P3P – A software suite for autonomous SHM of bridge networks

Enrique García-Macías[1,2], Antonello Ruccolo[3], Mariano Angelo Zanini[4], Carlo Pellegrino[4], Carmelo Gentile[3*], Filippo Ubertini[1] and Paolo Mannella[5]

[1]Department of Civil and Environmental Engineering, University of Perugia, Via G. Duranti, 93, Perugia, 06125, Italy.
[2]Department of Structural Mechanics and Hydraulic Engineering, University of Granada, Av. Fuentenueva sn, Granada, 18002, Spain.
[3]Department of Civil and Environmental Engineering (DICA), Politecnico di Milano, Milan, 20133, Italy.
[4]Department of Civil Environmental and Architectural Engineering, University of Padova, Via Marzolo 9, Padova, 35131, Italy.
[5]Directorate of Operation and Territorial Coordination, ANAS S.p.A., Via Marsala, 27, Roma, 00185, Italy.

*Corresponding author(s). E-mail(s): carmelo.gentile@polimi.it;
Contributing authors: enrique.garciamacias@unipg.it; antonello.ruccolo@polimi.it; marianoangelo.zanini@dicea.unipd.it; carlo.pellegrino@unipd.it; filippo.ubertini@unipg.it; p.mannella@stradeanas.it;

Abstract

This paper presents the development of a new software code for the fully autonomous management of permanent integrated Structural Health Monitoring (SHM) systems installed in highway bridge structures. The code was developed within the framework of a research project funded by Anas S.p.A., the largest public infrastructure manager in Italy. The software program includes all the necessary steps to conduct SHM within the statistical pattern recognition paradigm, including automated dynamic identification, modal tracking, filtering of environmental effects, and damage detection through novelty analysis. Additionally, the software suite includes specific modules for processing and analysis of seismic events and structural reliability analysis of bridges, as well as specific functionalities for span-wise identification of long multi-span bridges. Moreover, a novel automated density-based tracking algorithm is developed. The potential of P3P is illustrated through two real application case studies: (i) a long multi-span bridge, the Trigno V Bridge in Italy; and (ii) the Z-24 Bridge benchmark. This work demonstrates the effectiveness of the developed code for handling large monitoring databases within the framework of SHM as a statistical pattern recognition, and

currently P3P is in phase of being applied by Anas S.p.A for the management of a large number of bridges of the Italian roadway system.

Keywords: Bridges, Damage detection, OMA, Reliability, Seismic analysis, SHM Software development.

1 Introduction

Recent tragic collapses like the I-35 W Mississippi river in 2007 or the Genoa bridge in 2018 have revealed the fundamental challenge posed by ageing civil infrastructure and the need to prioritize its management in the political agenda. A remarkable evidence of this is the last Infrastructure Report Card recently released in 2021 by the American Society of Civil Engineers (ASCE) [1], which detected that 7.5% of the more than 600,000 American highway bridges are in poor conditions and estimated the nation's backlog of bridge repair at $123 billion. The EU-funded BRIME project [2] in 2001 revealed a similar picture across the European highway bridges, with deficiency rates among the surveyed countries ranging between 26 and 39%. This has promoted a vast volume of research on SHM in the last two decades, as well as the publication of a multitude of technical standards of SHM worldwide [3]. These comprise from the first SHM guide in 2001 by the SIS Canada Research Network [4], until the last guidelines of the monitoring of viaducts published by the Italian Ministry of Infrastructure and Transport [5]. There is also a growing awareness on the limitations of classical periodic bridge maintenance policies based on visual inspections to provide efficient assessment of early-stage pathologies. It is worth noting the survey report published by the US Department of Transportation in 2001 [6], which concluded that inspection reports may considerably differ depending on the lighting conditions, fear of traffic, and experience of the inspectors. Nonetheless, despite the large research efforts made in this field, the routine implementation of SHM remains marginal and classical periodic bridge maintenance policies continue to be the only basis for performance assessment [7]. Among the reasons explaining the slow technological transfer of SHM to industry, it is worth stressing the lack of standardized solutions for continuous data processing and automated data-driven decision-making, requiring in many cases the assistance of highly specialized professionals and in-house software platforms.

In its broadest sense, SHM can be defined as the process of implementing a continuous damage identification strategy for an engineering infrastructure. Damage can be conceived in this context as any disruptive event affecting the system's performance such as, for instance, alterations of the material and/or geometrical properties, or changes to the boundary conditions/connectivity [8]. In this framework, ambient vibration-based SHM has received most attention owing to its minimal intrusiveness, non-destructive character, and global damage identification capabilities [9]. These techniques exploit the ambient

vibrations of mechanical systems under normal operating conditions to extract their modal features (i.e. resonant frequencies, damping ratios and mode shapes) through Operational Modal Analysis (OMA) [10]. Then the appearance of damage affecting the stiffness and/or the energy dissipation properties of the system can be inferred from permanent variations in its modal signatures [11–14]. In this light, considerable research has been devoted in recent years to the development of automated OMA procedures to enable their use in continuous SHM schemes. In particular, a number of approaches have been proposed in the literature for the automated interpretation of stabilization diagrams obtained by time domain-based Stochastic Subspace Identification (SSI) methods. These include, among others, hierarchical clustering [15, 16], fuzzy clustering [17], Gaussian Mixture Models (GMM) [18], density-based clustering [19], and blind k-means [20]. Such techniques allow to conduct automated and quasi real-time identification of stable modal poles, which can be then used to trace time series of modal signatures of the monitored structure through a frequency tracking approach. Nonetheless, while highly effective for global damage assessment, vibration-based SHM faces major difficulties to detect local defects with limited impacts on the global stiffness. This circumstance suggests the need for incorporating diverse sensing solutions (e.g. dynamic, static, chemical) with superior local damage identification capabilities so to achieve both local and global damage identification. On this basis, a large variety of damage-sensitive features can be extracted from the monitoring signals, including probabilistic-based features, time- and frequency-domain features, modal properties and derived quantities (e.g. modal flexibility, mode shape curvatures [21]), as well as combined features obtained through data fusion [22, 23].

Numerous research works have also reported about the frequent existence of important effects of environmental and operational conditions (EOCs) (e.g. temperature, humidity, wind, traffic intensity) upon the structural response of civil engineering assets. A remarkable example was provided by Peeters and De Roeck [24] who found temperature-driven variations of up to 18% of the fundamental frequency of the well-known case study of the Z24-Bridge. Correlations between EOCs and the structural response may be disparate depending on manifold factors such as the construction materials, structural typology, connectivity, solar radiation, thermal capacitance, to mention a few. For instance, concrete and masonry structures often experience biphasic correlations for temperatures below and above zero. This is due to the formation of ice crystals in the material discontinuities under freezing air temperatures and the subsequent stiffening effect, which often manifests as sudden increases in the resonant frequencies of the structure (see e.g. [24–26]). The effects of EOCs translate into daily and seasonal trends in the monitoring data, which often mask the appearance of anomalies induced by damage and thus need to be filtered out to attain effective damage identification. This process, also referred to as data normalization, is usually conducted through statistical pattern recognition (SPR) and machine learning techniques [22]. In

general, data normalization consists in the construction of a statistical model mapping the selected monitoring features and trained over a certain baseline period in which the structure is assumed to remain in healthy conditions [27]. Data normalization approaches are generally classified as input-output (regression models) and output-only (latent-variable models) models whether data from EOCs are available or not, respectively [28]. Latent-variable statistical models have the key advantage of not requiring monitoring data from EOCs and thus imply lesser archive needs. Nonetheless, regression models are often required to properly interpret and quantify damage-induced anomalies [29].

In light of the previous discussion, along with the progressive cheapening of sensing technologies, it is clear that the use of integrated and dense SHM systems involving heterogeneous sensing solutions is becoming especially attractive. The management and processing of the monitoring data of such systems fall within the areas of Data Science and Big Data analysis and require expert engineering judgement given the extremely case-dependent character of SHM. This represents a formidable challenge from the algorithmic and software perspective, which hinders the technological transfer of SHM to routine engineering practice. Indeed, the number of commercial or open-source software codes for SHM management is rather scarce, and most existing solutions are in-house software codes developed by specialized research groups. Among the few existing ones, the most popular software programs such as MACEC [30], LMS Test.Lab [31], or ARTeMIS [32] are dedicated to dynamic testing of structures. While these codes provide extensive state-of-the-art OMA techniques, their ability to incorporate heterogeneous monitoring data and conduct data normalization/cleansing through SPR is considerably limited or infeasible. In light of these limitations, and with the aim of addressing the last prescriptions by the Italian guidelines of the monitoring of viaducts, this paper presents the development of a new software code for the autonomous management of integrated SHM systems within the SPR paradigm. The software, named P3P, is specifically designed for SHM of highway bridges, and represents an user-friendly and updated version of a previous software suite developed by some of the authors named MOVA/MOSS [33]. This new software platform was developed within a research project funded by Anas S.p.A, the largest Italian company dedicated to the management of the Italian roadway system. The software program includes automated OMA, frequency tracking, filtering of environmental effects, and damage detection through novelty analysis, which constitutes the most original contribution of the work, as well as two modules specific for processing and analysis of seismic events and structural reliability analysis of bridges. Ease of use, interpretability, and minimal intervention of expert judgement were cornerstone elements in the development of P3P. This motivated the development of a novel automated density-based tracking algorithm, as well as specific functionalities for span-wise identification of long multi-span bridges. In order to highlight the potential of P3P, this paper presents two application case studies: (i) a long multi-span bridge,

the Trigno V Bridge in the Abruzzo region of Italy; and (ii) the Z-24 Bridge benchmark [34, 35].

The remaining of the paper is organised as follows: Section 2 overviews the theoretical basis of continuous SHM involving dynamic- and non-dynamic monitoring data. Section 3 presents the software architecture of P3P. Section 4 presents the numerical results and discussion, finally, Section 5 concludes the paper.

2 Theoretical background

2.1 Automated OMA

2.1.1 Cov-SSI

Modal identification methods based on SSI are derived from the stochastic discrete-time state-space form of the equation of equilibrium of a linear-time-invariant dynamic system of n_2 degrees of freedom (DOFs) under white noise unknown excitation:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{w}(k), \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k), \end{aligned} \tag{1}$$

where $k \in \mathbb{N}$ is a generic time step (i.e. $t(k) = k\,\Delta t = k/f_s$ with $f_s = \Delta t^{-1}$ the sampling frequency), matrices $\mathbf{A} \in \mathbb{R}^{2n_2 \times 2n_2}$ and $\mathbf{C} \in \mathbb{R}^{l \times 2n_2}$ denote the state and output matrices of the system, respectively, and vectors $\mathbf{x} \in \mathbb{R}^{2n_2}$ and $\mathbf{y} \in \mathbb{R}^{l}$ stand for the state and observation vectors. Vectors $\mathbf{w}(k) \in \mathbb{R}^{2n_2}$ and $\mathbf{v}(k) \in \mathbb{R}^{l}$ represent white noise processes accounting for the unknown outputs as well as the effects of process noise (modelling inaccuracies) and measurement noise, respectively. Both noise vectors are assumed to be zero-mean realizations of stochastic processes with the following correlation matrices:

$$E\left(\begin{bmatrix}\mathbf{w}_p \\ \mathbf{v}_p\end{bmatrix}\begin{bmatrix}\mathbf{w}_q^{\mathrm{T}} & \mathbf{v}_q^{\mathrm{T}}\end{bmatrix}\right) = \begin{bmatrix}\mathbf{Q} & \mathbf{S} \\ \mathbf{S}^{\mathrm{T}} & \mathbf{R}\end{bmatrix}\delta_{pq}, \tag{2}$$

where sub-indexes p and q denote generic time instants, δ_{pq} is the Kronecker delta, and E is the expected value operator.

It can be demonstrated that the modal properties of the system can be obtained by the eigenvalue decomposition of matrix $\mathbf{A}$. Specifically, the structure's natural frequencies ω_i and damping ratios ζ_i can be extracted from the eigenvalues μ_i of matrix $\mathbf{A}$ as [36]:

$$\lambda_i = \frac{\ln(\mu_i)}{\Delta t} \Leftrightarrow \lambda_i = -\zeta_i\omega_i + \mathrm{i}\omega_i\sqrt{1-\zeta_i^2}, \tag{3}$$

with $\mathrm{i} = \sqrt{-1}$ being the imaginary unit. Moreover, the complex mode shapes $\boldsymbol{\varphi}_i$ of the system can be obtained from the eigenvectors $\boldsymbol{\phi}_i$ of $\mathbf{A}$ as:

$$\boldsymbol{\varphi}_i = \mathbf{C}\,\boldsymbol{\phi}_i. \tag{4}$$

The covariance-driven SSI (Cov-SSI) method is used in this work to identify the system's matrices because of its accuracy, noise-robustness, and computational efficiency. The Cov-SSI method exploits the covariances among the output measurements for positive time lags varying from Δt to $(2j_b-1)\Delta t$, represented by $\mathbf{R}_1$ to $\mathbf{R}_{2j_b-1}$, with $\mathbf{R}_i = E\left\{\mathbf{y}_{k+i}\mathbf{y}_k^{\mathrm{T}}\right\}$. Afterwards, the covariances are organized in a lj_b-by-lj_b block Toeplitz matrix as:

$$\mathbf{T}_{1|_{j_b}} = \begin{bmatrix} \mathbf{R}_{j_b} & \mathbf{R}_{j_b-1} & \dots & \mathbf{R}_1 \\ \mathbf{R}_{j_b+1} & \mathbf{R}_{j_b} & \dots & \mathbf{R}_2 \\ \dots & \dots & \dots & \dots \\ \mathbf{R}_{2j_b-1} & \mathbf{R}_{2j_b-2} & \dots & \mathbf{R}_{j_b} \end{bmatrix}. \tag{5}$$

To identify the system's matrices from $\mathbf{T}_{1|_{j_b}}$, Cov-SSI leverages an important property of stochastic state-space models relating the output covariances and the state vectors [37]:

$$\mathbf{R}_j = \mathbf{C}\mathbf{A}^{j-1}\mathbf{G}, \tag{6}$$

with $\mathbf{G} = E\left\{\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}}\right\}$ being the state covariance matrix. Applying the factorization in Eq. (6) to all the covariance matrices stored in the Toeplitz matrix in Eq. (5), $T_{1|_{j_b}}$ can be decomposed in the product of the following matrices:

$$\mathbf{T}_{1|_{j_b}} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \dots \\ \mathbf{C}\mathbf{A}^{j_b-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{j_b-1}\mathbf{G} & \dots & \mathbf{A}\mathbf{G} & \mathbf{G} \end{bmatrix} = \mathbf{O}\boldsymbol{\Gamma}. \tag{7}$$

The second equality in Eq. (7) defines the extended observability matrix $\mathbf{O}$ and the reversed extended stochastic controllability matrix $\boldsymbol{\Gamma}$. Matrix $\mathbf{O}$ is a column block matrix of j_b blocks with dimensions l-by-n (n is the dimension of the state-space model), while $\boldsymbol{\Gamma}$ is formed by j_b n-by-l matrices organized in a row. On the other hand, the singular value decomposition (SVD) of $\mathbf{T}_{1|_{j_b}}$ reads:

$$\mathbf{T}_{1|_{j_b}} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_2^{\mathrm{T}} \end{bmatrix} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^{\mathrm{T}}. \tag{8}$$

The number n of non-zero SVs (model order) concentrated in $\mathbf{S}_1$ gives the rank of the system (assuming $n < lj_b$). The comparison between Eqs. (7) and (8) allows to compute an estimate of the observability and the controllability matrices from the outputs of the SVD as:

$$\mathbf{O} = \mathbf{U}_1\mathbf{S}_1^{1/2}, \quad \boldsymbol{\Gamma} = \mathbf{S}_1^{1/2}\mathbf{V}_1^{\mathrm{T}}. \tag{9}$$

Once the observability and controllability matrices are obtained, the state-space model matrices $\mathbf{A}$ and $\mathbf{C}$ can be readily identified. On one hand, matrix $\mathbf{C}$ can be extracted from the first l lines of the observability matrix (see

Eq. (7)). On the other hand, a common approach to compute $\mathbf{A}$ is the Balanced Realization (BR) method, which exploits the shift structure of the observability matrix as [38]:

$$\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \dots \\ \mathbf{CA}^{j_b-2} \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \dots \\ \mathbf{CA}^{j_b-1} \end{bmatrix} \Leftrightarrow \mathbf{A} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \dots \\ \mathbf{CA}^{j_b-2} \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{C} \\ \mathbf{CA}^2 \\ \dots \\ \mathbf{CA}^{j_b-1} \end{bmatrix} = \mathbf{O}^{to\dagger} \mathbf{O}^{bo}, \tag{10}$$

where $\mathbf{O}^{to}$ and $\mathbf{O}^{bo}$ contain the first and the last $l(j_b-1)$ lines of $\mathbf{O}$, respectively, and symbol † stands for the Moore-Penrose pseudo-inverse.

2.1.2 Automated OMA

The Cov-SSI algorithm is controlled by two user-defined parameters: (i) the model order n, and (ii) the time-lag parameter j_b. The value of j_b defines number of block rows in the Toeplitz matrix and, therefore, determines the computational cost in the SVD in Eq. (8). Too small values of j_b may fail to identify low-frequency modes, while too large values increase the number of spurious modes and raise the computational burden of the identification. The value of j_b may be chosen by the following rule of thumb [39]:

$$j_b \geq \frac{f_s}{2\, f_o}, \tag{11}$$

with f_o the expected fundamental frequency of the system. Alternative approaches include the convergence analysis of the condition number of $\mathbf{T}_{1|_{j_b}}$ or certain modal estimates for increasing values of j_b (see e.g. [40]).

With regard to the model order, theoretically n can be defined by inspecting the number of non-zero singular values in the SVD of the Toeplitz matrix. Nevertheless, in practice, due to the presence of noise in the measurements and the circumstance that the system may not have a perfectly linear behaviour, singular values that should be zero usually present certain residual values. As a result, it is common to specify the model order in a conservative way with the aim of identifying weakly-excited modes (at least larger than twice the number of physical modes of the system). Nonetheless, this leads to the introduction of numerical modes (also called spurious or noise modes) induced by over-modelling of the noise in the output measurements. The most popular approach to separate physical from spurious modes consists in conducting the identification for increasing model orders within a certain interval $[n_{min}, n_{max}]$ with steps Δn. Let us denote $H_n = (n_{max} - n_{min})/\Delta n$ the number of model orders to be identified. A initial cleansing of the modal poles can be conducted at this stage by means of hard criteria (HC) eliminating complex conjugate modes and poles violating certain feasibility constraints, e.g. physically meaningful damping ratios or minimum modal phase collinearity (MPC) [41] values.

Then, the retained modal parameter estimates are represented together conforming a stabilization diagram (See Fig. 1 (a)). On this basis, stable poles can be identified as those abiding certain soft criteria (SC) in terms of tolerances of modal estimates (resonant frequency, damping ratios and mode shapes) obtained between consecutive model orders. This allows to distinguish physical from spurious modes by looking for vertical alignments of stable poles in the stabilisation diagram, that is to say, modes that consistently appear in most of the identifications.

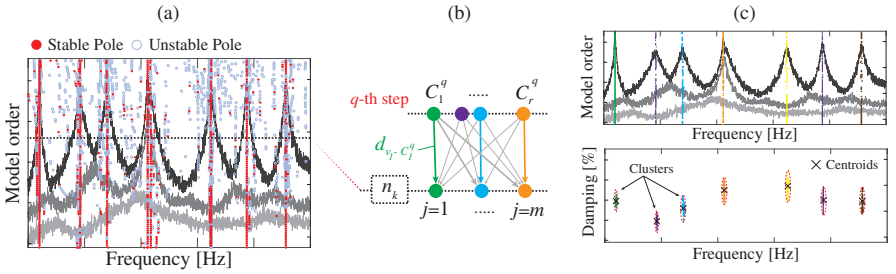

Fig. 1 Automated Cov-SSI: (a) stabilization diagrams, (b) modal distance, and (c) agglomerate hierarchical clustering.

While stabilisation diagrams provide a powerful tool to manually identify sets of physical modes, automated interpretation procedures are required for their use in continuous SHM. To do so, an agglomerate hierarchical clustering algorithm similar to the one proposed in reference [15] is implemented herein. The main goal is to group the previously identified stable modes into a set of homogeneous data clusters pertaining to the same structural mode. The algorithm starts by setting the first cluster seeds C_i^0, $i = 1, ..., r$ as the stable poles obtained for the maximum model order n_{max}, and initiating their cluster sizes as $cs_i = 1$. Afterwards, the algorithm sequentially explores the similarities between the existing clusters C_i^q $(q = 1, \ldots, H_n)$ and the stable poles identified with model orders from $n_{max} - \Delta n$ to n_{min}. Let us focus on the construction of a generic cluster C_i. At a generic q-th step, cluster C_i^q is compared against the m stable poles v_j, $j \in [1, ..., m]$, identified at the corresponding model order n_k, $k = 1 - (2 - n_{max}) \frac{q - H_n}{1 - H_n}$ (see Fig. 1) (b). In order to compare mode v_j and cluster C_i^q, a similarity metric is defined as:

$$d_{v_j - C_i^q} = \frac{\left| f_{v_j} - \overline{f}_{C_i^q} \right|}{\overline{f}_{C_i^q}} + 1 - MAC\left(\mathbf{\Theta}_{v_j}, \overline{\mathbf{\Theta}}_{C_i^q} \right), \tag{12}$$

where $\overline{f}_{C_i^q}$ and $\overline{\mathbf{\Theta}}_{C_i^q}$ denote the mean values of the frequencies and mode shapes of the poles conforming cluster C_i^q, respectively, while f_{v_j} and $\mathbf{\Theta}_{v_j}$ are the frequency and the mode shape of the j-th pole, v_j, respectively. Operator MAC in Eq. (12) stands for Modal Assurance Criterion [42]. Then, the similarities between C_i^q and the m poles are arranged in a similarity vector $\mathbf{d}_q$ as:

$$\mathbf{d}_q = \left[d_{v_1 - C_i^q}, d_{v_2 - C_i^q}, \ldots, d_{v_m - C_i^q} \right]^{\mathrm{T}}, \tag{13}$$

and the following rule is applied:

$$\begin{cases} C_i^{q+1} = C_i^q \cup \{v_h\}, \; cs_h = cs_h + 1 & \text{if } d_{q,h} = \min(\mathbf{d}_q) \leq d_{max} \\ C_i^{q+1} = C_i^q & \text{otherwise} \end{cases} \tag{14}$$

meaning that the h-th mode v_h (which is the closest one in terms of frequency and mode shape) is assigned to cluster C_i^q if the similarity value $d_{q,h}$ is smaller than an user-defined threshold d_{max}. If assigned, the centroid of the current cluster is updated in the next step C_i^{q+1} as the average value of frequencies, damping ratios, and mode shapes of the poles pertaining to this cluster. At this stage, poorly populated clusters can be considered as spurious modes and disregarded by checking if their cluster sizes are below a certain user-defined threshold cs_{min}, i.e. $cs_i \leq cs_{min}$. The algorithm terminates with a set of $C_i^{H_n}$ clusters characterized by their centroids and cluster sizes cs_i (see Fig. 1 (c)).

2.2 Feature Extraction

As anticipated in Section 1, it is convenient to count on heterogeneous damage-sensitive features to attain both local and global damage identification. In general, monitoring data in P3P are organised as static, environmental, and dynamic magnitudes. Given that static (e.g. displacements, crack amplitudes, tilts) and environmental data (e.g. humidity, temperature) usually exhibit low-varying trends, probabilistic-based metrics such as averages, max/min and root-mean-squared (rms) values suffice to obtain representative features. On the other hand, features from dynamic data can be extracted by probabilistic-based metrics and through OMA. The latter involves the application of the automated OMA procedure introduced in Section 2.1.2 to every ambient vibrations record acquired by the SHM system. This results in clusters of stable poles in time as shown in Fig. 2 (a) for the case of the Z-24 Bridge. A frequent situation in practice relates the circumstance that some clusters previously identified as physical poles do not actually represent any real mode of the monitored structure due to failures in the identification or the presence of non-stationary non-Gaussian excitations. Moreover, the identification may fail to identify some modes at times when these are weakly-excited (e.g. during night-time when excitations induced by traffic are usually low). To address these issues, it is necessary to implement a frequency tracking approach to extract time series of the modal signatures of the structure. A classical frequency tracking approach is to define a reference list or baseline modal properties extracted from a separate ambient vibration test (AVT) [27]. Once defined, the time series of modal features are traced by exploiting similarities between the reference modes and all the identified sets of modal characteristics. In this work, the similarity metric previously defined in Eq. (12) is adapted to compare

the sets of poles continuously extracted by the automated OMA procedure against the reference baseline. To avoid misclassification, the comparison is only conducted for sets of poles abiding with certain user-defined tolerances in terms of resonant frequencies, damping ratios and MAC values. In addition, P3P allows the user to select between static or dynamic frequency tracking. The latter procedure updates the reference resonant frequencies as new poles are added to the time series, which allows to accommodate fast trends in the modal signatures induced by EOCs (e.g. freezing conditions).

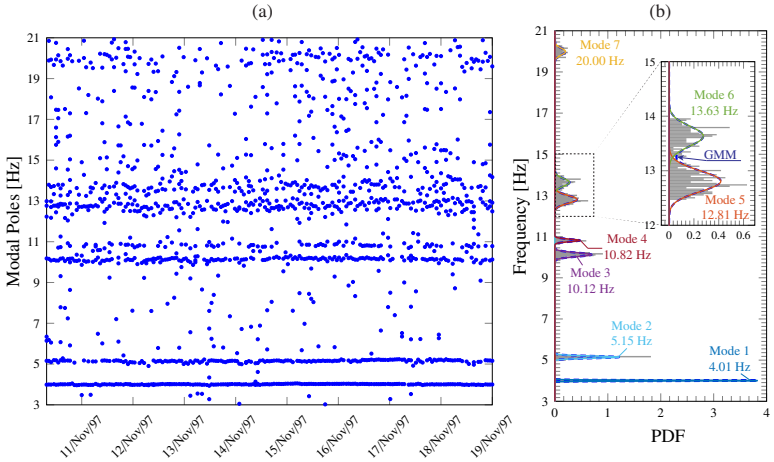

Fig. 2 Modal identification results of the first $N_t = 200$ acquisitions in the Z-24 Bridge (a), and automated identification of the reference modal baseline through parametric clustering (b).

With the aim of facilitating the frequency tracking process and avoiding the need for conducting an independent AVT, a novel density-based tracking procedure is proposed herein inspired by the work of Favarelli and Giorgetti [43]. The algorithm starts by automatically setting the reference baseline (starting phase) without no assumptions on the reference modal signatures or the number of modes to track. The idea is to analyse the probability distribution function (PDF) of the identified resonant frequencies in an initial period of N_t acquisitions collected in an observation vector $\mathbf{x}$ as shown in Fig. 2 (a) for the Z-24 Bridge ($N_t = 200$). In this case, the histogram of $\mathbf{x}$ in Fig. 2 (b) reveals seven areas of concentrated probability (modes of the PDF) corresponding to the first seven natural frequencies of the bridge in the broadband up to 21 Hz. To automatically identify these modes, a clustering procedure combining k-means and GMM is implemented. Firstly, k-means clustering [44] is used to partition the data into K clusters with maximum squared Euclidean distance between their centroids. To avoid the dependence on the initial seed of the k-means approach, the optimal clustering is selected as the best solution among 10 k-means repetitions with random centroid initialization. The best solution

is then used as the initial classification for the GMM. This model assumes that the probability density function $p(\mathbf{x})$ of the data set in $\mathbf{x}$ can be represented as a linear superposition of K Gaussian components as [44]:

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\mathbf{x}\middle|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right), \tag{15}$$

where each component of the mixture is defined as a Gaussian distribution $\mathcal{N}\left(\mathbf{x}\middle|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right)$ with mean and covariance matrix denoted by $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$, respectively. Parameters $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_K]^{\mathrm{T}}$ in Eq. (15) are called the mixing coefficients, and they range between 0 and 1 ($0 \leq \pi_k \leq 1$) and sum to one ($\sum_{k=1}^{K} \pi_k = 1$). The model parameters, $\boldsymbol{\mu}_k$,$\boldsymbol{\Sigma}_k$ and π_k, are fitted by minimizing the log-likelihood function:

$$\ln p(\mathbf{X}\middle|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^{t_p} \ln \left\{ \sum_{k=1}^{K} \pi_k \mathcal{N}\left(\mathbf{x}_n\middle|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right) \right\}. \tag{16}$$

The optimal solution for the parameters ($\boldsymbol{\mu}$,$\boldsymbol{\Sigma}$ and $\boldsymbol{\pi}$) minimizing Eq. (16) is estimated using the iterative Expectation-Maximization (EM) algorithm. In the expectation (E) step, the parameters (initial guess at the beginning of the E step) are held fixed and the posterior probability of assigning $\mathbf{x}_n$ to the k's cluster is given by the so-called responsibilities $\gamma(z_{nk})$ as:

$$\gamma(z_{nk}) = \frac{\pi_k \mathcal{N}\left(\mathbf{x}_n\middle|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right)}{\sum_{j=1}^{K} \pi_j \mathcal{N}\left(\mathbf{x}_n\middle|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j\right)}, \tag{17}$$

where z_{nk} is an element of a K-dimensional binary random variable $\mathbf{z}$ with a 1-of-K representation. Only one element in $\mathbf{z}$ is equal to 1 and all other elements are 0. Then, in the maximization (M) step, the parameters are re-estimated using the posterior probability calculated in the previous E step as follows:

$$\boldsymbol{\mu}_k^{new} = \frac{1}{N_k} \sum_{n=1}^{t_p} \gamma(z_{nk}) \mathbf{x}_n, \tag{18}$$

$$\boldsymbol{\Sigma}_k^{new} = \frac{1}{N_k} \sum_{n=1}^{t_p} \gamma(z_{nk}) \left(\mathbf{x}_n - \boldsymbol{\mu}_k^{new}\right) \left(\mathbf{x}_n - \boldsymbol{\mu}_k^{new}\right)^{\mathrm{T}}, \tag{19}$$

$$\pi_k^{new} = N_k/N, \quad N_k = \sum_{n=1}^{t_p} \gamma(z_{nk}). \tag{20}$$

The process is iterated using the updated data values until the convergence of the log likelihood is met. Since the number of modes in the reference baseline is unknown, the previous procedure is conducted sequentially for increasing

numbers of target clusters from $K = 1$ to K_{max}, with K_{max} being conservatively larger than the maximum number of expected modes. Once obtained, the optimal number of clusters and the corresponding baseline modes are obtained through the GMM with minimum Bayesian Information Criterion (BIC) [45]. In the case of the Z-24 Bridge, K_{max} was set to 12. It is noted in Fig. 2 (b) how the proposed procedure accurately identifies the seven natural modes of the bridge. On this basis, the reference baseline modes are defined according to the average values of the modal poles pertaining to the identified mixtures.

Once the reference baseline is identified, it follows the online phase where the reference modes are tracked throughout the whole monitoring period. To do so, a Gaussian kernel is defined for all the reference resonant frequencies in the baseline as:

$$\mathcal{N}\left(\theta;\ f^{ref}, \sigma_r\right) = e^{-\frac{1}{2\sigma_r^2}\left(\theta - f^{ref}\right)^2}, \tag{21}$$

with f^{ref} and σ_r being the r-th reference resonant frequency and the associated standard deviation σ_r. Parameter σ_r controls the kernel width, and, as a rule of thumb, it can be selected as the standard deviations obtained by the mixtures identified in the previous GMM. Then, to compare the identified poles at a generic time instant q, the following distance metric is also defined:

$$dm = (1-\alpha)\frac{\left|f^p - f^{ref}\right|}{f^{ref}} + \alpha\left[1 - MAC\left(\boldsymbol{\Theta}^p, \boldsymbol{\Theta}^{ref}\right)\right], \tag{22}$$

with superscripts ref and p relating the corresponding magnitudes to the reference mode and a generic pole, and coefficient α being a weighting factor between the contribution of the resonant frequencies and mode shapes. At a generic step N_a in the tracking process, consider a number n_{ref} reference modes and n_p poles abiding certain under-defined tolerances with respect to the baseline modes. Then, the distances between the retained poles with respect to the baseline modes are computed and stored in a distance matrix $\mathbf{D} \in \mathbb{R}^{n_{ref} \times n_p}$. On this basis, the pole with minimum distance to every reference mode is selected and stored in observation matrix $\mathbf{F}$ containing the time series of tracked modes:

$$\mathbf{F} = \begin{bmatrix} f_1^{(1)} & \dots & f_1^{(N_a)} \\ \dots & \dots & \dots \\ f_{n_{ref}}^{(1)} & \dots & f_{n_{ref}}^{(N_a)} \end{bmatrix}. \tag{23}$$

The previous tracking procedure can be also made dynamic by updating the reference resonant frequencies every time a new pole is collected in $\mathbf{F}$. To do so, following the Bayes' rule and considering conjugate prior and posterior distributions, the mean values of the Gaussian kernels in Eq. (21) are updated as f_{up}^{ref}:

$$f_{up}^{ref} = \sigma_{up}\left(\frac{f^p}{\sigma_f^2} + \frac{f^{ref}}{\sigma_r^2}\right), \tag{24}$$

where σ_f denotes the standard deviation of a zero-mean Gaussian likelihood function. Large values of σ_f decrease smoothness but allow capturing sudden changes of modal frequencies. The standard deviation of the Gaussian kernels in Eq. (21) are not updated throughout the tracking process to keep the kernel width constant.

2.3 Statistical Pattern Recognition

From the previous section, the presence of damage can be inferred from the analysis of an observation matrix $\mathbf{Y}$ containing the selected damage-sensitive features (static and/or dynamic features). Nonetheless, given the masking effects of EOCs, these quantities cannot be directly used for damage identification purposes. Instead, the variability in $\mathbf{Y}$ due to benign EOCs needs to be filtered out through data normalization as sketched in Fig. 3 (a). To do so, a certain SPR model is trained to reproduce the features in $\mathbf{Y}$ obtained in a baseline in-control period, usually referred to as the *training period*, when the structure is assumed to remain in healthy condition. This must be long enough to cover the full range of environmental conditions, including both daily and seasonal fluctuations (typically one year). On this basis, the residuals $\mathbf{E} \in \mathbb{R}^{n \times N}$ between the features in $\mathbf{Y}$ and the predictions made by the statistical model $\widehat{\mathbf{Y}}$ result more convenient for damage identification:

$$\mathbf{E} = \mathbf{Y} - \widehat{\mathbf{Y}}, \tag{25}$$

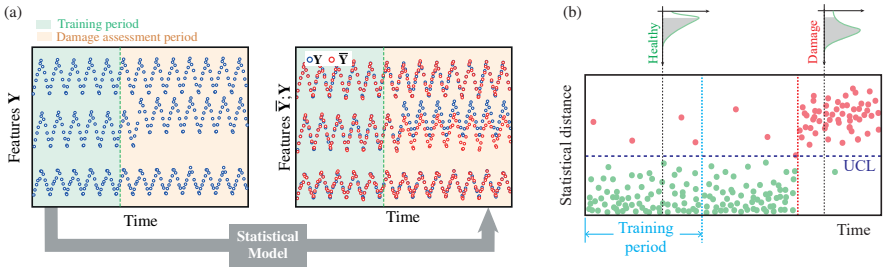

Fig. 3 Flowchart of data normalization through statistical pattern recognition (a), and schematic representation of damage detection using control charts.

Note that matrix $\widehat{\mathbf{Y}}$ condenses the part of the variance of the features in $\mathbf{Y}$ driven by EOC. Hence, when the system remains healthy, matrix $\mathbf{E}$ only contains the residual variance stemming from modelling errors. Conversely, if a certain damage develops, this only affects the data contained in $\mathbf{Y}$ while matrix $\widehat{\mathbf{Y}}$ remains unaltered. Therefore, matrix $\mathbf{E}$ concentrates the damage-induced variance and is apt for being used for damage identification.

Although a wide variety of SPR models have been proposed in the literature, Multiple Linear Regression (MLR) and Principal Components Analysis (PCA) are included in P3P as the most commonly used input-output and

output only ones, respectively. Because of space constraints, their theoretical description is omitted in this paper, and interested readers may refer to references [46, 47] for further details. Concatenations of these two models, as well as cluster-wise non-linear extensions of MLR and PCA are also available in P3P following an equivalent approach to the one presented in Section 2.2 (see reference [33] for further details). Finally, a simplified clustering approach for bi-linear SPR is also implemented in P3P by creating local models for damage-sensitive features with temperatures below and above zero.

2.4 Novelty Analysis

Once the residual error matrix $\mathbf{E}$ is computed, the appearance of damage can be detected through novelty analysis of control charts as sketched in Fig. 3 (b). These furnish in time a certain statistical distance accounting for disturbances in the distribution of the residuals contained in $\mathbf{E}$. By defining an in-control region, the appearance of out-of-control processes, possibly associated to damage, is detected in the shape of data points violating the in-control region. A wide variety of control charts has been reported in the literature, although the Hotelling's T^2 control chart is the most extensively used in the realm of SHM. The statistical distance assessed by the Hotelling's T^2 control chart [48] is defined as:

$$T_i^2 = r \left(\overline{\mathbf{E}} - \overline{\overline{\mathbf{E}}}\right)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} \left(\overline{\mathbf{E}} - \overline{\overline{\mathbf{E}}}\right), \; i = 1, 2, \ldots, N/r, \tag{26}$$

where r is an integer parameter referred to as subgroup size, $\overline{\mathbf{E}}$ is the mean of the residuals in the subgroup of the last r observations, while $\overline{\overline{\mathbf{E}}}$ and $\boldsymbol{\Sigma}_0$ are the mean values and the covariance matrix of the residuals statistically estimated in the training period. The statistical distance in Eq. (26) is positive by definition, so an in-control region can be defined by an interval $[0, UCL]$, with the upper control limit UCL relating the statistical distance associated with a certain confidence level for the distribution of data within the training period. On this basis, control charts constitute an easily automatable and intuitive tool to conduct damage detection. The user can identify the appearance of an anomalous condition through persistent accumulations of data points violating the in-control region. Furthermore, approaches for automated identification of anomalies in control charts can be also found in the literature (see e.g. [33]).

2.5 Seismic record analysis and structural reliability assessment

Additional modules that can provide the infrastructure owner with useful information for implementing decisions in both emergency and ordinary conditions are also included in P3P. Specifically, a seismic record analyser has been developed to collect ground-shaking time-histories defined as those acceleration time series recorded by the sensors installed on the monitored bridge exceeding a certain user-defined acceleration threshold. Such a tool analyses them

and produces a quick seismic report describing the main features of the strong motion and conducts a comparison with the seismic hazard curve prescribed by the Italian Building Code for the site of interest. In this way, the infrastructure owner can readily quantify the severity of the event and the need of a carrying out a subsequent post-event visual inspection. The strong motions stored in P3P are processed in both time and frequency domains to get relevant intensity measure (IM) parameters, able to briefly describe the impact of the earthquake at the bridge site. These include the peak ground acceleration (PGA), velocity (PGV), displacement (PDG), the root mean square acceleration (RMSA) and the Arias Intensity (AI) values. The seismic record analyser also provides the response spectra of the recorded ground motions in both spectral acceleration and displacement values for both the horizontal and vertical components of the ground shaking. In addition, it is also possible to quantify the return period of the event based on the comparison of the hazard curve for the site of interest in terms of PGA and the specific value of such IM that characterizes the record. Finally, if there are sensors located at the base and top of the piers, P3P also processes the time histories to estimate the real drift demand time history. The maximum drift demand is then compared with a certain user-defined drift threshold value to have a preliminary estimate of the potential damage level that an inspector can expect to observe during post-event visual surveys.

An additional tool is also included in P3P to assess the structural reliability of bridge decks at the Ultimate Limit State considering as possible failure mechanism both ductile and fragile collapses (bending and shear actions, respectively). Specifically, the current version of the code only allows such type of analysis for pre-stressed reinforced concrete simply-supported girders, although the authors intend to develop an extension also for continuous decks. In detail, the user has to input 4 different sets of data that will be post-processed to quantify the structural reliability index β according to the Eurocode 0 format, i.e. deck geometrical characteristics, material properties, beam cross-section and reinforcement positions. Random variables are modelled according to the prescriptions provided by the Joint Committee for Structural Safety (JCSS), to get via the use of the First-Order Reliability Method (FORM) an estimate of the structural reliability index β.

3 Software architecture

The typical work-flow of an SHM system permanently installed in a bridge and managed by P3P is sketched in Fig. 4. In general, the monitoring system consists of an integrated sensor network deployed on the structure comprising dynamic, environmental and static sensors. An in-place data acquisition system (DAQ) permanently collects the monitoring data from the sensor network, and monitoring records of certain time duration are transferred and stored in a server or in the cloud. Here the data is continuously accessed by P3P which automatically performs all the steps involved in SHM as a SPR problem

as overviewed above. To do so, P3P distinguishes between acceleration and non-acceleration data, and the user can define independent tags and units for acceleration, EOC, and static sensors. Acceleration signals are introduced in a sequential process involving: (a) signal-processing of acceleration signals, (b) system identification through automated OMA, and (c) frequency tracking. At this stage, the user can identify subgroups of accelerometers corresponding to separate spans in multi-span bridges, which allows to conduct span-wise OMA and frequency tracking. In parallel, probabilistic-based features (mean and rms) are extracted from both acceleration and non-acceleration data. This process constitutes the feature extraction phase in P3P (i). Finally, once the user selects the set of desired damage-sensitive features, P3P conducts (ii) data normalization and (iii) anomaly detection. The outcome of this process is the generation and automated updating of multiple control charts assessing the desired sets of damage-sensitive features.

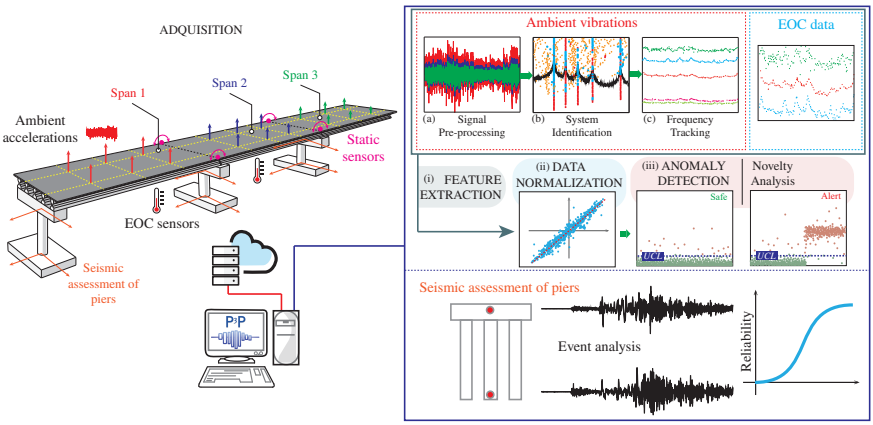

Fig. 4 Schematic of a permanent SHM system installed in a bridge structure using P3P.

While originally developed using MATLAB environment, a stand-alone version of P3P has been developed in C++ with a compact graphical user interface (GUI) as shown Fig. 5. From this interface, the user can access 10 main modules, namely:

1. Project configuration: Definition of local or FTP directories where acceleration data are stored, together with file configuration properties (extension, heading lines, decimal separator and text delimiter), log folder, naming protocols, and outputs settings.
2. Modal Geometry: The user can introduce formatted geometry text files for the representation of natural mode shapes. A large variety of geometrical entities can be defined, including nodes, lines, colour planes, acceleration channels (tag, node and measurement direction through three direction cosines), and kinetic equations among nodes. The code also

allows to identify sets of channels belonging to different spans in multi-span bridges. Finally, the user can also set certain channels dedicated to seismic analysis (typically at the base and top of piers).

3. Signal pre-processing: This module allows to visualize the acceleration time signals in the time- and frequency-domain (pre- and post-filtering). The software interface provides a library of filters to minimize the effects of noise and the presence of abnormal events, namely signal downsampling, linear detrend, Butterworth band-pass filtering, and Hanning window filtering of signal spikes.
4. System Identification - OMA: This module includes the automated procedure previously introduced in Section 2.1. Additionally, the user can also apply manual or automated Frequency Domain Decomposition (FDD) for fast assessment and verification. The identification results are presented in the shape of tabulated data, stabilization diagrams, exportable reports, and the histogram representation of the MAC matrix. Moreover, this interface counts on a specific section for mode shape representation, including the animation of complex modes, complexity plots, and display of quality factors such as the Mode Phase Collinearity (MPC) and Mean Phase Deviation (MPD) [16].
5. Static/Environmental data: The user can introduce as many environmental and static sensors as desired by providing a reference tag, measurement unit, naming protocol, and the directory that contains the record files.
6. Modal tracking: This module includes the frequency tracking approaches previously presented in Section 2.2. Only if the manual algorithm is selected, the user must introduce a reference set of frequencies. In either algorithm, a set of tolerances on allowable maximum relative frequency variations and minimum MAC values must be introduced. The outcome of this module is the extraction of the time series of resonant frequencies, damping ratios, and mode shapes of the structure.
7. Statistical Pattern Recognition: This module implements the SPR models previously overviewed in Section 2.3. The user can define different statistical models accounting for distinct damage-sensitive features. Specifically, this interface includes the following features: filling of missing data through autoregressive modelling, estimators/predictors definition, and clustering analysis to create cluster-wise non-linear SPR models. Every model, identified by an user-defined label name, is used for novelty analysis in the subsequent damage identification module.
8. Continuous SHM: This last step performs online novelty analysis through control charts based on the previously defined damage-sensitive features. It allows to manage permanent SHM systems in real-time and in a completely autonomous way. The interface includes real-time graphs of the time series of the estimators and acceleration data, as well as the corresponding control charts. Every time a new data file is found, the record is processed following the settings of the previous modules, the interface

graphs are updated, and the software performs damage detection based on all the previously defined SPR models.

9. Seismic analysis: Every time a seismic event is detected as acceleration signals overpassing a certain user-defined acceleration threshold, a computer file identifying the seismic event is generated. Then, the user can conduct the analysis of the registered seismic event, including its comparison with the design spectra prescribed by the Italian seismic code (NTC 2008), and the computation of meaningful engineering parameters such as Arias intensity, cumulative absolute velocity (CAV), response spectra, as well as the displacement drifts between the base and top of the piers.
10. Reliability: Finally, this module allows to compute the reliability index of the bridge under study by introducing its geometrical properties (dimensions, number of girders, typology and loading conditions), material parameters (concrete and steel classes), and the cross-section of the girders (dimensions and position of reinforcements).

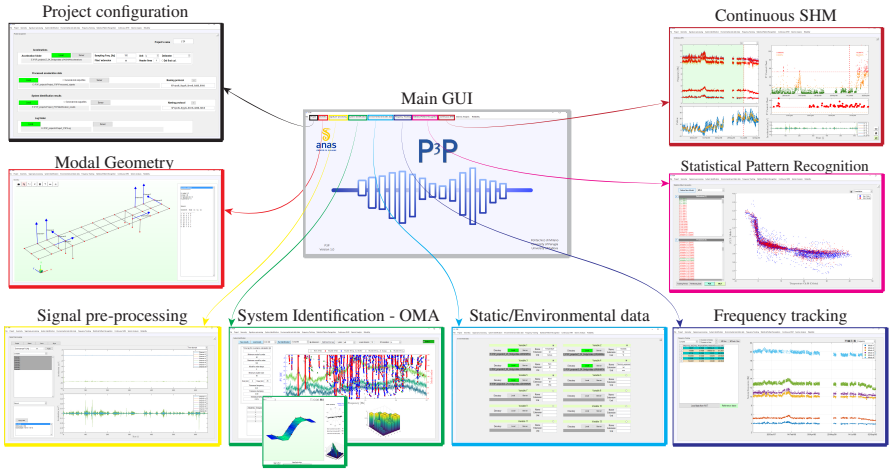

Fig. 5 Software architecture of P3P.

4 Case Studies and Discussion

This section presents the application results of P3P to two case studies: the Trigno V Bridge in the Abruzzo region of Italy (Fig. 6 (a,c)) and the Z-24 Bridge benchmark (Fig. 6 (b,d)). The first case study illustrates the challenges involved in the OMA of long multi-span simply supported bridges. On the other hand, the Z-24 Bridge is analysed to demonstrate the potential of P3P for automated damage identification.

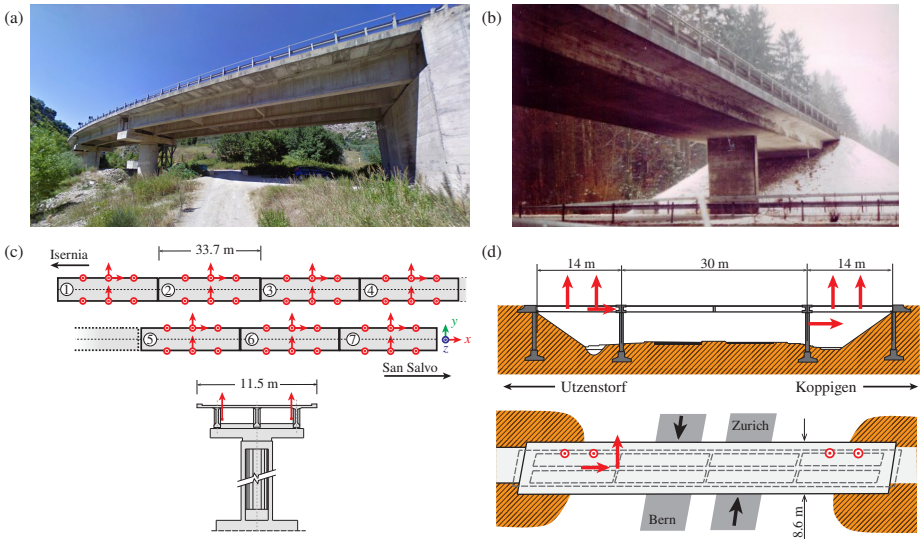

Fig. 6 Case studies: Trigno Bridge (a,c), Z-24 Bridge (b,d) [49].

4.1 Trigno Bridge

The Trigno V Bridge is a seven span simply supported concrete-girder highway bridge (7×33.7 m long) located in the Italian region of Abruzos between the municipalities of Isernia and San Salvo (Fig. 6 (a)). The isostatic spans are nominally equal with length of 33.7 m and width of 11.5 m (Fig. 6 (c)). Each span is made of three longitudinal I-shaped pre-stressed girders separated by 3.8 m between axes and connected to a 0.25 m thick reinforced concrete deck slab and transverse tie diaphragms with a spacing of 11.2 m. Within the development framework of P3P, a dense dynamic monitoring system was installed in the bridge on October 13$^{\text{th}}$ 2021. The system comprised 50 MEMS-based accelerometers covering 81 measurement channels (see Fig. 6 (c)). Four asynchronous 30 min long acquisitions were acquired between 11:00 a.m. and 12:30 p.m. at a sampling rate of 200 Hz and under normal operating conditions, with wind and traffic as the main sources of excitation. In this work, only the vertical accelerations are processed. This amounts to 42 channels (6 channels per span) as illustrated in the modal geometry defined in P3P and shown in Fig. 7.

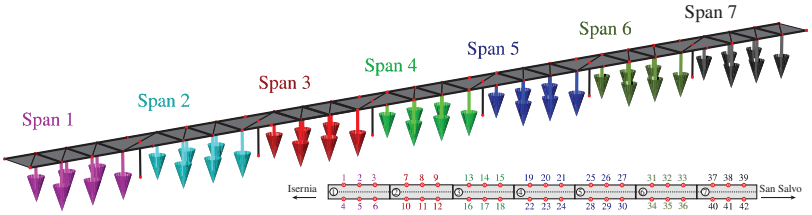

Fig. 7 Modal geometry of the Trigno Bridge.

Ambient vibrations have been processed including a linear detrend, second-order Butterworth band-pass filter with cut-off frequencies of 1 Hz and 15 Hz, and down-sampling to 30 Hz. An example of the processed acceleration time series and the corresponding power spectral density (PSD) curves for the second span of the Trigno Bridge are furnished in Figs. 8 (a) and (b). Then, the automated OMA procedure previously introduced in Section 2.1 is applied to identify the modal signatures of all the spans. It is important to note that the modal identification of the whole bridge considering simultaneously all the measurement channels is extremely challenging, since it is well-known that repetitive or quasi-periodic structures exhibit clustered modes with closely spaced natural frequencies and modes with similar wavelengths [50]. To avoid such difficulties, the modal identification is conducted at the span level with the aim of conducting span-wise damage detection. In this light, Fig. 9 (a) shows the stabilization diagram obtained from the second span of the Trigno Bridge using the processed time signals previously reported in Fig. 8. The modal identification in all the spans has been conducted with time lag parameters j_b ranging from 24 (0.8 s) to 44 (1.45 s), and model orders between 2 and 120 with steps of 2. For the elimination of spurious poles, HC include maximum damping ratios of $\zeta_{max} = 10\%$ and minimum MPC values of 80%. Besides, a minimum cluster size of 20 and a maximum distance $d_{max} = 0.07$ have been set for the clustering of stable poles. The identification results in Fig. 9 (a) reveal the existence of five alignments of stable poles. Is is observed in Fig. 9 (b) that the automated procedure previously presented in Section 2.1.2 succeeds at identifying five clusters of stable poles. Note that the two first clusters (with frequencies of 3.88 and 4.16 Hz) correspond to two clear peaks clearly observable in both the singular values (SVs) of the spectral matrix (inserted in the background of Fig. 9 (b) for clarity purposes) and the PSD curves from Fig. 8 (b). Indeed, these clusters are highly populated with stable poles identified at most of the considered model orders. Conversely, the remaining poles in the frequency broadband between 8 and 14 Hz are noticeably less populated, with a considerable number of misclassification at several model orders. This fact agrees with the inspection of the PSD curves or the SVs, which, despite exhibiting energy concentrations at frequencies neighbouring the identified clusters, do not present clear resonant peaks. Since the traffic was considerably intense during the acquisitions, this is conceivably ascribed to

the interaction with the passing vehicles, which may be masking the structural vibrations in this frequency range.

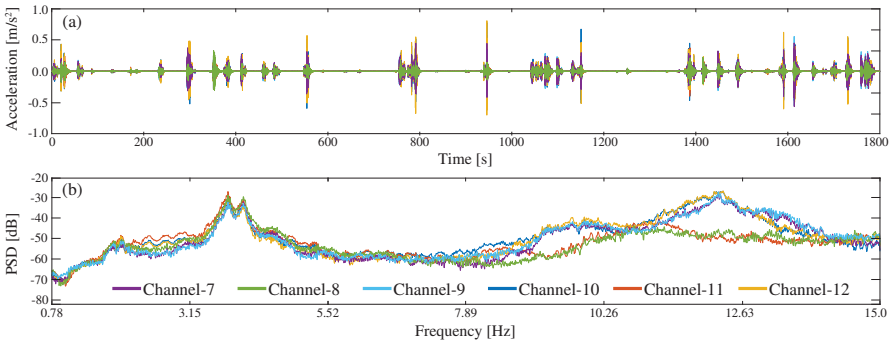

Fig. 8 Ambient vibrations recorded in the second span of the Trigno Bridge (a), and corresponding PSD curves (4096 data points, frequency resolution of 7.32E-3 Hz, 11:00 p.m., October 13th) (b).

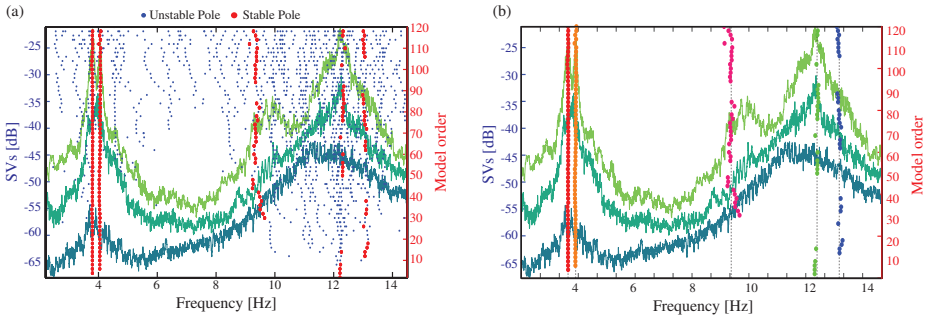

Fig. 9 Modal Identification of the second span of the Trigno Bridge (11:00 p.m., October 13th 2021): (a) Stabilization diagram, and (b) clusters of identified physical modes.

Following the previous procedure and after interpretation of the natural mode shapes, four vibration modes have been consistently found in all the spans of the bridge as illustrated in Fig. 10. The identification results averaged over all the acquisitions are reported in Table 1, and the resonant frequencies of all the spans of the Trigno Bridge are also illustrated in Fig. 11 for easy interpretation of the results. These include two closely spaced modes, namely a first-order flexural mode (3.8-3.9 Hz) and a first-order torsional mode (4.0-4.2 Hz), a second-order torsional mode (9.3-12.3 Hz), and a second-order flexural mode (12.0-13.2 Hz). Note in Table 1 that, indeed, all the spans exhibit very close resonant frequencies (especially the first two modes) and mode shapes with very similar wavelengths (Fig. 10), which justifies the use of

span-wise identification. Finally, note that low variability was found between the resonant frequencies identified over all the acquisitions (average variability of 0.4%). Conversely, the identification of structural damping is considerably more uncertain, with large variability ranges as common in OMA of moderately rigid civil engineering structures.

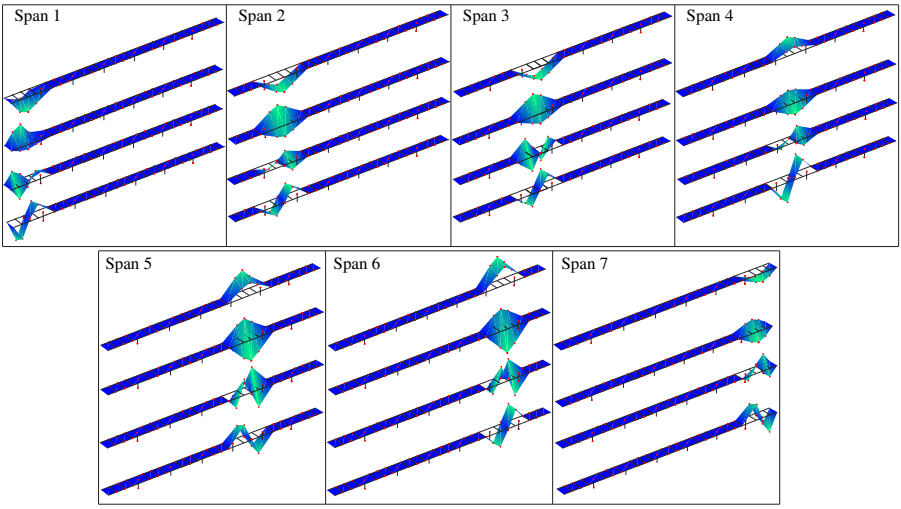

Fig. 10 Natural mode shapes of the Trigno Bridge (11:00 p.m., October 13$^{\text{th}}$ 2021).

Table 1 Modal identification results of the Trigno V Bridge (October 13$^{\text{th}}$).

	Mode No.	f_i [Hz]	ζ_i [%]		Mode No.	f_i [Hz]	ζ_i [%]
Span 1	1	3.87 ± 3.4E-01%	1.62 ± 7.8%	Span 5	1	3.85 ± 1.9E-01%	1.72 ± 23.9%
	2	4.16 ± 7.3E-02%	2.57 ± 10.1%		2	4.13 ± 8.4E-01%	2.57 ± 22.2%
	3	10.06 ± 8.2E-01%	2.62 ± 57.4%		3	9.74 ± 3.2E-01%	2.09 ± 13.2%
	4	13.05 ± 9.5E-02%	1.34 ± 25.9%		4	12.76 ± 5.8E-01%	1.85 ± 16.5%
Span 2	1	3.79 ± 3.6E-02%	1.36 ± 6.6%	Span 6	1	3.82 ± 1.8E-01%	1.47 ± 9.9%
	2	4.06 ± 4.6E-01%	2.39 ± 49.0%		2	3.98 ± 4.1E-01%	3.49 ± 7.7%
	3	9.31 ± 8.2E-01%	3.25 ± 10.0%		3	9.84 ± 1.9E+00%	4.84 ± 15.4%
	4	12.20 ± 2.6E-01%	1.39 ± 7.9%		4	12.49 ± 7.8E-01%	1.53 ± 25.2%
Span 3	1	3.85 ± 2.5E-01%	1.24 ± 25.8%	Span 7	1	3.89 ± 3.7E-01%	1.88 ± 42.8%
	2	4.11 ± 2.9E-01%	1.60 ± 45.5%		2	4.13 ± 6.0E-01%	1.95 ± 17.9%
	3	9.32 ± 2.7E-01%	2.14 ± 27.8%		3	12.26 ± 9.7E-02%	1.91 ± 45.2%
	4	12.04 ± 2.5E-01%	1.61 ± 30.3%		4	13.15 ± 2.5E-01%	2.18 ± 12.1%
Span 4	1	3.84 ± 2.3E-01%	1.61 ± 16.8%				
	2	4.09 ± 8.8E-01%	4.25 ± 17.8%				
	3	11.19 ± 2.1E-01%	1.61 ± 36.7%				
	4	12.09 ± 3.6E-01%	1.33 ± 25.0%				

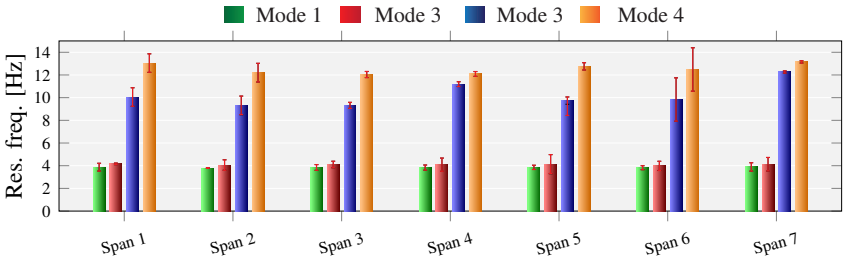

Fig. 11 Resonant frequencies of the first four resonant frequencies of the spans of the Trigno Bridge (error bars indicate the uncertainty in the identifications [%]).

4.2 The Z-24 Bridge

This last section demonstrates the ability of P3P to detect damage through the well-known benchmark of the Z-24 Bridge. The bridge, before demolition in 1998, was extensively instrumented and tested with the purpose of providing a benchmark case study for vibration-based SHM as reported for the first time by Peeters and De Roeck [34]. 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 144 m, as depicted in Fig. 6 (b). A continuous monitoring system was installed in the bridge from November 11^{th} 1997 until September 10^{th} 1998. Artificially controlled damage was introduced in the bridge in the last month of operation. In particular, 14 different damage conditions were induced in the bridge, including different settlements, tilting of one of the piers, failure of anchor heads and rupture of tendons [35, 49]. The monitoring database is available to the scientific community by the structural mechanics section of KU Leuven. It includes 5556 computer files of ambient vibration records with 65536 samples ($\approx$ 11 min), acquired at a sampling rate of 100 Hz with a periodicity of 1 hour from November 10^{th} 1997 until September 4^{th} 1998. Accelerations were recorded by 8 sensors deployed on the bridge according to the layout in Fig. 6 (d). In addition, the database also includes records of environmental data, comprising temperature (46 channels), wind and humidity sensors at several locations.

The monitoring data were introduced in P3P to conduct damage identification according to the SPR framework previously overviewed in Section 2. Acceleration records were subject to a filtering process including linear detrend, second-order Butterworth filtering with cut-off frequencies of 1 and 30 Hz, and signal down-sampling to 60 Hz. Then, the modal signatures of the bridge were identified following the previously introduced automated Cov-SSI procedure with identification parameters $j_b = 120$ (2 s), $d_{max} = 0.03$, $\zeta_{max} = 10\%$, and model orders ranging from 2 to 100 with steps of 2. The resulting set of modal clusters in time is furnished in Fig. 12. The time series of physically representative modes of the bridge were extracted through the automated frequency tracking approach previously introduced in Section 2.2. To do so, the reference baseline obtained by a GMM and previously reported

in Fig. 2 has been used to trace the time series of the seven modal signatures of the bridge throughout the whole monitoring period. In the analyses, the model parameters of the Guassian kernels used for the tracking were extracted from the GMM, and the standard deviation of the likelihood function used for the updating of the reference resonant frequencies (mean values of the Guassian kernels) was set to $\sigma_f = 0.55$. Also, the tolerances for relative variations of resonant frequencies were set between 3% to 5%. It is noted in Fig. 12 that the proposed algorithm successfully traces the time series of resonant frequencies all throughout the monitoring period (average success rate of 75.77%).

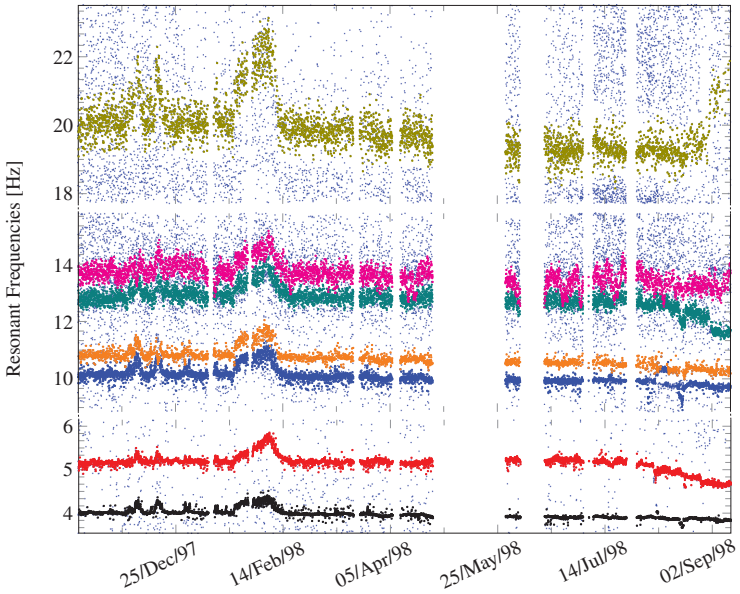

Fig. 12 Time series of resonant frequencies of the Z-24 Bridge identified by automated density-based frequency tracking.

The correlation between the fundamental frequency and the environmental air temperature is depicted in Fig. 13 (a). It is clear the existence of a non-linear biphasic correlation for temperatures below and above zero. This fact explains the presence of sudden shifts in the natural frequencies previously reported in Fig. 12 (see e.g. January 1998). As previously hypothesized by many other authors, this phenomenon is explained by the stiffening of the asphalt layer during freezing conditions. The results in Figs. 12 and 13 are clear examples of the need for identifying and eliminating the variance induced by EOCs before conducting damage identification. To do so, a cluster-wise non-linear MLR model has been trained considering the first four resonant frequencies of the bridge as estimators, and the recordings by 5 temperature sensors as predictors (air temperature, temperature in Web south, pavements,

and soffit, see reference [51] for further details). The monitoring data have been trivially separated in two clusters containing temperatures below and above zero, and a MLR model has been constructed through least squares fitting over a training period defined from November 10th 1997 until July 14th 1998 (4404 samples). Reasonably good fittings were found between the predictions of the MLR model and the experimental data as shown in Fig. 13 (a) (coefficient of determination R^2=0.8 and root-mean-squared error (RMSE) of 4.3E-2). Once trained, the MLR model is used to reconstruct the time series of resonant frequencies throughout the damage assessment period as shown in Fig. 13.

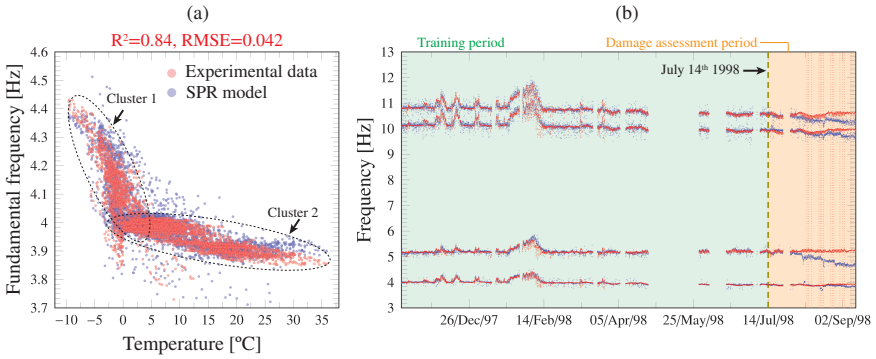

Fig. 13 Correlation between the fundamental frequency of the Z-24 Bridge and air temperature (over the training period from November 10th 1997 until July 14th 1998) (a). Predictions of the first four resonant frequencies of the Z-24 Bridge by MLR (b). Red dotted lines in (b) denote the dates when the artificial damage was induced in the bridge.

Finally, Fig. 14 shows the Hotelling's T^2 control chart of the residuals from Fig. 13. In this case, given the limited number of data-points in the damage assessment period, a subgroup size $r = 1$ has been defined in Eq. (26) for better illustration. The UCL has been set with a confidence level of 90% and estimated through the empirical analysis of the cumulative distribution of the data points in the training period. The appearance of the induced damage is clearly noted in Fig. 14 in the shape of a persistent accumulation of data points above the UCL. In particular, after the first damage scenario (August 10th 1998, (D1)-Lowering of pier of 20 mm), the number of data points violating the UCL raises to 67.5%. It is also noticeable in Fig. 14 (b) that persistent shifts are found in the data-points after some major damage events (indicated with inserts in the figure), which evidence the appearance of accumulated damage. To highlight this aspect, the moving median of the data points in the control chart (time window of 72 data points, 3 days) is included in Fig. 14. It can be observed that the moving median exhibits three major shifts (after D1, D3 and D5) followed by intermediate progressive raises as a result of intermediate pathologies. These results demonstrate the ability of P3P to handle large databases of heterogeneous monitoring data and effectively detect damage and, to some extent, track its severity.

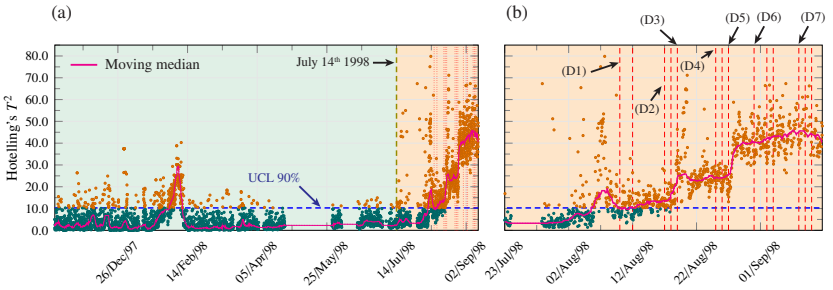

Fig. 14 Control chart for damage identification of the Z-24 Bridge (a) and zoom view of the damage assessment period (b). Red dashed lines indicate the dates when the artificial damage was induced in the bridge ((D1)-lowering of pier 20 mm, (D2)-lowering of pier 80 mm, (D3) - tilt of foundation, (D4) - spalling of concrete, (D5) - landslide at abutment, (D6) - failure of concrete hinge, (D7) - rupture of tendons).

5 Conclusions

This work has presented the development of a new software suite called P3P for the autonomous management of integrated SHM systems within the SPR paradigm. With the main purpose of addressing the last prescriptions by the Italian guidelines of the monitoring of viaducts, the software code has been developed with state-of-the-art SHM techniques within the SPR paradigm. Specifically, the code includes specific modules for automated OMA, frequency tracking, filtering of environmental effects, and damage detection through novelty analysis, as well as two specific modules for processing and analysis of seismic events and structural reliability analysis of bridges. Furthermore, a novel automated density-based tracking algorithm has been developed and validated. The potential of P3P has been appraised through two real case studies: (i) the Trigno V Bridge, and (ii) the Z-24 Bridge benchmark. The first case study has been selected to highlight the ability of the developed code to handle dense dynamic-testing campaigns and perform OMA of long multi-span bridges. In this regard, the developed software allows the user to identify subsets of channels pertaining to each span of the bridge and conduct span-wise OMA. Numerical results and discussion have been reported to illustrate the difficulties involved in the identification of quasi-periodic structures exhibiting dense clusters of modal poles with close frequencies and mode shapes with similar wavelengths. In this context, span-wise OMA becomes convenient for an easy identification and the possibility of conducting span-wise damage identification. Finally, the Z-24 Bridge benchmark has been analysed to demonstrate the potential of the proposed automated density-based frequency tracking, as well as to investigate the ability of P3P to identify progressive damage. To do so, a bi-linear MLR model has been constructed to filter out the effects of EOCs upon the resonant frequencies of the bridge. Then, the cleansed time series of resonant frequencies have been analysed through novelty analysis.

The presented results demonstrate the ability of P3P to filter out benign variances in the monitoring data induced by EOC and effectively conduct damage detection/quantification through the inspection of quality control charts.

Acknowledgements. The financial support and access to the monitoring data of the Trigno Bridge by Anas S.p.A are gratefully acknowledged. Furthermore, the structural mechanics section of KU Leuven is gratefully acknowledged for providing access to the Z24 Bridge dataset.

References

[1] ASCE, 2021 Report Card for America's Infrastructure, Technical Report, American Society of Civil Engineers, 2021.

[2] R. Woodward, D. W. Cullington, A. F. Daly, P. R. Vassie, P. Haardt, R. Kashner, R. Astudillo, C. Velando, B. Godart, C. Cremona, Bridge management in Europe (BRIME)–Deliverable D14–Final Report, Technical Report, 2001.

[3] F. Moreu, X. Li, S. Li, D. Zhang, Technical specifications of structural health monitoring for highway bridges: New Chinese structural health monitoring code, Front Built Environ 4 (2018) 10.

[4] ISIS Canada, Guidelines for Structural Health Monitoring, ISIS Canada, 2001, pp. 539–541.

[5] Ministero delle Infrastrutture e dei Trasporti Consiglio Superiore dei Lavori Pubblici, Linee guida per la classificazione e gestione del rischio, la valutazione della sicurezza ed il monitoraggio dei ponti esistenti, 2020.

[6] M. Moore, B. M. Phares, B. Graybeal, D. Rolander, G. Washer, J. Wiss, Reliability of visual inspection for highway bridges, volume I, Technical Report, Turner-Fairbank Highway Research Center, 2001.

[7] K. Gkoumas, F. L. Marques Dos Santos, M. Van Balen, A. Tsakalidis, A. Ortega Hortelano, M. Grosso, G. Haq, F. Pekár, Research and innovation in bridge maintenance, inspection and monitoring, Publications Office of the European Union (2019).

[8] R. Fuentes, E. J. Cross, P. A. Gardner, L. A. Bull, T. J. Rogers, R. J. Barthorpe, H. Shi, N. Dervilis, C. R. Farrar, K. Worden, Structural Health Monitoring and Damage Identification, Handbook of Experimental Structural Dynamics (2020) 1–72.

[9] C. P. Fritzen, Vibration–based structural health monitoring–concepts and applications, in: Key Engineering Materials, volume 293, Trans Tech Publ, 2005, pp. 3–20.

[10] F. B. Zahid, Z. C. Ong, S. Y. Khoo, A review of operational modal analysis techniques for in-service modal identification, J Braz Soc Mech Sci Eng 42 (2020) 1–18.

[11] A. Saisi, C. Gentile, M. Guidobaldi, Post-earthquake continuous dynamic monitoring of the Gabbia Tower in Mantua, Italy, Constr Build Mater 81 (2015) 101–112.

[12] C. Gentile, A. Saisi, Continuous dynamic monitoring of a centenary iron bridge for structural modification assessment, Front Struct Civ 9 (2015) 26–41.

[13] W. Soo Lon Wah, Y. T. Chen, G. W. Roberts, A. Elamin, Separating damage from environmental effects affecting civil structures for near real-time damage detection, Struct Health Monit 17 (2018) 850–868.

[14] F. Ubertini, N. Cavalagli, A. Kita, G. Comanducci, Assessment of a monumental masonry bell-tower after 2016 Central Italy seismic sequence by long-term SHM, B Earthq Eng 16 (2018) 775–801.

[15] F. Ubertini, C. Gentile, A. L. Materazzi, Automated modal identification in operational conditions and its application to bridges, Eng Struct 46 (2013) 264–278.

[16] G. Zini, M. Betti, G. Bartoli, A quality-based automated procedure for operational modal analysis, Mech Syst Signal Process 164 (2022) 108173.

[17] P. É. Charbonnel, Fuzzy-driven strategy for fully automated modal analysis: Application to the SMART2013 shaking-table test campaign, Mech Syst Signal Process 152 (2021) 107388.

[18] P. Cheema, M. M. Alamdari, G. A. Vio, F. L. Zhang, C. W. Kim, Infinite mixture models for operational modal analysis: An automated and principled approach, J Sound Vib 491 (2021) 115757.

[19] Y. He, J. P. Yang, X. F. Li, A three-stage automated modal identification framework for bridge parameters based on frequency uncertainty and density clustering, Eng Struct 255 (2022) 113891.

[20] E. M. Tronci, M. De Angelis, R. Betti, V. Altomare, Multi-stage semi-automated methodology for modal parameters estimation adopting parametric system identification algorithms, Mech Syst Signal Process 165 (2022) 108317.

[21] V. R. Gharehbaghi, E. Noroozinejad Farsangi, M. Noori, T. Y. Yang, S. Li, A. Nguyen, C. Málaga-Chuquitaype, P. Gardoni, S. Mirjalili, A critical review on structural health monitoring: definitions, methods, and

perspectives, Arch Comput Method E (2021) 1–27.

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

[23] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, D. J. Inman, A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications, Mech Syst Signal Process 147 (2021) 107077.

[24] B. Peeters, G. De Roeck, One–year monitoring of the Z24–Bridge: environmental effects versus damage events, Earthq Eng Struct Dyn 30 (2001) 149–171.

[25] A. Kita, N. Cavalagli, F. Ubertini, Temperature effects on static and dynamic behavior of Consoli Palace in Gubbio, Italy, Mech Syst Signal Process 120 (2019) 180–202.

[26] P. Gardner, L. A. Bull, N. Dervilis, K. Worden, Domain-adapted Gaussian mixture models for population-based structural health monitoring, J Civ Struct Health Monit (2022) 1–11.

[27] E. García-Macías, F. Ubertini, Integrated SHM Systems: Damage Detection Through Unsupervised Learning and Data Fusion, in: Structural Health Monitoring Based on Data Science Techniques, Springer, 2022, pp. 247–268.

[28] C. Cremona, J. Santos, Structural health monitoring as a big-data problem, Struct Eng Int 28 (2018) 243–254.

[29] E. García-Macías, F. Ubertini, Least Angle Regression for early-stage identification of earthquake-induced damage in a monumental masonry palace: Palazzo dei Consoli, Eng Struct 259 (2022) 114119.

[30] E. Reynders, M. Schevenels, G. De Roeck, MACEC 3.2: A Matlab toolbox for experimental and operational modal analysis-User's manual, Katholieke Universiteit, Leuven, 2011.

[31] L. M. S. Test, Lab Modal Analysis User Manual, Siemens Industry Software NV, 2015.

[32] ARTeMIS Extractor Pro User's Manual, Release 3.5, Structural Vibration Solutions, Aalborg, Denmark, 2006.

[33] E. García-Macías, F. Ubertini, MOVA/MOSS: Two integrated software solutions for comprehensive Structural Health Monitoring of structures, Mech Syst Signal Process 143 (2020) 106830.

[34] B. Peeters, G. De Roeck, One-year monitoring of the Z 24-Bridge: environmental effects versus damage events, Earthq Eng Struct Dyn 30 (2001) 149–171.

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

[36] F. Magalhães, Á. Cunha, Explaining operational modal analysis with data from an arch bridge, Mech Syst Signal Process 25 (2011) 1431–1450.

[37] P. Van Overschee, B. De Moor, Subspace identification for linear systems: Theory–Implementation–Applications, Springer Science & Business Media, 2012.

[38] C. Rainieri, G. Fabbrocino, Operational modal analysis of civil engineering structures, Springer, New York 142 (2014) 143.

[39] E. Reynders, G. De Roeck, Reference-based combined deterministic-stochastic subspace identification for experimental and operational modal analysis, Mech Syst Signal Process 22 (2008) 617–637.

[40] C. Rainieri, G. Fabbrocino, Influence of model order and number of block rows on accuracy and precision of modal parameter estimates in stochastic subspace identification, Int J Lifecycle Performance Engineering 1 (2014) 317–334.

[41] R. S. Pappa, K. B. Elliott, A. Schenk, Consistent-mode indicator for the eigensystem realization algorithm, J Guid Control Dyn 16 (1993) 852–858.

[42] M. Pastor, M. Binda, T. Harčarik, Modal assurance criterion, Procedia Eng 48 (2012) 543–548.

[43] E. Favarelli, A. Giorgetti, Machine learning for automatic processing of modal analysis in damage detection of bridges, IEEE Trans Instrum Meas 70 (2020) 1–13.

[44] C. M. Bishop, N. M. Nasrabadi, Pattern recognition and machine learning, volume 4, Springer, 2006.

[45] G. Schwarz, Estimating the dimension of a model, Ann Stat (1978) 461–464.

[46] B. Peeters, System identification and damage detection in civil engeneering, Ph.D. thesis, KU Leuven, 2000.

[47] F. Magalhães, Operational modal analysis for testing and monitoring of bridges and special structures, Ph.D. thesis, University of Porto, 2010.

[48] H. Hotteling, Multivariate quality control, illustrated by the air testing of sample bombsights, Techniques of statistical analysis (1947) 111–184.

[49] G. Steenackers, P. Guillaume, Structural health monitoring of the Z24 Bridge in presence of environmental changes using modal analysis, in: Proceedings of IMAC, volume 23.

[50] J. Zhang, E. Reynders, G. De Roeck, G. Lombaert, Model updating of periodic structures based on free wave characteristics, J Sound Vib 442 (2019) 281–307.

[51] EMPA work package, SIMCES Task A1 & A2 Longterm Monitoring and Bridge Tests, Technical Report, Dübendorf, Switzerland, 1999.