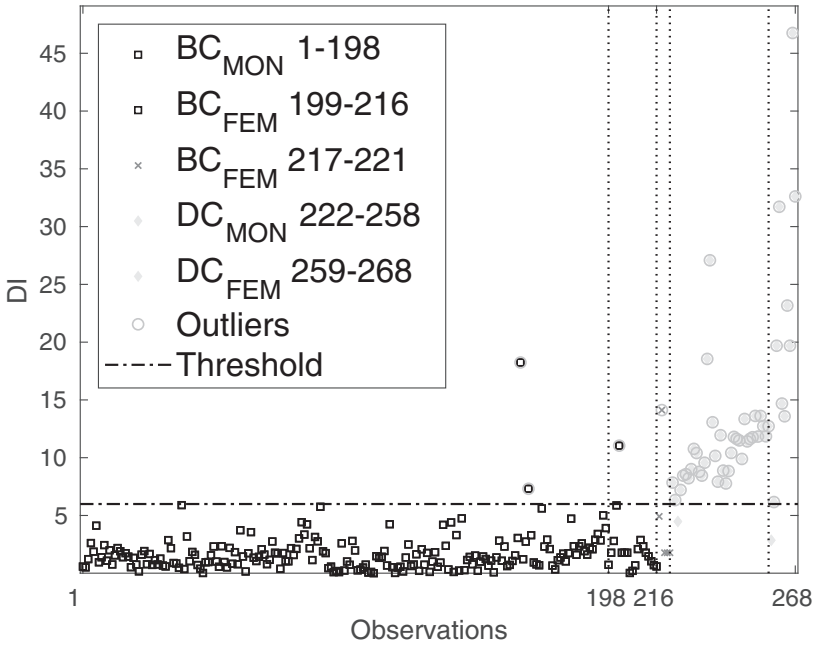
The DIs obtained from the observations in the test matrix ($\mathbf{Z}^{268\times 2}$) along with a threshold based on a level of significance of 5% are depicted in Fig. 7(b). The number of Type I and Type II errors are summarized in Table 5. In the observations in the range of 1–198, used in the training phase, there were four Type I errors as expected due to the probabilistic nature of the defined threshold; note that jumps caused by the freezing effects were significantly attenuated, which means the algorithm was able to normalize the training observations and remove the freezing temperature effects. In terms of classification performance, and for the observations simulated using the FE model, the algorithm output three Type I errors associated with the addition of mass (201, 202, and 218) and one Type II error associated with an observation from a damaged state condition (259). In particular, Type I errors associated with the addition of mass were expected because similar state conditions were not used in the training

Table 5. Damage detection for each training strategy in terms of Type I and Type II errors

	MON		FEM		
Strategy	Type I	Type II	Type I	Type II	Total
Training 1	4	1	3	1	9
Training 2	2	1	2	1	6
Training 3	2	1	2	0	5



(a)



(b)

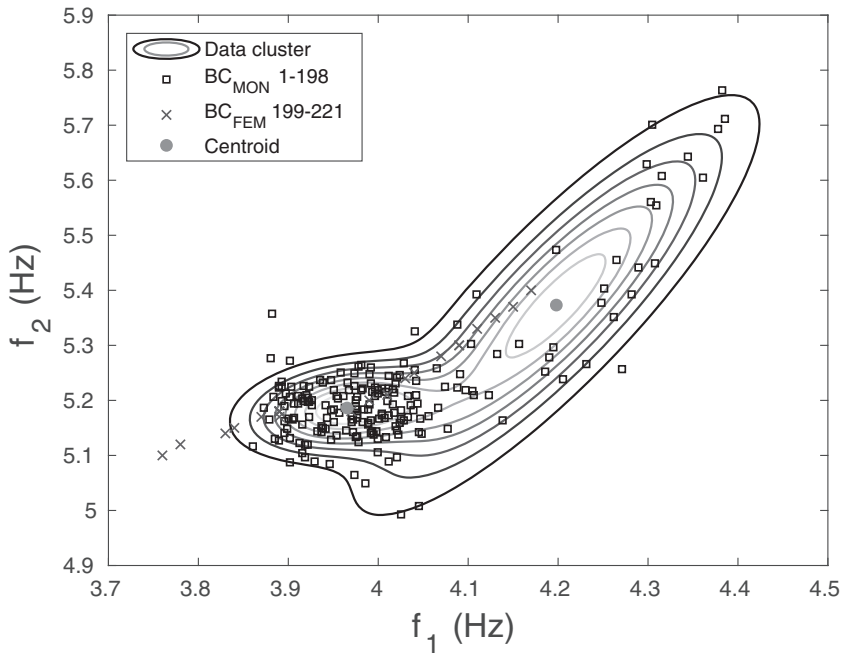
Fig. 8. Training 2: (a) clustering performance based on undamaged monitoring and FE model data; and (b) outlier detection. BC = baseline condition.

matrix, because the bridge had no traffic during the 1-year monitoring program.

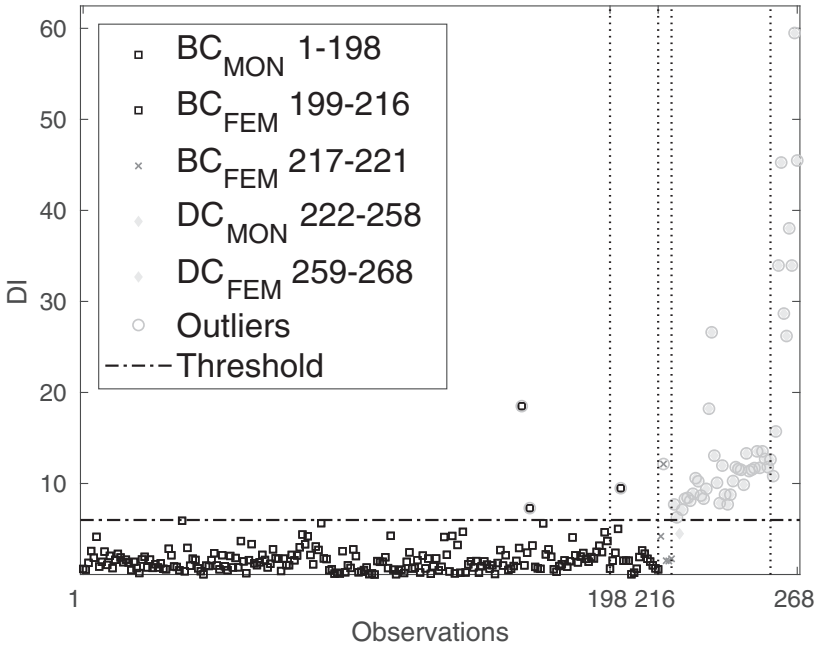
Training Process Based on Both Monitoring and FE Model Data: Training 2

To improve the learning process and the classification performance of the GMM-based algorithm, as originally proposed in this paper, the training matrix was then composed of both monitoring and FE model data ($\mathbf{X}^{216\times2}$). The new data set also included undamaged observations related to the addition of mass and temperature changes (in the range of –5°C to 0°C) simulated by the FE model. Some undamaged observations (217–221), from the FE model, were not included in the training matrix to test the robustness of the algorithm to classify undamaged observations similar to those used in the training process. The clustering performance is presented in Fig. 8(a). The FE model data with temperature changes were within the two clusters found, which showed the effectiveness of the FE model to mimic the response of the bridge.

The DIs obtained from the test matrix ($\mathbf{Z}^{268\times2}$) along with a threshold based on a level of significance of 5% are depicted in Fig. 8(b); Table 5 summarizes the Type I and Type II errors. Comparing Training strategies 1 and 2, one can infer some improvements in the damage-classification accuracy in terms of Type I errors.



(a)



(b)

Fig. 9. Training 3: (a) clustering performance based on undamaged monitoring and FE model data; and (b) outlier detection. BC = baseline condition.

Actually, it is possible to observe an improvement of the characterization of the baseline condition when the structure is affected by temperature changes, because the number of Type I errors for the monitoring data decreased from four to two, which means the FE model embedded relatively well the stiffness of the bridge for freezing temperatures (Observations 90 and 136 were correctly classified as undamaged ones); additionally, Observation 201 with simulated mass changes was then correctly classified as undamaged.

Training Process Based on All Undamaged Data from Monitoring and FE Model: Training 3

In pattern recognition and machine learning, in general, the more data that are available in the training process, the better the classification performance is. Therefore, in an attempt to improve the classification performance (by reducing the number of both Type I and Type II errors), all observations from the undamaged condition (both monitoring and FE model) were used in the training process ($\mathbf{X}^{221\times2}$). The clustering performance is presented in Fig. 9(a), and the DIs obtained from the test matrix ($\mathbf{Z}^{268\times2}$) along with the threshold based on the same level of significance are depicted in Fig. 9(b). Table 5 also summarizes the Type I and Type II errors.

Comparing the this training strategy (Training 3) with the previous one (Training 2), one can observe a single improvement; in particular, Observation 259 (pier settlement of 2 cm) was correctly classified as damaged. This improvement was associated with the refining of the learning process through data from other operational and environmental conditions. Nevertheless, the FEM observations from mass changes were not correctly classified (202 and 218), which may be related to the fact that the number of observations used in the training process was not enough to significantly change the learning parameters of the MLA.

Conclusions

This study explored the integration of FE modeling and machine-learning techniques into a hybrid approach for damage detection. This approach can be applied to any structure but is particularly useful in large-scale civil engineering structures, where the data provided by the SHM system is typically limited in scope (undamaged conditions under regular loading). Here, the FE model was used first to provide simulated data from scenarios not recorded in the monitoring sessions. These scenarios included the response of the undamaged structure under operational and environmental variability. The enriched data were then fed into a MLA to improve the damage classification. Second, damaged FE data were also generated to test the MLA.

It was shown that a FE model, calibrated with monitoring data, is able to recover the measured data with sufficient precision, as well as to generate new data from undamaged and damaged scenarios. The results showed that the MLA is able to correctly classify undamaged observations generated by the FE model, even when they are not used in the training process; additionally, they also show that the damage-detection performance is improved when the learning process incorporates data from the FE model. In conclusion, this study confirmed the effectiveness of the FE model to generate structural scenarios not measured by monitoring systems and to improve the learning process of MLAs for damage detection.

References

Barthorpe, R. J. 2010. "On model- and data-based approaches to structural health monitoring." Ph.D. thesis, Univ. of Sheffield.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 2008. *Time series analysis: Forecasting and control.* 4th ed. Hoboken, NJ: John Wiley & Sons, Inc.

Catbas, F. N., H. B. Gokce, and D. M. Frangopol. 2013. "Predictive analysis by incorporating uncertainty through a family of models calibrated with structural health-monitoring data." *J. Eng. Mech.* 139 (6): 712–723. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000342.

Charles, R. F., S. W. Doebling, and D. A. Nix. 2001. "Vibration–based structural damage identification." *Philos. Trans. R. Soc. London, Ser. A* 359 (1778): 131–149. https://doi.org/10.1098/rsta.2000.0717.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum likelihood from incomplete data via the EM algorithm." *J. R. Stat. Soc. Ser. B Stat. Method.* 39 (1): 1–38. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

EMPA. 1999. *Long term monitoring and bridge tests. Rep. No. 168349/20e.* Dübendorf, Switzerland: Project SIMCES.

Figueiredo, E., and E. Cross. 2013. "Linear approaches to modeling nonlinearities in long-term monitoring of bridges." *J. Civ. Struct. Health Monit.* 3 (3): 187–194. https://doi.org/10.1007/s13349-013-0038-3.

Figueiredo, E., I. Moldovan, and M. B. Marques. 2013. *Condition assessment of bridges: Past, present, and Future—A complementary approach.* Lisboa, Portugal: University Católica Editora.

Figueiredo, E., G. Park, C. R. Farrar, K. Worden, and J. Figueiras. 2011. "Machine learning algorithms for damage detection under operational and environmental variability." *Struct. Health Monit.* 10 (6): 559–572. https://doi.org/10.1177/1475921710388971.

Grätsch, T., and K. Bathe. 2005. "A posteriori error estimation techniques in practical finite element analysis." *Comput. Struct.* 83 (4–5): 235–265. https://doi.org/10.1016/j.compstruc.2004.08.011.

Gresil, M., B. Lin, Y. Shen, and V. Giurgiutiu. 2011. "Predictive modelling of space structures for SHM with multiple PWAS transducers." In *Proc., ASME 2011 Conf. on Smart Materials, Adaptive Structures and Intelligent Systems.* Scottsdale, AZ: ASME.

Liu, Y., and S. Zhang. 2017. "Probabilistic baseline of finite element model of bridges under environmental temperature changes." *Comput.-Aided Civ. Infrastruct. Eng.* 32 (7): 581–598. https://doi.org/10.1111/mice.12268.

Malekzadeh, M., G. Atia, and F. N. Catbas. 2015. "Performance-based structural health monitoring through an innovative hybrid data interpretation framework." *J. Civ. Struct. Health Monit.* 5 (3): 287–305. https://doi.org/10.1007/s13349-015-0118-7.

Masciotta, M.-G., L. F. Ramos, P. B. Lourenço, M. Vasta, and G. de Roeck. 2016. "A spectrum-driven damage identification technique: Application and validation through the numerical simulation of the Z24 Bridge." *Mech. Syst. Signal Process.* 70–71: 578–600. https://doi.org/10.1016/j.ymssp.2015.08.027.

McLachlan, G. J., and D. Peel. 2000. *Finite mixture models.* New York: John Wiley & Sons, Inc.

Mirzaee, A., R. Abbasnia, and M. Shayanfar. 2015. "A comparative study on sensitivity-based damage detection methods in bridges." *Shock Vib.* 2015: 120630. https://doi.org/10.1155/2015/120630.

Neves, C. 2017. "Structural health monitoring of bridges: Model-free damage detection method using machine learning." Licentiate thesis in structural engineering and bridges, KTH Royal Institute of Technology, School of Architecture and the Built Environment.

Peeters, B. 2000. "System identification and damage detection in civil engineering." Ph.D. thesis, Katholieke Univ. Leuven.

Peeters, B., and G. de Roeck. 1999. "Reference-based stochastic subspace identification for output-only modal analysis." *Mech. Syst. Signal Process.* 13 (6): 855–878. https://doi.org/10.1006/mssp.1999.1249.

Peeters, B., and G. de Roeck. 2001. "One-year monitoring of the Z24-Bridge: Environmental effects versus damage events." *Earthquake Eng. Struct. Dyn.* 30 (2): 149–171.

Santos, A., E. Figueiredo, and J. Costa. 2015. "Clustering studies for damage detection in bridges: A comparison study." In *Proc., 10th Int. Workshop on Structural Health Monitoring*, 1165–1172. Stanford, CA: Stanford Univ.

Santos, A., E. Figueiredo, M. F. M. Silva, C. S. Sales, and J. C. W. A. Costa. 2016. "Machine learning algorithms for damage detection: Kernel-based approaches." *J. Sound Vib.* 363: 584–599. https://doi.org/10.1016/j.jsv.2015.11.008.

Smith, I. F., and S. Saitta. 2008. "Improving knowledge of structural system behavior through multiple models." *J. Struct. Eng.* 134 (4): 553–561. https://doi.org/10.1061/(ASCE)0733-9445(2008)134:4(553).

Sohn, H. 2007. "Effects of environmental and operational variability on structural health monitoring." *Philos. Trans. R. Soc. A* 365 (1851): 539–560. https://doi.org/10.1098/rsta.2006.1935.

Watson, D. K., and R. K. N. D. Rajapakse. 2000. "Seasonal variation in material properties of a flexible pavement." *Can. J. Civ. Eng.* 27 (1): 44–54. https://doi.org/10.1139/l99-049.

Wursten, E. R. G., and G. D. Roeck. 2014. "Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification." *Struct. Health Monit.* 13 (1): 82–93. https://doi.org/10.1177/1475921713502836.

Zienkiewicz, O. C., R. L. Taylor, and J. Zhu. 2013. *The finite element method: Its basis and fundamentals.* Amsterdam, Netherlands: Elsevier.