Engineering Structures 366 (2026) 123284

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

Population-based structural health monitoring of historical masonry bridges using spaceborne InSAR measurements☆

Wael Alahmad [a], Said Quqa [b,*], Cristina Gentilini [a]

[a] Department of Architecture, University of Bologna, Viale del Risorgimento 2, Bologna, 40136, Italy
[b] Department of Civil, Chemical, Environmental, and Materials Engineering, University of Bologna, Viale del Risorgimento 2, Bologna, 40136, Italy

ARTICLE INFO

Keywords:
Masonry arch bridge
Satellite data
Transfer learning
Long short-term memory
Population-based structural health monitoring
Damage detection

ABSTRACT

Italy's railway network includes nearly 10,000 masonry arch bridges, many of which now carry loads far exceeding their original design while undergoing material degradation, posing a major infrastructure management challenge. Conventional structural health monitoring relies on costly on-site instrumentation that cannot be deployed at this scale. This paper proposes a population-based SHM framework that identifies anomalous deformation in masonry arch bridges from satellite interferometric (InSAR) measurements, while keeping physically meaningful, engineering-interpretable indicators. For each bridge, Sentinel-1 line-of-sight displacements and environmental variables (air temperature and soil moisture) are condensed into compact statistical features describing the global displacement distribution, independent of the uncertain position of individual InSAR scatterers. Subspace alignment, a feature-based domain adaptation technique, is used to harmonize the baseline feature distributions across structures with different geometry. To overcome the scarcity of labeled anomalies, a simplified parametric bridge model is used to generate a population of synthetic bridges with imposed anomaly scenarios, on which a bidirectional long short-term memory regression network is trained. Through simulation-to-real transfer, the network is then applied to real data. The model outputs two physically interpretable descriptors, namely anomaly intensity and spatial extent. Validation is carried out on eight masonry arch railway bridges in Liguria, Italy. A persistent localized settlement identified on one structure is consistent with independent velocity maps. The framework offers a scalable, non-invasive tool for regional-scale screening and inspection prioritization of heritage bridge stocks.

1. Introduction

Masonry arch bridges, a hallmark of historical civil engineering, exemplify the structural efficiency and longevity achievable through traditional construction techniques. In Italy alone, the railway network comprises nearly 10,000 such bridges [1]. However, the discrepancy between the loads these structures were originally designed to carry and the significantly increased demands they now endure, along with material degradation and shifting environmental conditions, raises substantial concerns about their structural integrity. Ensuring their continued safety requires the implementation of innovative and scalable monitoring techniques.

Structural health monitoring (SHM) provides valuable information on the condition of civil structures and historical monuments through nondestructive methods [2–8]. A broad range of sensing technologies has been used to extract the structural displacements on which many SHM methods rely, from contact instrumentation, such as displacement transducers, and point-based geodetic techniques such as GNSS, to non-contact and low-cost alternatives. Among the latter, vision-based systems can measure multiple points simultaneously from a single camera [9], while smartphone-based sensing leverages the ubiquity of embedded MEMS sensors and cameras to ease large-scale data collection [10]. However, traditional SHM techniques generally involve high costs due to on-site sensor deployment and data management, and even these emerging low-cost solutions still require physical access to each structure or proximity to the monitored components.

Among emerging technologies, satellite-based radar imaging is notable for its growing availability and improving spatial and temporal resolution. This technique enables the estimation of millimeter-level deformations by leveraging the fundamentals of radar interferometry, offering a powerful tool for long-term monitoring [11–14].

Initially, interferometric synthetic aperture radar (InSAR) was used primarily to monitor changes in the ground surface, such as subsidence

☆ This article is part of a Special issue entitled: 'Masonry arch bridges' published in Engineering Structures.

* Corresponding author.
E-mail address: said.quqa2@unibo.it (S. Quqa).

https://doi.org/10.1016/j.engstruct.2026.123284

Available online 2 July 2026

and landslides [11,12], playing a crucial role in the assessment and mitigation of natural hazards. More recently, its effectiveness has been demonstrated in measuring infrastructure displacements that are highly correlated with on-site measurements and meaningful for structural analysis [15]. This level of reliability is attained despite the several noise sources that affect interferometric measurements, including atmospheric phase delays, temporal and geometric decorrelation, orbital and residual topographic errors, and phase-unwrapping errors [13], whose impact is largely mitigated by multi-temporal processing techniques. These advances make it possible to assess the structural condition of civil infrastructures and historical buildings [6,16,17].

A growing body of works has demonstrated the relevance of InSAR for heritage conservation. Selvakumaran et al. [16] developed an early anomaly detection framework and tested it on the Tadcaster Bridge, which revealed precursory deformation signals more than a month before its collapse. Alani et al. [17] integrated InSAR data with ground-penetrating radar to assess the 13th-century Aylesford Bridge, relating observed displacements to river water level variations. Gagliardi et al. [6] monitored the historic Rochester Bridge, detecting persistent subsidence patterns and identifying potential instability zones. Recent applications have also extended to heritage buildings: Bonaldo et al. [18] used SBAS-InSAR to monitor Palazzo Primoli in Rome, while Caprino et al. [19] applied MT-InSAR to the Scrovegni Chapel in Padua, identifying seasonal motion patterns and advocating for satellite–ground hybrid SHM frameworks.

In parallel, InSAR data have increasingly been used to support structural analysis and safety assessment. Guzman-Acevedo et al. [20] integrated Sentinel-1 observations with finite element modeling to estimate a risk-based safety index for the El Carrizo Bridge after seismic retrofitting. Farneti et al. [21,22] combined InSAR data with applied element method simulations to analyze the progressive failure of the Albiano-Magra Bridge.

Beyond these descriptive and physics-based applications, a further line of research has employed data-driven regression to cope with the limited and environmentally affected nature of SAR-derived bridge displacements. Entezami et al. [23] introduced an adaptive ensemble regression method, based on Gaussian process regression, to predict displacement time series from a limited number of SAR images while simultaneously suppressing environmental and operational variability. In a closely related work, Entezami et al. [24] proposed a two-stage deep regression framework coupling regressor selection with long short-term memory networks to predict and normalize SAR-derived displacement responses under both measured and unmeasured environmental and operational effects. These studies confirm the value of regression-based normalization and temporal learning for satellite-based bridge monitoring. However, they are formulated for response prediction at the level of individual structures.

A notable advantage of satellite-based approaches is the ease of obtaining displacement measurements over wide regions, enabling consistent and cost-effective monitoring of large infrastructure networks. In this context, population-based SHM (PBSHM) offers a promising, yet still underexplored, framework for InSAR-driven applications [25,26]. Unlike traditional SHM approaches that focus on individual assets, PBSHM leverages statistical relationships across groups of similar structures to generalize diagnostic information from a subset of the population to the rest of the individuals. This allows, for instance, the use of data from a bridge with a known damage condition to assess whether other bridges exhibit analogous anomalies. When structural variability limits direct comparisons, transfer learning techniques can be used to align the statistical distributions of features [27], enabling knowledge transfer across the population.

The first application of PBSHM using satellite data was recently proposed by Quqa et al. [25], who analyzed a set of similar bridges along the same river and exploited their shared environmental exposure to isolate structure-specific anomalies. That study demonstrated the ability to discriminate among different anomaly types in real datasets, including settlements, uplifts, flood-related effects, and data outliers, by transferring information from a population of synthetic bridge models.

Despite the strong potential of PBSHM, feature harmonization techniques can reduce the direct physical interpretability of the resulting indicators, complicating threshold definition and informed maintenance prioritization. This limitation also hampers the quantitative characterization of detected anomalies, such as the estimation of settlement magnitude, which is more naturally expressed in the original displacement domain.

This study proposes a PBSHM framework for detecting localized settlements in masonry bridges using satellite-derived displacement measurements and environmental data, while restoring engineering interpretability through a compact and physically meaningful anomaly index. For each bridge in the monitored population, displacement and environmental observations, entirely obtained from satellite sources, are condensed into a consistent set of instantaneous statistical features. This formulation enables bridges with different geometries and sizes to be compared through a uniform description of their global displacement response.

Based on these features, a regression-based learning model is designed to characterize structural anomalies in terms of physically interpretable quantities, namely the intensity and spatial extent of settlements or uplifts.

To address the limited availability of labeled structural anomalies in real monitoring data, the proposed framework introduces a simplified parametric bridge model. The model is conceived as an archetype of masonry bridges and is defined through an analytical formulation that requires only coarse geometric information representative of the monitored population. The model is further refined through calibration with real measurements to reproduce the dominant environmentally driven displacement behavior more accurately. Then, it is used to generate a wide range of synthetic anomalous scenarios.

To enable simulation-to-real knowledge transfer, subspace alignment, a transfer learning technique, is employed to harmonize feature distributions across bridges and to project both synthetic and real data onto a shared latent space. Here, the regression model, implemented as a bidirectional long short-term memory neural network, is trained exclusively on the synthetic scenarios. Subsequently, the model is applied to real bridge data projected onto the same latent space to identify potential anomalies. The network outputs continuous estimates of both the magnitude and spatial extent of settlement/uplift, thereby mapping latent-space patterns back onto physically interpretable quantities that are directly relevant for engineering assessment and maintenance decision-making.

Section 2 details the proposed PBSHM methodology. Section 3 describes the study area and the datasets used in the analysis. Section 4 introduces the simplified parametric bridge model adopted for synthetic data generation. Section 5 presents and discusses the obtained results, and Section 6 concludes the paper and outlines directions for future research.

2. Proposed methodology

The proposed PBSHM framework consists of the following main steps:

1. Feature extraction: transformation of satellite displacement time series and environmental measurements into a compact set of statistical descriptors for each bridge;
2. Feature harmonization: application of subspace alignment (SA) to harmonize feature distributions across different bridges (both real and synthetic) and enable population-level comparison;
3. Anomaly characterization: quantitative characterization of structural anomalies using synthetic data labels.

Beyond the new case study examined in this work, the main methodological contributions are twofold. The first concerns the anomaly characterization step, described in detail in Section 2.3, which introduces a regression-based strategy to enable quantitative inference across bridges with heterogeneous geometries. The second novelty is the simplified parametric bridge model specifically developed for masonry arch bridges, presented in Section 4.

2.1. Feature extraction

The proposed framework relies on two primary data sources: satellite-derived displacement measurements and environmental variables, both obtained from open-access Earth Observation services. Displacement data are provided by the European Ground Motion Service (EGMS), a Copernicus initiative based on radar imagery acquired by the Sentinel-1 constellation [28]. Sentinel-1 operates in C-band at a central frequency of 5.405 GHz (wavelength 5.6 cm) and acquires images in Interferometric Wide Swath (IW) mode, with a spatial resolution of approximately 5 m (range) × 20 m (azimuth) over a 250 km swath.

This study exploits the Level-2b calibrated product, which delivers geodetically consistent line-of-sight (LOS) displacement time series for coherent targets (persistent scatterers, PSs) identified through PS-InSAR processing [29,30]. The third EGMS data release is used, covering the period from January 2019 to December 2023. However, only the interval from January 2019 to December 2021 is processed in the present analyses, as the revisit cycle during this period is nearly uniform at six days. Following the in-flight anomaly of Sentinel-1B on December 23, 2021, the revisit interval increased to 12 days [31], resulting in a change in temporal sampling that could affect the performance of the proposed BiLSTM network.

Only descending-orbit acquisitions (north-to-south trajectory) are considered. Because Sentinel-1 data are acquired in bursts across adjacent sub-swaths, slight variations in acquisition geometry (mainly related to track and incidence angle) occur between sub-swaths. To enhance spatial coverage, data from multiple sub-swaths are jointly exploited in this study.

Assuming that bridge displacements are predominantly vertical, LOS measurements are converted into estimated vertical displacements according to [21,32]:

$$\hat{d}(\mathrm{PS}_j, t) = \frac{d_{\mathrm{LOS}}(\mathrm{PS}_j, t)}{\cos\beta} \tag{1}$$

where $d_{\mathrm{LOS}}(\mathrm{PS}_j, t)$ is the LOS displacement of the jth PS at time t, and β is the incidence angle. The resulting displacement time series for each satellite track are then linearly interpolated to a daily temporal resolution.

A major challenge when using InSAR for bridge monitoring is the difficulty of associating PSs with specific structural components, especially in medium-resolution datasets where target localization uncertainty is significant [28]. To overcome this limitation, the proposed framework adopts a holistic feature-based approach that summarizes the global displacement response of each bridge [25]. Accordingly, for each bridge and each time step, statistical features are computed from the aggregated PS displacement records, considering all available satellite tracks after temporal interpolation. The selected features, listed in Table 1, characterize the spatial distribution of displacement within the bridge footprint and form the basis for subsequent harmonization and anomaly characterization.

These features characterize the global displacement distribution within the bridge footprint without requiring the assignment of each PS to a specific structural component. The mean (f_1) and median (f_7) quantify the overall displacement of the monitored portion, the median being the outlier-robust counterpart of the mean; the standard deviation (f_2) reflects the spatial heterogeneity of the displacement field, distinguishing uniform from differential movements; the skewness (f_3) detects asymmetric responses, such as settlement concentrated on one portion of the structure; the kurtosis (f_4) is sensitive to localized anomalies affecting only a few PSs, which produce heavy-tailed distributions; the minimum and maximum (f_5, f_6) capture the extreme deformation values; and the displacement with maximum temporal variance (f_8) identifies the most responsive measurement within the footprint.

Table 1
Set of selected features.

Feature	Description
f_1	Mean of the instantaneous displacement distribution
f_2	Standard deviation of the instantaneous displacement distribution
f_3	Skewness of the instantaneous displacement distribution
f_4	Kurtosis of the instantaneous displacement distribution
f_5	Minimum instantaneous displacement
f_6	Maximum instantaneous displacement
f_7	Median of the instantaneous displacement distribution
f_8	Displacement with maximum temporal variance
f_9	Air temperature at bridge location
f_{10}	Soil moisture at bridge location

Environmental variables are also incorporated within the feature list to contextualize displacement behavior. Specifically, air temperature at 2 m height and volumetric soil moisture in the upper 0–7 cm layer are retrieved from ERA5-Land, which provides hourly global data at a spatial resolution of 9 × 9 km [33], obtained from the Open-Meteo historical weather API [34]. These variables are downsampled to the same temporal grid as the displacement series.

To provide a consistency check for the adopted environmental inputs, the ERA5-Land variables were compared with local on-site measurements available from the regional meteo-hydrological monitoring portal of Regione Liguria [35]. The comparison was carried out using the Genova-Pontedecimo station, managed by ARPAL CFMI-PC and located in the Polcevera basin about 500 m from Bridge 5 (see Section 3). Because the portal provides only recent records, the comparison covers a limited interval. Nonetheless, the results in Fig. 1 indicate good consistency between the local observations and the adopted large-scale environmental descriptors. Specifically, Fig. 1(a) compares the on-site and ERA5-Land air temperatures, which are in close agreement (Pearson correlation 0.905). Since soil moisture is not measured at the station, Fig. 1(b) instead compares the locally measured river water level with the ERA5-Land soil-moisture variable. The two series both reflect the hydrological state of the Polcevera basin and are indeed strongly correlated (Pearson correlation 0.708). These results indicate that the adopted reanalysis variables capture a substantial part of the local environmental variability.

2.2. Feature harmonization

Differences in bridge geometry, orientation, and local environmental exposure introduce structure-specific variability in the extracted features, which may hinder direct cross-bridge comparison. To mitigate this effect, feature harmonization is performed using SA [36]. The performance of SA for domain adaptation in this context is reported in Appendix A, together with a comparison against alternative strategies.

Within the taxonomy of Pan and Yang [27], the adopted strategy is a homogeneous, feature-based, transductive domain-adaptation method: the learning task (regression of anomaly intensity and extent) is shared across domains. The domains (the different bridges, and synthetic versus real data) are represented in the same feature space,

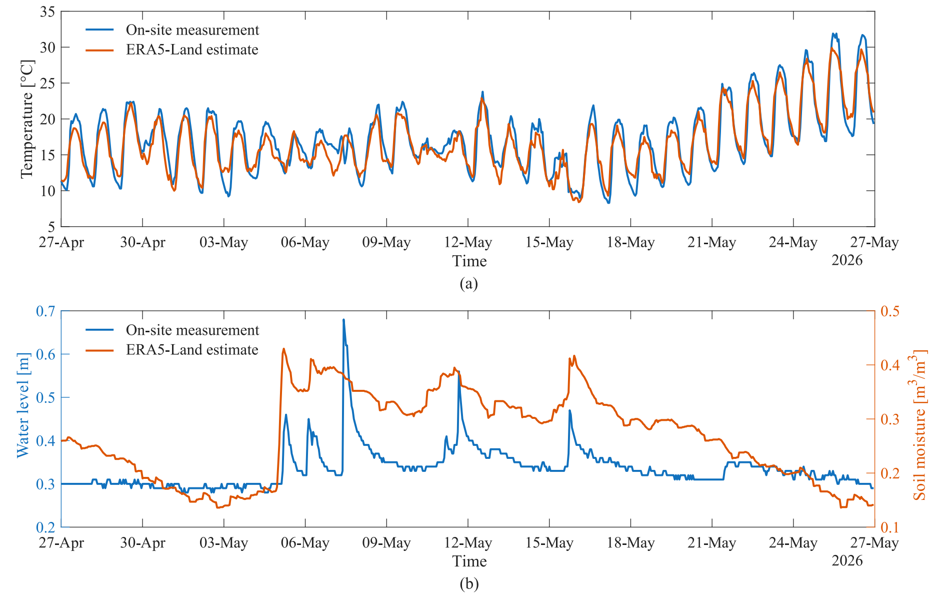

Fig. 1. Comparison between on-site environmental measurements from the Regione Liguria monitoring portal and the corresponding ERA5-Land variables used in this study, near Bridge 5: (a) air temperature measured at the Genova-Pontedecimo station versus ERA5-Land air temperature, (b) river water level measured at the Genova-Pontedecimo hydrometric station versus ERA5-Land soil moisture.

while labels are available only in the synthetic source domain. The quantity transferred across structures is the feature representation, that is the dominant directions of baseline variability and the associated environmental-displacement relationship learned on the synthetic data (*what* to transfer). This is achieved by linearly aligning the baseline feature subspaces of the different bridges into a common latent space (*how* to transfer, formalized below). The transfer is justified by the fact that bridges share the same structural typology and regional environmental exposure, and hence a comparable baseline behavior (*when* to transfer).

SA operates under the assumption that, although feature distributions may differ across bridges, the dominant patterns of baseline variability share a common structure. For each bridge s, features extracted during a selected baseline interval are assembled into a matrix $\mathbf{F}_s \in \mathbb{R}^{D\times K}$, where $K = 10$ is the number of features (see Table 1) and D is the number of time samples. Each feature is standardized by removing its mean and scaling to unit variance. The corresponding normalization parameters are stored and applied consistently to all subsequent data.

Principal component analysis (PCA) is then applied to $\mathbf{F}_s$ to obtain the basis $\mathbf{V}_s \in \mathbb{R}^{K\times r}$ spanning the dominant r directions of baseline variability. One bridge is selected as the reference (source) domain, with basis $\mathbf{V}_1$, while all other bridges act as target domains. Alignment is achieved through the transformation

$$\mathbf{W}_s = \mathbf{V}_s\mathbf{V}_s^\top\mathbf{V}_1 \tag{2}$$

which maps the target subspace onto the reference one. For a new feature vector $\mathbf{f}_s(t) \in \mathbb{R}^{K\times 1}$, the projection into the aligned latent space is given by

$$\mathbf{p}_1(t) = \mathbf{f}_1^\top\mathbf{V}_1 \tag{3a}$$

$$\mathbf{p}_s(t) = \mathbf{f}_s^\top\mathbf{W}_s \tag{3b}$$

In this work, no dimensionality reduction is performed, and the full feature space is retained ($r = K$).

It is important to emphasize that SA aligns baseline variability. Nevertheless, it does not guarantee that identical physical anomalies, such as the same settlement magnitude, map to the same position in the latent space when they affect bridges with different geometries (e.g., due to variations in pier height, see Section 3). This limitation precludes the use of classification strategies based on absolute latent-space proximity, as adopted in Alahmad et al. [37]. Nevertheless, SA remains essential to reduce inter-bridge bias and establish a common reference frame in which subsequent learning-based anomaly characterization can be performed (see Appendix A).

2.3. Anomaly characterization

Anomaly characterization is based on the observation that, once baseline variability has been harmonized, structural anomalies appear as persistent deviations from the baseline statistical behavior governed by environmental drivers. While the directions and amplitudes of these deviations in the latent space may vary across bridges due to geometric heterogeneity, they consistently alter the statistical structure of the extracted features, becoming more easily learnable by a regression model.

Accordingly, a bidirectional long short-term memory (BiLSTM) network is adopted for anomaly characterization. This regression-based approach removes the need to define discrete anomaly labels a priori and enables continuous estimation of physically meaningful anomaly characteristics. When trained on harmonized features derived from a synthetic population of structurally heterogeneous bridges, the model learns the range of correlations between environmental drivers and displacement-based features, thereby supporting quantitative characterization of anomalies in real bridges through simulation-to-real knowledge transfer.

Structural anomalies detectable with medium-resolution InSAR data (restricted here to the vertical displacement component) are assumed to primarily manifest as settlements or uplifts. The BiLSTM network is therefore designed to simultaneously estimate (i) the anomaly intensity, expressed as the average settlement or uplift magnitude, and (ii) the spatial extent of the affected region, quantified as the percentage of involved PSs. This study assumes that PSs are approximately uniformly distributed along the deck. Under this condition, the percentage of affected PSs provides a consistent proxy for the fraction of the structure affected by the anomaly, so that the extent label retains a comparable physical interpretation across bridges of different geometry.

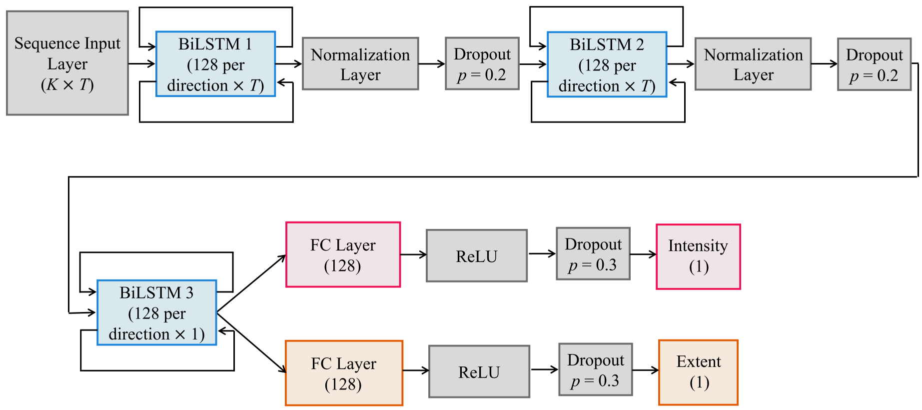

Fig. 2. Neural network architecture.

By processing feature sequences in both forward and backward temporal directions, the bidirectional architecture exploits the full temporal context of each analysis window, facilitating discrimination between persistent, non-environmental deviations associated with structural anomalies and transient or cyclic variations driven by environmental effects.

The architecture of the proposed dual-output regression network is summarized below:

- Input layer: A sequence input layer receives a fixed-length window of $T = 12$ days, each represented by 10 harmonized features. This window length was chosen so that each analysis window carries information related to 3 or 4 independent Sentinel-1 acquisitions. At the nominal 6-day revisit, the window contains 2 acquisitions in most cases. In addition, because the series is interpolated to a daily grid, the values within the window also reflect the trends toward the immediately preceding and following acquisitions.
- Recurrent backbone: The backbone consists of three BiLSTM layers, each with 128 units per direction (256-dimensional concatenated output). The first two BiLSTM layers operate in "sequence mode", producing a feature vector at each time step; each of these layers is followed by layer normalization and dropout ($p = 0.20$) to enhance stability and generalization. The third BiLSTM layer operates in "last mode", returning a single 256-dimensional context vector summarizing the entire T-step input sequence. A performance comparison between the selected BiLSTM-based backbone and other alternatives is reported in Appendix A.
- Dense prediction heads: The final context vector is fed into two independent branches. Each branch comprises a fully connected (FC) layer with 128 units, a ReLU activation, and a dropout layer ($p = 0.30$). Each branch ends with a scalar neuron: the first predicts the anomaly intensity (settlement or uplift), and the second predicts the spatial extent of the affected region.

Fig. 2 shows the complete architecture, highlighting the recurrent backbone with its normalization and dropout blocks, and the dual-head design used for simultaneous multi-output regression.

The two scalar outputs are concatenated to form the prediction vector $[\hat{y}_{\text{int}}, \hat{y}_{\text{ext}}]$. Training is performed by minimizing a weighted mean-squared error (MSE) over a mini-batch of B windows. For each window i, the network outputs predictions $\hat{y}_{\text{int},i}$ and $\hat{y}_{\text{ext},i}$, with ground-truth values $y_{\text{int},i}$ and $y_{\text{ext},i}$. The batch loss is defined as

$$\mathcal{L} = \alpha \cdot \frac{1}{B}\sum_{i=1}^{B}\left(\hat{y}_{\text{int},i} - y_{\text{int},i}\right)^2 + \beta \cdot \frac{1}{B}\sum_{i=1}^{B}\left(\hat{y}_{\text{ext},i} - y_{\text{ext},i}\right)^2 \quad (4)$$

where $\alpha, \beta \geq 0$ are task weights satisfying $\alpha + \beta = 1$. In this study, equal weighting is adopted, i.e., $\alpha = \beta = 0.5$. Before computing the loss, both target variables (intensity and extent) are normalized to the interval $[0, 1]$. All network hyperparameters are selected based on performance on a validation dataset, as discussed in Appendix B, and are reported in detail in Section 5.1.

The training data for the network are generated using the simplified parametric bridge model described in Section 4. Before introducing this model and the specific training dataset, the next section presents the real case study on which the parametrization is based.

Table 2
Details of the considered bridges.

Label	Latitude	Longitude	N. of arches	N. of PSs
Bridge 1	44.46010	8.75314	6	24
Bridge 2	44.46074	8.76878	9	63
Bridge 3	44.44951	8.81954	8	74
Bridge 4	44.47247	8.89263	8	85
Bridge 5	44.48332	8.89573	6	21
Bridge 6	44.50854	8.90746	6	12
Bridge 7	44.52673	8.91307	7	68
Bridge 8	44.61596	8.66013	6	22

3. Case study

This study focuses on eight historical masonry arch railway bridges located near Genoa, in northwestern Italy. The structures are distributed across rural and semi-urban settings, ranging from the inland hills around Novi Ligure to the coastal areas of Genoa and Voltri (Fig. 3).

These bridges are representative examples of early 20th-century transportation infrastructure and share broadly similar structural configurations, though they differ in total length, height, number of arches, and PS density. Table 2 reports the geographical coordinates of each bridge, along with the number of arches covered by PSs and the total number of PSs considered in this study. It is important to note that, for some structures, PSs are available only on a portion of the bridge. Consequently, the number of arches listed in Table 2 refers exclusively to the part of the structure where PSs are present, which may not coincide with the full arch count.

This diverse portfolio provides a suitable testbed for evaluating how environmental drivers (such as air temperature and soil moisture) affect deformation patterns, and for studying how settlements or uplifts influence the statistical distribution of displacement.

4. Synthetic parametric bridge model

A parametric model built on the formulation proposed by Alahmad et al. [37,38] is adapted in this study to reproduce the dominant physical mechanisms governing vertical displacements in masonry arch

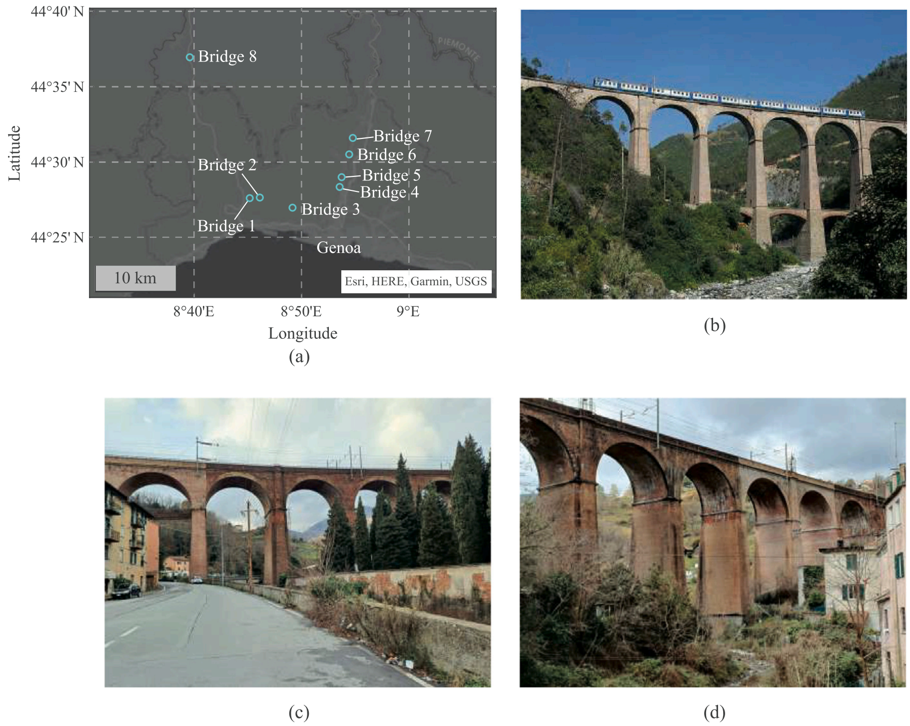

Fig. 3. Bridge portfolio: (a) location of the studied bridges, (b) view of Bridge 3, (c) view of Bridge 4, and (d) view of Bridge 7.

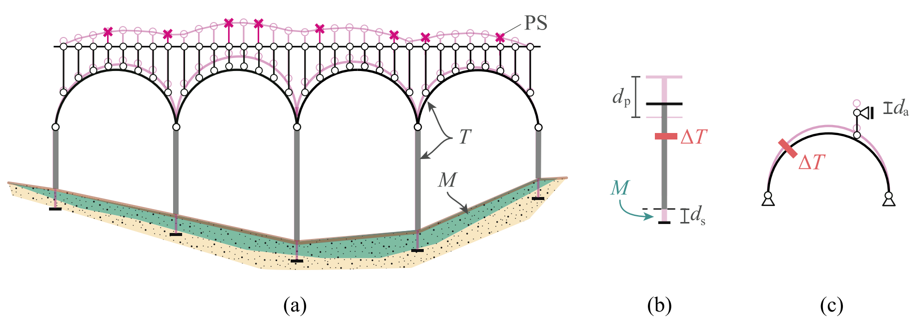

Fig. 4. Scheme of the parametric model: (a) global model, (b) detail of the pier, and (c) detail of the arch.

bridges. Rather than a continuum or finite-element representation of the masonry, the bridge is described through a reduced, beam-like kinematic model in which the deck displacement at each PS is expressed as a closed-form function of a few geometric quantities and environmental drivers. This choice deliberately avoids the definition of homogenized masonry material properties, whose reliable identification for historical masonry would require extensive in-situ surveys and material testing rarely available at the regional scale targeted here. Instead, the combined effect of material behavior, cross-sectional geometry, and boundary conditions is lumped into a small set of effective coefficients, which are calibrated directly against the measured displacement time series. These coefficients are therefore phenomenological rather than physical quantities, and the model is required only to reproduce the dominant global vertical displacement trends with the accuracy and spatial sampling of Sentinel-1 InSAR, not the local stress–strain response of the masonry. Residual discrepancies between the synthetic and real displacements are further mitigated by the subspace-alignment step, which harmonizes the feature distributions before the regression model is trained for anomaly characterization.

The role of this model is twofold. First, it enables controlled assessment of the PBSHM methodology introduced in Section 2 by generating synthetic displacement responses under realistic environmental conditions. Second, it provides the synthetic training data required for the simulation-to-real learning strategy adopted to characterize anomalies in real bridges.

Several modifications are introduced to tailor the original formulation to the context of railway bridges. In contrast to the historical portico structure considered in Alahmad et al. [37,38], the bridges examined here are characterized by piers with non-uniform heights, which significantly influence thermally induced displacement patterns. To account for this effect, the model assigns a specific pier height to each support, thereby better reflecting the actual bridge geometry. Deformations of infill walls are neglected, as their contribution is assumed to be small compared to the vertical displacements of persistent scatterers induced by pier deformation. Furthermore, owing to the large cross-sectional dimensions of bridge components, humidity-related effects are considered negligible relative to temperature-driven mechanisms. Consequently, only air temperature and soil moisture are retained as environmental drivers in the adapted model. A schematic representation of the resulting model configuration is provided in Fig. 4.

In the model, a set of geometric parameters must be defined, including the height of each pier (h_p), arch span (l), total bridge length (l_b), and number of arches (N_a). These parameters can be selected to provide a representative description of the bridges under monitoring,

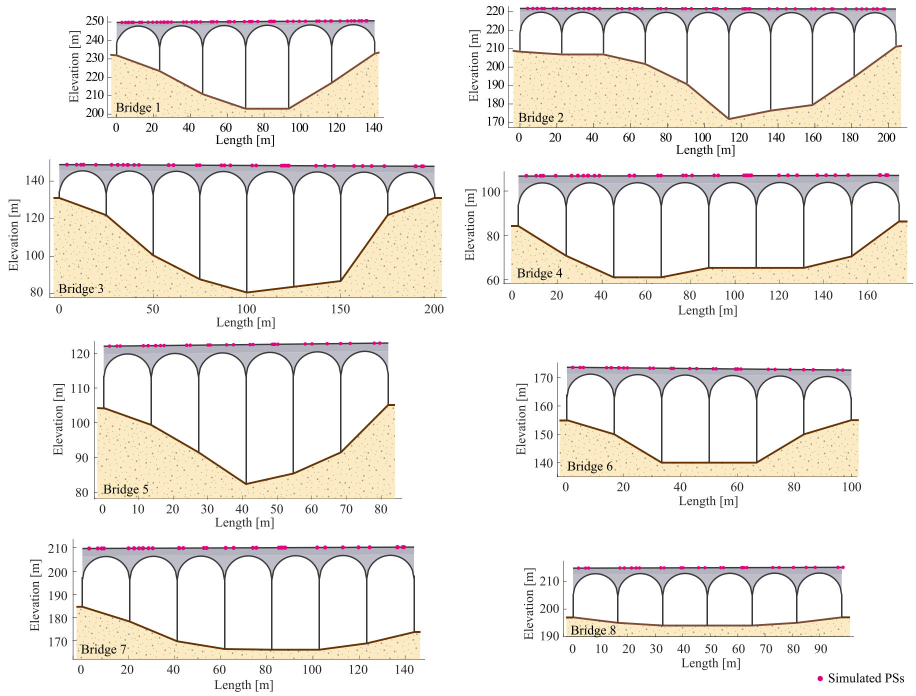

Fig. 5. Geometry of the simplified models. The distribution of PSs is represented with dots on the decks.

allowing the synthetic displacement fields to remain compatible with the spatial distribution of PSs observed in real data.

In the proposed formulation, the vertical displacement of each PS on the bridge deck results from the combined action of four phenomena:

1. Temperature-induced expansion/contraction of piers: Each pier is modeled as an axially expanding or contracting element in response to air temperature. The vertical displacement of a pier is expressed as

$$d_{\mathrm{p}}(t) = \alpha_T h_{\mathrm{p}} T(t) \tag{5}$$

where $T(t)$ is the air temperature, h_{p} the pier height, and α_T an equivalent linear thermal expansion coefficient. Given the relatively low temporal resolution of Sentinel-1 displacement data, thermal inertia is neglected.

2. Temperature-induced deformation of arches: Arch deformation is modeled as the result of uniform thermal strain along the arch axis, superimposed on the horizontal thrust response. Each arch is idealized as semicircular, and due to the repetitive geometry, each pier lies on a symmetry axis, constraining the arch to vertical displacement only:

$$d_{\mathrm{a}}(\theta_j, t) = \alpha_T\, T(t) \left(\pi l \sin^2\theta_j + y_j\right) \tag{6}$$

where l is the arch span, θ_j is the angular position of the jth point on the arch relative to the springing point, and y_j is the vertical coordinate of that point.

3. Soil-moisture-induced displacement: Soil–structure interaction is represented through a simplified formulation in which the displacement response is driven by the measured soil moisture $M(t)$:

$$d_s(t) = \alpha_M\, M(t - \tau_M) \tag{7}$$

where α_M is a stiffness-like coefficient and τ_M a time-delay parameter accounting for the lag between moisture variation and its structural effect. Although simplified, this formulation captures the observed correlation between vertical displacement and soil moisture. A more detailed explanation of this model is reported in Alahmad et al. [37].

4. Measurement noise: A stochastic term ϵ is added to emulate variability and uncertainty in real InSAR measurements.

The total vertical displacement modeled for PS_j at time t is therefore the sum of the four mentioned contributions (see Fig. 4), namely:

$$d(\mathrm{PS}_j, t) = d_{\mathrm{p}}(t) + d_{\mathrm{a}}(\theta_j, t) + d_{\mathrm{s}}(t) + \epsilon(t) \tag{8}$$

4.1. Model calibration

Calibration of the environmental parameters α_T, α_M, and τ_M can be performed to further reduce discrepancies between the synthetic displacement response and that observed on real bridges.

A realistic population of synthetic bridges is constructed by developing a calibrated model for each bridge described in Section 3. Geometric parameters are extracted using the measurement tools available in Google Earth Pro [39]. Fig. 5 presents the resulting simplified representations of the synthetic bridge population.

The spatial distribution of simulated PSs does not replicate that observed in the real datasets (illustrated for an example bridge and different satellite tracks in Fig. 6). Instead, a uniform density of 5 PSs per span is adopted, with scatterer locations assigned randomly along the deck. This density is derived from the ratio between the total number of PSs detected across the eight real bridges and the total number of spans, providing a representative average for the synthetic population.

Environmental parameters can be calibrated through a grid-search procedure. In this study, candidate values for the thermal expansion coefficient are sampled within the interval $\alpha_{\mathrm{T}} = [10^{-6}, 10^{-5}]°\mathrm{C}^{-1}$, while values for the soil-moisture coefficient are selected within $\alpha_{\mathrm{M}} = [-0.009, -0.001]$ m. Both intervals are discretized into 20 equally spaced values. For the time-delay parameter τ_{M}, positive lags up to 100 days (i.e., accounting for delayed structural response to environmental variations) are tested in increments of 10 days. The candidate intervals for the calibration parameters were selected empirically after preliminary tests over wider ranges. Since the optimal values for most bridges converged within the intervals finally adopted, the same bounds were used for all bridges to ensure a consistent calibration procedure.

Fig. 6. Distribution of PSs on Bridge 4 in two different satellite tracks (66 and 168).

For each parameter combination, the discrepancy between the synthetic and real responses is quantified through the root mean square error (RMSE) between the mean vertical displacement of all available PSs located within the bridge footprint, $\mu_r(t)$, and that of all simulated PSs distributed along the corresponding bridge model, $\mu_s(t)$. This choice is consistent with the proposed methodology, which relies not on individual PS positions but on statistical descriptors of the displacement distribution at the bridge level, and therefore does not require the simulated PSs to coincide spatially with the real ones. The comparison is carried out over the calibration interval from April 20, 2019, to September 25, 2021, after converting the measured LOS displacements into vertical displacements according to Eq. (1).

The cost function minimized during calibration is defined as

$$J(\alpha_T, \alpha_M, \tau_M) = \sqrt{\frac{1}{t_{tot}} \sum_{t=1}^{t_{tot}} \left(\mu_s(t) - \mu_r(t)\right)^2} \tag{9}$$

with:

$$\mu_s(t) = \frac{1}{N_s} \sum_{j=1}^{N_s} d(\mathrm{PS}_j, t) \tag{10a}$$

$$\mu_r(t) = \frac{1}{N_r} \sum_{j=1}^{N_r} \hat{d}(\mathrm{PS}_j, t) \tag{10b}$$

Here, t_{tot} is the total number of time samples for real displacement measurements $\hat{d}(\mathrm{PS}_j, t)$ in the calibration interval, while N_s and N_r denote the number of PSs used in the simulation and those available in the real dataset, respectively. Any linear trend (such as a uniform movement affecting the entire monitored area) is removed from $\mu_r(t)$ before fitting. Thus, in Eq. (10b), $\mu_r(t)$ represents the detrended mean displacement. On the other hand, $\mu_s(t)$ has no linear trend by construction.

Table 3 reports the optimized parameters for each synthetic bridge model, together with the number of simulated PSs (N_s) and the corresponding RMSE values. Measurement noise is subsequently added to the time histories of individual simulated displacement measures as a zero-mean Gaussian random variable, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, with $\sigma_\varepsilon = 2.5$ mm. This value is consistent with the displacement uncertainty typically reported for highly coherent Sentinel-1 InSAR observations [30].

A qualitative illustration of the performance of the parametric bridge model is shown in Fig. 7 for Bridges 1 and 2, where the detrended mean displacement obtained from the real measurements is compared with that generated by the calibrated model. The synthetic time series reproduce the seasonal trends observed in the data, yielding an average RMSE of approximately 2 mm (Table 3). These results indicate that, despite its simplified formulation, the parametric model captures the dominant environmental mechanisms governing vertical bridge displacements.

Table 3
Calibrated parameters of the simulated bridges.

Bridge label	α_T [°C^{-1}]	α_M [m]	τ_M [days]	N_s	RMSE [m]
Bridge 1	9.5263×10^{-6}	−0.0090	70	30	0.0016
Bridge 2	7.1579×10^{-6}	−0.0090	100	45	0.0013
Bridge 3	5.7368×10^{-6}	−0.0090	80	40	0.0014
Bridge 4	1×10^{-6}	−0.0052	100	40	0.0011
Bridge 5	1×10^{-6}	−0.0010	0	30	0.0020
Bridge 6	7.1579×10^{-6}	−0.0090	50	30	0.0036
Bridge 7	1×10^{-6}	−0.0018	100	35	0.0013
Bridge 8	7.1579×10^{-6}	−0.0082	50	30	0.0020

Notably, the calibrated parameters reported in Table 3 vary substantially across bridges. Since the environmental drivers are taken from ERA5-Land reanalysis models, which integrate satellite observations with complex hydro-geological models, the resulting parameters should not be interpreted as physical quantities in a strict sense. Instead, they act as effective coefficients that encapsulate a range of complex environmental influences, with the purpose of reproducing the displacement trends as accurately as possible.

Moreover, the values of α_T for Bridges 4, 5, and 7 converge to the lowest bound imposed during calibration. Although smaller values might yield slightly improved fits, the objective of the framework is not to achieve a perfect bridge-by-bridge representation. Rather, it aims to generate realistic synthetic data reflecting plausible environmental responses across the population. From this perspective, the calibrated models for Bridges 4, 5, and 7 can be viewed as proxies for other structures exhibiting similar displacement behavior.

4.2. Simulated anomaly scenarios

Once the synthetic models are available, they are used to simulate structural anomalies that affect the displacements observable in satellite measurements, for example by imposing a vertical translation on a subset of piers to represent settlement or uplift. While environmental effects evolve over time, the imposed structural anomalies are modeled as static and persistent offsets superimposed on the displacement time series. Table 4 summarizes the anomaly types modeled in this study. Each scenario is simulated at four intensity levels (5, 10, 15, and 20 mm), thereby defining the synthetic training dataset used by the neural network described in Section 2.3 to detect and characterize analogous anomalies in real structures. These intensity levels were chosen to be compatible with the millimeter-level sensitivity of InSAR measurements while remaining representative of early-stage phenomena, whose overall displacement effect is of the same order of magnitude as the environmental oscillations (which produce displacement variations of about 10 mm, as shown in Fig. 7).

Different structural anomalies, including scour-induced movements, foundation displacements, and landslide-driven effects, produce deck deformations that InSAR observes as line-of-sight displacements. Because the proposed methodology relies exclusively on the vertical component of the displacement, each anomaly is represented only through its vertical displacement signature, independently of the mechanism that generated it. The modeled scenarios are thus intentionally agnostic to the underlying mechanism, and the attribution of the physical cause of an anomaly is deferred to subsequent targeted inspection. Consequently, mechanisms that generate differential movements of the deck, such as pier tilting or asymmetric foundation settlement, are captured by the framework and reported generically as settlement or uplift rather than classified by type.

Notably, since the modeled bridges have different numbers of spans, and PSs are randomly distributed over the simplified bridge models, the percentage of PSs affected by each anomaly scenario varies from bridge to bridge. Table 5 reports, for each structure, the proportion of PSs falling within the region influenced by 1-, 2-, or 3-pier anomalies. This variability introduces diversity in the synthetic training data and supports the generalization capability of the trained regressor across bridges with heterogeneous geometries.

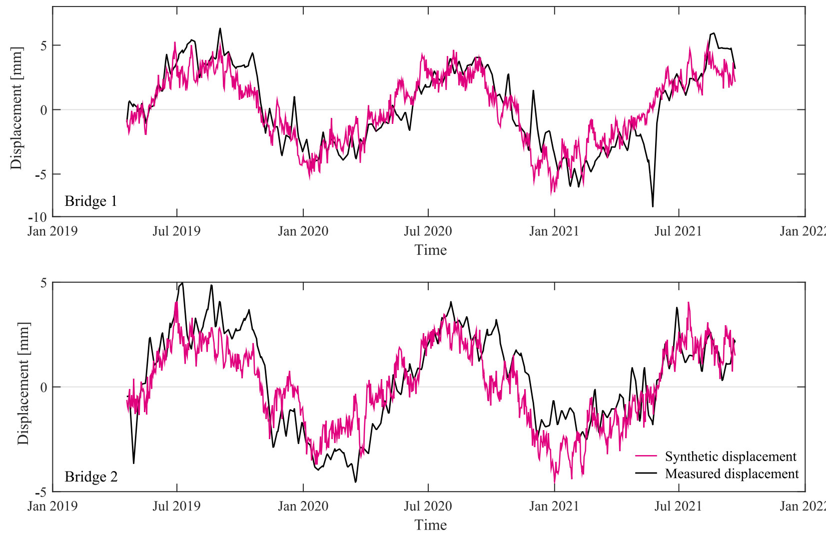

Fig. 7. Comparison between the detrended mean of all real and simulated PS displacements after calibration for Bridges 1 and 2.

Table 4
Anomaly scenarios modeled in this study. Each scenario is simulated at four intensity levels (5, 10, 15, and 20 mm).

Anomaly label	Description
S1	Settlement of one pier
U1	Uplift of one pier
S2	Settlement of two piers
U2	Uplift of two piers
S3	Settlement of three piers
U3	Uplift of three piers

Table 5
Percentage of PSs affected under pier-shift scenarios in Table 4.

Bridge label	1 pier (%)	2 piers (%)	3 piers (%)
Bridge 1	33.33	56.67	70.00
Bridge 2	17.78	31.11	37.78
Bridge 3	25.00	40.00	47.50
Bridge 4	35.00	42.50	60.00
Bridge 5	33.33	56.67	70.00
Bridge 6	33.33	46.67	70.00
Bridge 7	25.71	37.14	62.86
Bridge 8	33.33	46.67	70.00

5. Application

5.1. Population of synthetic bridges

This section presents the results obtained by evaluating the proposed methodology on the population of synthetic bridges introduced in Section 4, using imposed anomaly scenarios to assess its performance. The testbed dataset is generated using the calibrated parametric models driven by real environmental data. For each bridge, 600-day displacement time histories starting from January 10, 2019, are simulated for a healthy baseline condition and for all anomaly scenarios listed in Table 4, each treated independently.

For each bridge and anomaly scenario, daily anomaly-sensitive features, comprising both statistical descriptors and environmental variables, are extracted following the procedure described in Section 2.1. These features are then harmonized using the corresponding healthy baseline condition, in accordance with the SA strategy outlined in Section 2.2. Notably, the source domain needs to be chosen as the bridge providing the richest labeled information on the anomaly scenarios of interest. Since all synthetic bridges considered in this study were generated with the same anomaly classes and intensity levels (see Section 4), they contain equivalent labeled information. In the present study, this source domain is instantiated by the synthetic dataset of Bridge 1.

The effect of SA on the baseline features is first illustrated on Bridge 1 by comparing its real and synthetic baseline data. Fig. 8 reports a pairwise feature comparison, in which the red markers correspond to the real measurements and the blue markers to the corresponding calibrated synthetic model. In the original physical feature space (gray region), denoted by f_i, the real and synthetic data overlap partially, indicating a residual mismatch between the measured and simulated baseline responses. After SA (pink region), the projected features p_i exhibit visibly improved agreement, particularly in the marginal distributions along the diagonal of the figure, where the real and synthetic distributions become more closely aligned.

Fig. 9 shows instead a two-dimensional visualization of the harmonized features projected into the latent space for two representative bridges (Bridges 1 and 2), considering anomaly types S1 and U1 at different intensity levels. Despite corresponding to the same physical anomaly scenarios, the features populate distinct regions of the latent space as a result of differences in bridge geometry. This behavior demonstrates that anomaly characterization based solely on latent-space position, such as approaches relying on a k-Nearest Neighbors (KNN) classifier, is not applicable in the present case [37]. Consequently, the regression-based neural network introduced in Section 2.3 is required to enable quantitative anomaly characterization.

In order to train and test the proposed dual-output regressor, the entire synthetic dataset is split into 75,000 samples for training (using Bridges 1, 2, 4, 5, and 6), 15,000 for validation (Bridge 3), and 15,000 for testing (Bridge 8). Bridge 7 will be used later for different analyses.

During the training process, both target variables (anomaly intensity and spatial extent) are normalized to the interval $[0, 1]$ to ensure stable optimization and consistent scaling. The Adam optimizer is used for training, with a mini-batch size of 32, shuffling performed at each epoch, and ℓ_2 regularization set to $\lambda = 10^{-3}$. A piecewise learning-rate schedule is adopted, whereby the learning rate is multiplied by a factor of 0.6 every four epochs. Validation is performed every 78 iterations, and early stopping is activated after 10 consecutive validation checks without improvement. The initial learning rate is set to 10^{-3}, and the loss function is defined as an equally weighted joint regression over anomaly intensity and spatial extent. The selection of network hyperparameters is further discussed in Appendix B. Training terminates after epoch 13 (2418 iterations out of a maximum of 29,250), with a final learning rate of 2.16×10^{-4}.

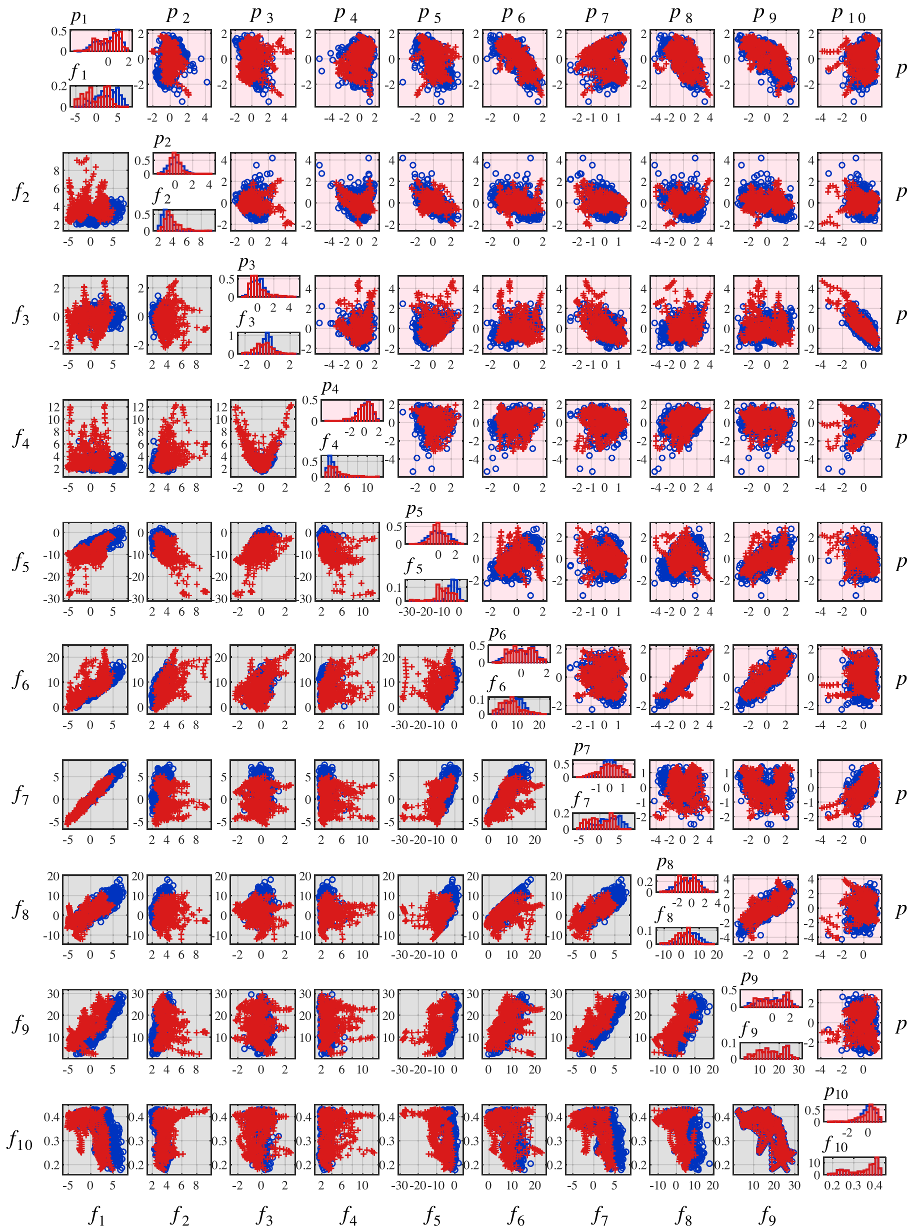

Fig. 8. Pairwise comparison of the real and synthetic baseline features for Bridge 1, before alignment (original features f_i, gray region) and after SA (projected features p_i, pink region). Blue circles denote the synthetic data from the calibrated model and red crosses the real measurements. The diagonal panels show the corresponding marginal distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

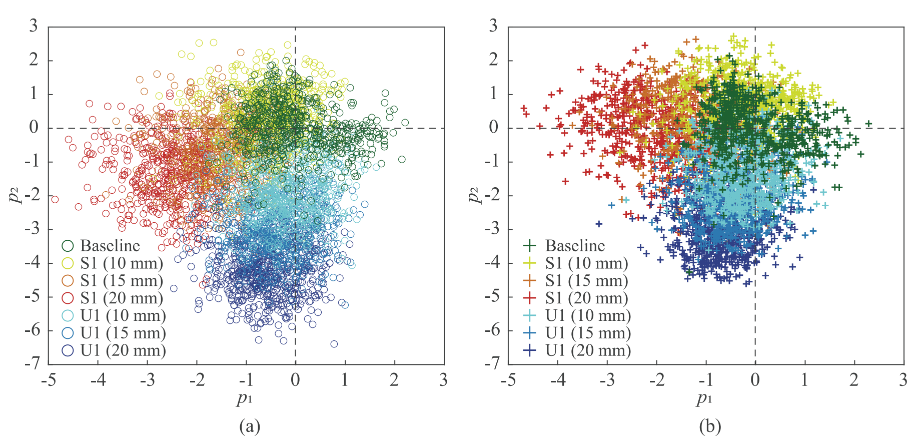

Fig. 9. Latent space distributions ($p_1 - p_2$ plane) of anomaly types S1 and U1 at different intensities for (a) Bridge 1 and (b) Bridge 2.

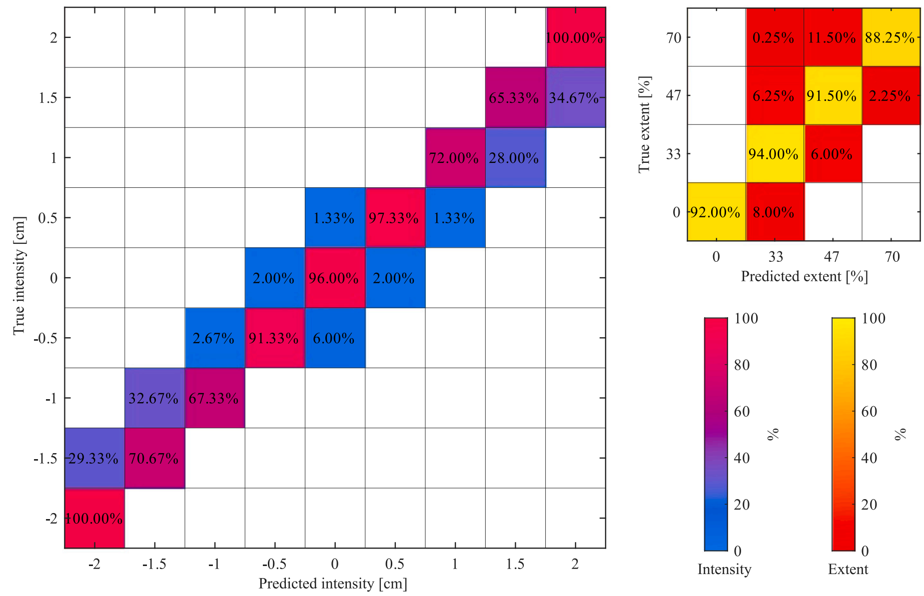

Fig. 10. Confusion matrix for anomaly intensity and spatial extent.

Table 6
Classification-style performance for intensity and extent predictions (averaged across all classes).

Output	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Intensity	83.52	84.92	84.44	83.71
Extent	91.28	93.52	91.44	92.34

To evaluate the regression performance with easily interpretable metrics, Table 6 reports accuracy, precision, recall, and F1 score for both outputs, evaluated on the test bridge (Bridge 8). These quantities are computed per class and then averaged. The network predictions are first mapped back from the normalized interval $[0,1]$ to the original target space ($[-20,20]$ mm for intensity and $[0,100]$% for the extent), and then discretized onto the nearest admissible labels, namely $[-20,-15,-10,-5,0,5,10,15,20]$ mm for the intensity and $[0,33,47,70]$% for the extent output, where 0 denotes no affected PSs and 33, 47, and 70% correspond to the original fractions of affected PSs associated with anomaly in one, two, and three piers for the test bridge, respectively.

Although such metrics are typically defined for the classification tasks, this mapping enhances the interpretability of the results, as regression-oriented metrics such as RMSE would depend heavily on the normalization of targets to the $[0,1]$ interval and would therefore provide less intuitive insight into anomaly characterization performance.

Fig. 10 presents the corresponding confusion matrices, providing a detailed class-by-class evaluation of the predictive precision of the model.

To further assess the generalization capability of the proposed model, an additional independent test is conducted using previously unseen data from Bridge 7, which was excluded from all training, validation, and initial testing stages. This test set comprises incremental settlement and uplift scenarios involving adjacent bridge piers that differ from those considered in the initial synthetic dataset. Specifically, six anomaly cases are simulated, each initiating on December 10, 2020. In all scenarios, the anomaly intensity increases linearly over time, reaching a maximum value of 20 mm on September 11, 2023. Notably, the synthetic datasets are generated up to this date because, once the parametric model is calibrated, only environmental drivers are required to simulate the structural response.

Fig. 11 presents the predicted anomaly intensity (first column) and spatial extent (second column). Across all scenarios, the anomaly intensity is estimated with good accuracy over the full range of simulated values, although variability makes intensities below approximately 5 mm more difficult to detect reliably. In contrast, predictions of the affected spatial extent exhibit higher variability at low anomaly intensities and become progressively more stable as the intensity increases, with a marked stabilization observed for intensities of approximately 5 mm and above.

These results suggest a practical operational strategy: anomaly detection should be based on thresholding the estimated intensity (e.g., around 5 mm in this case, according to the prediction stability), while the predicted extent should be interpreted only for instances in which the intensity exceeds this threshold, that is, when the extent estimate becomes informative.

It is worth emphasizing that, although feature harmonization projects the raw measurements into a latent space where direct physical interpretation is no longer explicit and the effects of anomalies are not directly comparable across structures, the proposed neural network produces outputs that remain directly interpretable in engineering terms and applicable across the entire bridge population. Appendix A.3 presents the same analyses performed without applying SA to the features. The reduced performance observed in that case indicates that, even when physically meaningful features are used, inter-bridge discrepancies hinder the learning process and make generalization more challenging in the absence of feature harmonization.

5.2. Population of real bridges

This section presents the results obtained by applying the proposed methodology to the real data collected from the monitored bridges described in Section 3. For each bridge, features are first extracted from the satellite displacement time series and environmental variables following the procedure outlined in Section 2.1.

To enable simulation-to-real knowledge transfer, all real-bridge features are projected onto the latent space defined in Section 5.1 by the synthetic training data. Specifically, the real features obtained

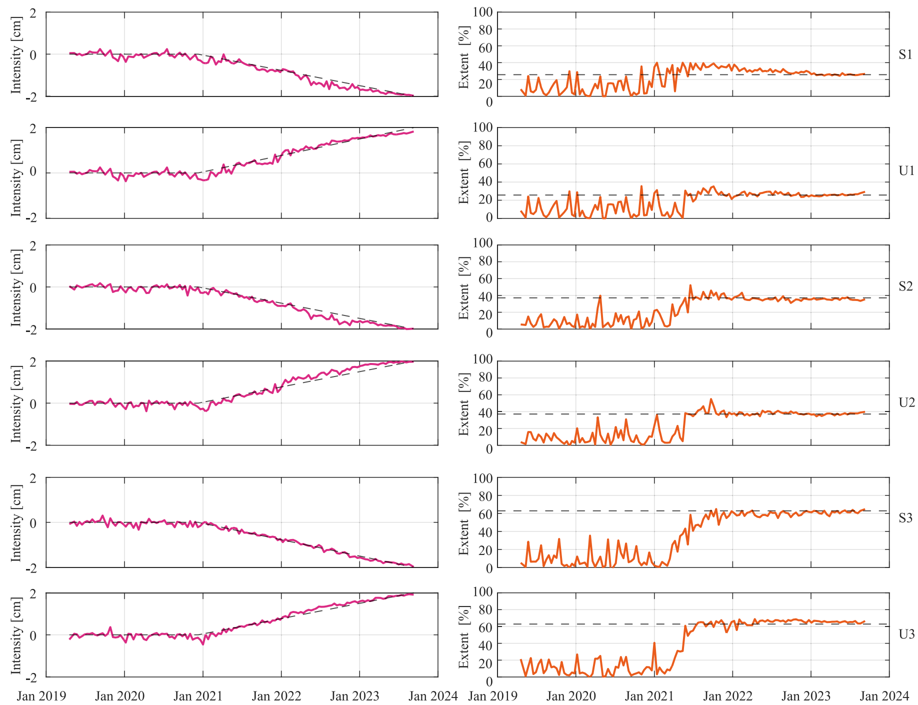

Fig. 11. Predicted anomaly intensity (left column) and spatial extent (right column) for Bridge 7 obtained from the regression model for scenarios affecting one, two, and three adjacent piers, respectively, equivalent to 25.71%, 37.14%, and 62.86% of the available PSs. Dashed lines indicate the reference values.

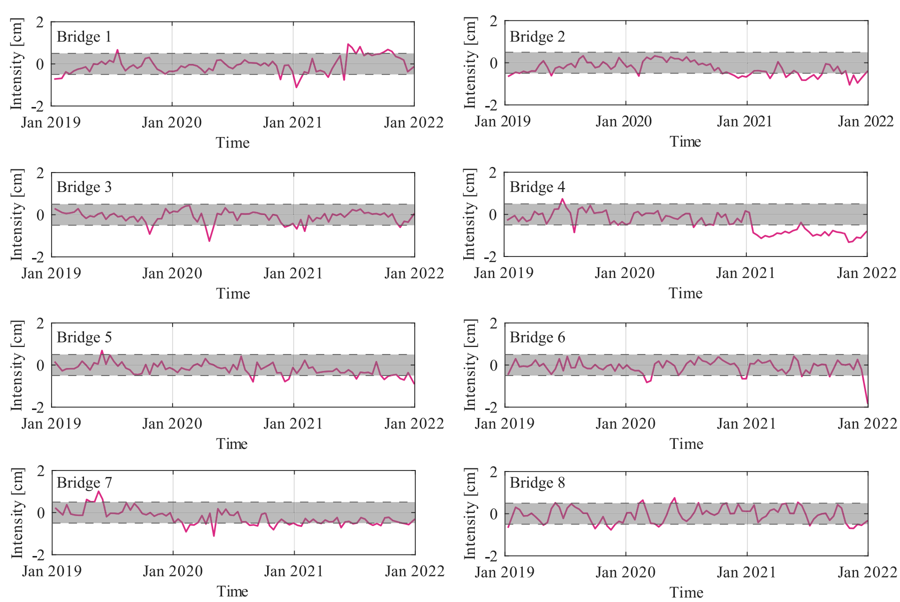

Fig. 12. Predicted anomaly intensity obtained by the regression neural network for the real monitored bridges over the period 2019–2021.

between January 10, 2019, and August 31, 2020 (600 daily instances) are used to estimate the projection matrices $\mathbf{W}_s$ (see Eq. (2)), under the assumption that all bridges are in a healthy condition during this baseline interval. On the other hand, the source domain is defined by the synthetic dataset of Bridge 1, which is used to compute the reference subspace $\mathbf{V}_1$ over the same temporal window. Projecting the complete set of real-bridge features onto this synthetic-derived latent space enables the direct application of the regression model trained on synthetic data.

The resulting predictions are summarized in Fig. 12, which reports the estimated anomaly intensity for each monitored bridge over the entire observation period from 2019 to the end of 2021.

A single sensitivity threshold of 5 mm is adopted for all bridges and is indicated by the gray shaded region in Fig. 12. This value corresponds to the minimum anomaly intensity that can be meaningfully identified in the synthetic dataset and adequately captures the range of variability of the anomaly index during the baseline interval. When this threshold is applied, only Bridge 4 exhibits persistent settlement intensity levels

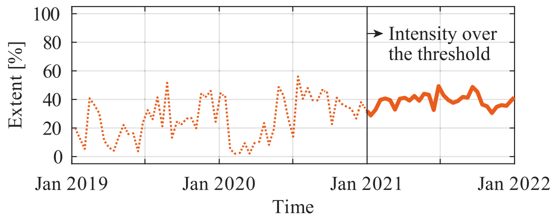

Fig. 13. Predicted spatial extent obtained by the regression neural network for Bridge 4 over the period 2019–2021.

exceeding the threshold, with estimated magnitudes on the order of 7–10 mm starting in early 2021.

Following the approach described in Section 5.1, the predicted spatial extent for Bridge 4 is examined and reported in Fig. 13. Consistent with the synthetic results, the predicted extent stabilizes once the anomaly intensity exceeds the sensitivity threshold, reaching values of approximately 40%. By design, the proposed methodology relies on bridge-level statistical features rather than on the position of individual PSs, owing to the limited geolocation accuracy of medium-resolution InSAR measurements. Consequently, the detected anomaly cannot be attributed to a specific structural component.

Although no official sources report confirmed settlement mechanisms for the investigated structures, the surrounding area is known to be highly active in terms of landslide processes [40]. A detailed inspection of the EGMS velocity maps for Bridge 4, derived from descending-orbit acquisitions of tracks 66 and 168, indicates that a cluster of PSs located on the southern side of the structure exhibits higher negative velocities, between −2 and −5 mm/year, Fig. 14. Consistently, the mean displacement time series computed for this subset of 20 southern scatterers shows a pronounced displacement on the order of 5–10 mm occurring at the beginning of 2021, particularly for measurements associated with track 168.

While the most evident cluster of high-velocity scatterers is located on the southern portion of the bridge, additional scatterers exhibiting moderate negative velocities (on the order of −3 mm/year) are also observed in the northern part of the structure. The combined contribution of these scatterers likely explains the slightly larger extent predicted by the model.

It is important to emphasize that the presence, origin, and structural implications of the detected anomaly cannot be conclusively assessed based solely on satellite observations. Nevertheless, the results demonstrate that the proposed methodology is able to automatically detect and characterize anomalous deformation patterns directly from the processed data, without requiring manual inspection of velocity maps for individual bridges or detailed analysis of displacement time histories, tasks that typically demand the processing and interpretation of several years of InSAR data on a case-by-case basis. In this sense, the proposed framework provides an effective large-scale screening tool to automatically flag potentially critical structures within a monitored population. Any subsequent interpretation of the detected anomalies, or assessment of their structural relevance, should be supported by targeted visual inspections or complementary monitoring techniques, such as on-site instrumentation or higher-resolution InSAR products.

6. Conclusions

This study demonstrates the effectiveness of integrating satellite-based interferometric measurements with a population-based structural health monitoring framework for the detection and quantitative characterization of anomalous deformation patterns in masonry bridges, interpreted as indicators of potentially developing mechanisms. The proposed methodology combines domain adaptation through subspace alignment with a bidirectional long short-term memory regression model trained exclusively on synthetic data, enabling simulation-to-real knowledge transfer to real InSAR displacement time series.

Results obtained on a population of eight masonry arch bridges in Liguria show that the simplified parametric bridge models are able to reproduce the dominant environmentally driven vertical displacement behavior with an average root mean square error of approximately 2 mm. These models provide realistic synthetic training data that encapsulate the main correlations between environmental drivers and structural response without requiring detailed material properties or high-fidelity numerical modeling.

When applied to both synthetic and real datasets, the proposed regression framework successfully estimates two physically interpretable anomaly descriptors: intensity (average settlement or uplift magnitude) and spatial extent (percentage of affected persistent scatterers).

In the real monitoring application, the method consistently identifies a persistent settlement pattern affecting one bridge in the population, with predicted intensity exceeding the sensitivity threshold (5 mm) from early 2021 and stabilizing at values on the order of 7–10 mm. The estimated spatial extent converges to approximately 40% of the available persistent scatterers once the intensity threshold is exceeded. Independent inspection of EGMS velocity maps and displacement time histories confirms the presence of coherent negative displacement trends in subsets of scatterers, supporting the consistency of the detected anomaly across different data releases.

A key outcome of this work is the demonstration that meaningful population-level inference does not require geometric invariance of

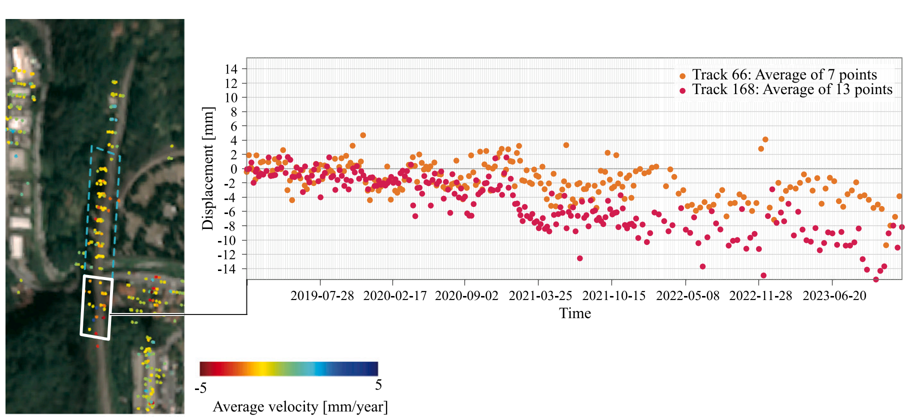

Fig. 14. Velocity map and average displacement time history of the southern portion of Bridge 4 obtained from EGMS.

anomaly patterns in the latent space. Instead, comparability is enforced at the level of baseline behavior through feature harmonization, while anomaly characterization is learned statistically by the regression model. This design enables quantitative inference across structurally heterogeneous bridges, overcoming the limitations of classification strategies based solely on latent-space proximity and avoiding the need for long historical records or detailed structural models.

Overall, the proposed framework provides a scalable, non-invasive, and data-efficient tool for regional-scale screening and prioritization of bridge assets using satellite observations. Importantly, the method is not intended to replace detailed structural assessment or to conclusively identify damage mechanisms, but rather to automatically highlight anomalous deformation patterns present in the data and support informed decision-making for further inspection.

Future developments will focus on extending the framework to additional classes of structures and anomaly types, as well as incorporating ascending and descending InSAR geometries to improve directional sensitivity. In particular, 3D displacement reconstruction is expected to become increasingly feasible with the ongoing evolution of satellite InSAR, through the combined use of ascending and descending acquisitions and the advent of emerging constellations such as IRIDE. These missions will provide higher spatial resolution and additional non-polar viewing geometries, enabling improved recovery of the north–south displacement component, to which near-polar SAR missions are largely insensitive.

CRediT authorship contribution statement

Wael Alahmad: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Said Quqa:** Writing – original draft, Validation, Methodology, Investigation, Conceptualization. **Cristina Gentilini:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Fig. 3(b–d) is reproduced under the CC BY-SA 4.0 license. The original images are available on Wikimedia Commons:

- Fig. 3(b): credited to Giorgio Stagni (October 2, 2011), https://commons.wikimedia.org/wiki/File:Genova_Granara_viadotto_ferroviario.jpg.
- Fig. 3(c): credited to Arbalete (December 28, 2023), https://commons.wikimedia.org/wiki/File:Ferrovia_succursale_dei_Giovi_-_viadotto_Romairone.jpg.
- Fig. 3(d): credited to Arbalete (December 28, 2023), https://commons.wikimedia.org/wiki/File:Ferrovia_succursale_dei_Giovi_-_viadotto_Costagiutta.jpg.

Appendix A. Comparison of regression models and domain-adaptation strategies

To support the architectural and methodological choices adopted in this study, the proposed framework is compared against alternative components under identical conditions. All alternatives share the same overall processing pipeline. Only the recurrent backbone Appendix A.1 or the feature-harmonization step Appendix A.2 is changed. The alternative models are not individually optimized, so the results are intended as a component-level comparison rather than an exhaustive benchmark. Performance is reported with the classification-style metrics defined in Section 5 and evaluated on the test bridge.

Table A.7
Component-level comparison of temporal backbones (test bridge). The best value in each column is highlighted in bold.

Model	Intensity		Extent	
	Acc. (%)	F1 (%)	Acc. (%)	F1 (%)
BiLSTM (used)	83.52	83.71	**91.28**	**92.34**
LSTM	**89.68**	**88.48**	90.08	88.60
GRU	77.92	78.15	88.48	85.63
ANN	66.16	64.44	72.82	55.79

Table A.8
Comparison of domain-adaptation strategies (test bridge). The *None* row corresponds to the pipeline with feature harmonization removed, analyzed in detail in Appendix A.3. The best value in each column is highlighted in bold.

Strategy	Intensity		Extent	
	Acc. (%)	F1 (%)	Acc. (%)	F1 (%)
SA (used)	83.52	83.71	**91.28**	**92.34**
None	61.92	58.62	79.76	61.25
NCA	79.92	77.34	76.72	58.92
NCORAL	**90.08**	**88.93**	88.08	90.56

A.1. Recurrent backbone

Table A.7 reports the accuracy and F1 score obtained using the same architecture shown in Fig. 2, replacing only the recurrent backbone with the alternative described here. In the recurrent variants (LSTM and gated recurrent units — GRU), the three BiLSTM layers are replaced by three unidirectional LSTM (or GRU) layers with the same number of units (128) per layer, retaining the same arrangement: the first two in sequence-to-sequence mode, each followed by layer normalization and dropout, and the third in sequence-to-vector mode. On the other hand, a fully-connected artificial neural network (ANN) is selected as a non-temporal baseline. It processes the instantaneous feature vector of a single day through three fully connected layers of 128 units with ReLU activations and dropout. In all cases, the two fully connected regression heads and the optimization settings are identical to those of the proposed model.

The ANN performs markedly worse on both outputs, confirming that the temporal structure of the displacement features is essential. Among the recurrent alternatives, the unidirectional LSTM attains the best intensity scores, whereas the proposed BiLSTM attains the best spatial-extent scores, which is the more challenging output.

A.2. Domain-adaptation strategy

The feature harmonization step is then compared with alternative domain adaptation strategies, namely, normal condition alignment (NCA) and normal correlation alignment (NCORAL) [41], and with the case in which no adaptation is applied. For each strategy, only the harmonization step is replaced by the corresponding method, while the remaining stages of the pipeline are left unchanged. Table A.8 presents the performance results obtained using the cases described above.

All domain-adaptation strategies improve over the no-adaptation case. SA and NCORAL are the strongest options and follow the same trade-off observed for the backbones: NCORAL is slightly superior for the intensity output, whereas SA achieves the best spatial-extent estimation.

A.3. Effect of feature harmonization

This section evaluates the impact of the feature harmonization step by repeating the methodology described in Section 5.1 with SA removed from the processing pipeline. In this configuration, the BiLSTM regressor is trained using the same network architecture and

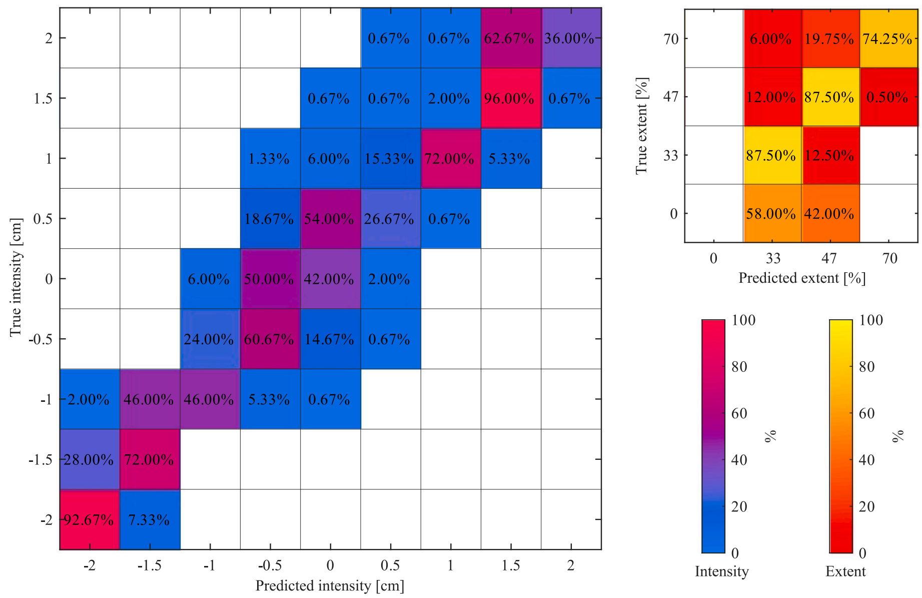

Fig. A.15. Confusion matrix for anomaly intensity and spatial extent without feature harmonization.

Table A.9
Classification-style performance for intensity and extent predictions without feature harmonization (averaged across all classes).

Output	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Intensity	61.92	64.84	60.44	58.62
Extent	79.76	61.73	62.31	61.25

hyperparameters selected in Appendix B, as well as the same dataset partitioning strategy.

Model evaluation follows the same classification-style aggregation adopted for Table 6, whereby the continuous regression outputs are discretized into intensity and extent classes to compute accuracy, average precision, average recall, and average F1 score. Table A.9 reports the full set of performance metrics obtained without feature harmonization.

Removing SA results in a marked degradation of predictive performance. In particular, accuracy for anomaly intensity decreases from 83.52% to 61.92%, while accuracy for spatial extent drops from 91.28% to 79.76%. Similar trends are observed for the remaining metrics, with average precision, recall, and F1 score consistently lower in the absence of feature harmonization. The most pronounced degradation is observed for the extent F1 score, which decreases from 92.34% to 61.25%.

These results highlight the critical role of SA in harmonizing feature distributions across domains, thereby improving the consistency of the regression mapping and reducing misclassification. This behavior is further illustrated by the confusion matrices shown in Fig. A.15, where the absence of feature harmonization leads to increased confusion among neighboring classes, particularly for intermediate anomaly levels.

Finally, the regression outputs are examined for the same synthetic bridge scenarios presented in Fig. 11. Fig. A.16 reports the corresponding predictions obtained without applying SA. Compared with the harmonized case discussed in Section 5.1, the predictions exhibit reduced contrast between different anomaly severities and a less stable match between imposed and predicted spatial extent. These observations are consistent with the quantitative performance degradation reported in Table A.9.

Appendix B. Hyperparameter search and architecture selection

A grid search is performed over a family of bidirectional LSTM architectures based on the structure shown in Fig. 2, which comprises stacked recurrent blocks followed by dual regression heads. The objective is to identify an architecture and set of hyperparameters that are optimal for the application considered in this study. In all configurations, the final recurrent layer is fixed as a BiLSTM operating in "last mode" (sequence-to-vector), while the grid explores variations in the number and size of the preceding sequence-output BiLSTM layers. The hyperparameter space explored included:

- Number of sequence-output BiLSTM blocks: $1, 2, 3$.
- Hidden units per direction in each BiLSTM: $64, 128, 256$.
- Input sequence length L: $6, 12$ days.
- Width of the dense prediction heads: $128, 256, 512$.
- Dropout probability after layer normalization: $0.10, 0.20, 0.30$.
- ℓ_2 regularization coefficient λ: 1×10^{-4}, 3×10^{-4}, 1×10^{-3}.

For each configuration, the dataset is segmented into non-overlapping windows of length L, and the network is trained using identical optimization settings (see Section 5). Model performance is evaluated using the validation mean squared error (MSE), computed jointly over the two regression outputs (anomaly intensity and spatial extent). The same validation split is used across all experiments to ensure a fair comparison.

Model selection is primarily based on the validation MSE, with lower values indicating better performance. When multiple configurations yield comparable results within the reporting precision, preference is given to the simpler architecture, as measured by the total number of learnable parameters. This parsimony criterion favors models that retain predictive accuracy while reducing model complexity, memory requirements, and inference cost.

Data availability

Data will be made available on request.

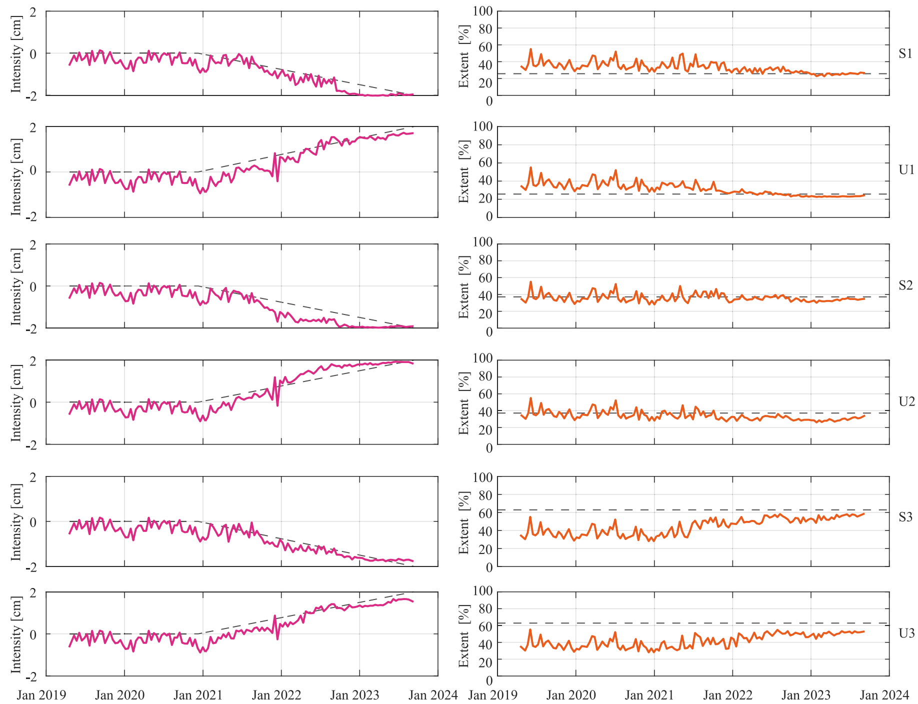

Fig. A.16. Predicted anomaly intensity (left column) and spatial extent (right column) for Bridge 7 obtained from the regression model without feature harmonization for scenarios affecting one, two, and three adjacent piers, respectively. Dashed lines indicate the reference values.

References

[1] Santis SD, de Felice G. Overview of railway masonry bridges with a safety factor estimate. Int J Archit Herit 2014;8:452–74. http://dx.doi.org/10.1080/15583058.2013.826298.

[2] Gentile C, Saisi A. Ambient vibration testing of historic masonry towers for structural identification and damage assessment. Constr Build Mater 2007;21:1311–21. http://dx.doi.org/10.1016/j.conbuildmat.2006.01.007.

[3] Ramos LF, Marques L, Lourenço PB, De Roeck G, Campos-Costa A, Roque J. Monitoring historical masonry structures with operational modal analysis: Two case studies. Mech Syst Signal Process 2010;24:1291–305. http://dx.doi.org/10.1016/j.ymssp.2010.01.011.

[4] Ubertini F, Cavalagli N, Kita A, Comanducci G. Assessment of a monumental masonry bell-tower after 2016 central Italy seismic sequence by long-term SHM. Bull Earthq Eng 2018;16:775–801. http://dx.doi.org/10.1007/s10518-017-0222-7.

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

[6] Gagliardi V, Ciampoli LB, D'Amico F, Alani AM, Tosti F, Benedetto A. Multi-temporal SAR interferometry for structural assessment of bridges: The rochester bridge case study. ASCE Library; 2021, p. 308–19. http://dx.doi.org/10.1061/9780784483510.028.

[7] Modesti M, Gentilini C, Palermo A, Reynders E, Lombaert G. A two-step procedure for damage detection in beam structures with incomplete mode shapes. J Civ Struct Health Monit 2025;15:287–306. http://dx.doi.org/10.1007/s13349-024-00839-0.

[8] Entezami A, Sarmadi H, Behkamal B. Long-term health monitoring of concrete and steel bridges under large and missing data by unsupervised meta learning. Eng Struct 2023;279:115616. http://dx.doi.org/10.1016/j.engstruct.2023.115616.

[9] Feng D, Feng MQ. Vision-based multipoint displacement measurement for structural health monitoring. Struct Control Health Monit 2016;23:876–90.

[10] Sarmadi H, Entezami A, Yuen KV, Behkamal B. Review on smartphone sensing technology for structural health monitoring. Measurement 2023;223:113716.

[11] Ferretti A, Prati C, Rocca F. Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. IEEE Trans Geosci Remote Sens 2000;38:2202–12. http://dx.doi.org/10.1109/36.868878.

[12] Ferretti A, Prati C, Rocca F. Permanent scatterers in SAR interferometry. IEEE Trans Geosci Remote Sens 2001;39:8–20. http://dx.doi.org/10.1109/36.898661.

[13] Ferretti A, Monti-Guarnieri A, Prati C, Rocca F, Massonnet D. InSAR principles: Guidelines for SAR interferometry processing and interpretation. ESA Publications; 2007.

[14] Selvakumaran S, Rossi C, Marinoni A, Webb G, Bennetts J, Barton E, Plank S, Middleton C. Combined InSAR and terrestrial structural monitoring of bridges. IEEE Trans Geosci Remote Sens 2020;58:4933–46. http://dx.doi.org/10.1109/TGRS.2020.2979961.

[15] Giordano PF, Kwapisz M, Miano A, Liuzzo R, Vorwagner A, Limongelli MP, Prota A, Ralbovsky M. Monitoring of a multi-span prestressed concrete bridge using satellite interferometric data and comparison with on-site sensor results. Struct Concr 2025;26:5430–53.

[16] Selvakumaran S, Plank S, Geiß C, Rossi C, Middleton C. Remote monitoring to predict bridge scour failure using Interferometric Synthetic Aperture Radar (InSAR) stacking techniques. Int J Appl Earth Obs Geoinf 2018;73:463–70. http://dx.doi.org/10.1016/j.jag.2018.07.004.

[17] Alani AM, Tosti F, Ciampoli LB, Gagliardi V, Benedetto A. An integrated investigative approach in health monitoring of masonry arch bridges using GPR and InSAR technologies. NDT E Int 2020;115:102288. http://dx.doi.org/10.1016/j.ndteint.2020.102288.

[18] Bonaldo G, Caprino A, Lorenzoni F, da Porto F. Monitoring displacements and damage detection through satellite MT-InSAR techniques: A new methodology and application to a case study in Rome (Italy). Remote Sens 2023;15. http://dx.doi.org/10.3390/rs15051177.

[19] Caprino A, Bonaldo G, Lorenzoni F, da Porto F. Application of multi-temporal InSAR (MT-InSAR) for structural monitoring: the case study of Scrovegni Chapel in Padova. Procedia Struct Integr 2023;44:1578–85. http://dx.doi.org/10.1016/j.prostr.2023.01.202.

[20] Guzman-Acevedo GM, Vazquez-Becerra GE, Quintana-Rodriguez JA, Gaxiola-Camacho JR, Anaya-Diaz M, Mediano-Martinez JC, Viramontes FJC. Structural health monitoring and risk assessment of bridges integrating InSAR and a calibrated FE model. Structures 2024;63:106353. http://dx.doi.org/10.1016/j.istruc.2024.106353.

[21] Farneti E, Cavalagli N, Costantini M, Trillo F, Minati F, Venanzi I, Ubertini F. A method for structural monitoring of multispan bridges using satellite InSAR data with uncertainty quantification and its pre-collapse application to the Albiano-Magra Bridge in Italy. Struct Health Monit 2023a;22:353–71. http://dx.doi.org/10.1177/14759217221083609.

[22] Farneti E, Cavalagli N, Venanzi I, Salvatore W, Ubertini F. Residual service life prediction for bridges undergoing slow landslide-induced movements combining satellite radar interferometry and numerical collapse simulation. Eng Struct 2023b;293:116628. http://dx.doi.org/10.1016/j.engstruct.2023.116628.

[23] Entezami A, Behkamal B, De Michele C, Mariani S. Displacement prediction for long-span bridges via limited remote sensing images: An adaptive ensemble regression method. Measurement 2025a;245:116567. http://dx.doi.org/10.1016/j.measurement.2024.116567.

[24] Entezami A, Behkamal B, De Michele C, Mariani S. A kernelized deep regression method to simultaneously predict and normalize displacement responses of long-span bridges via limited synthetic aperture radar images. Struct Health Monit 2025b;1–22. http://dx.doi.org/10.1177/14759217241309075.

[25] Quqa S, Palermo A, Ubertini F, Marzani A. Regional-scale bridge condition monitoring using InSAR displacements and environmental data. Struct Health Monit 2024;24:14759217241302369. http://dx.doi.org/10.1177/14759217241302369.

[26] Quqa S, Lasri O, Delo G, Giordano PF, Surace C, Marzani A, Limongelli MP. Regional-scale bridge health monitoring: Survey of current methods and roadmap for future opportunities under changing climate. Struct Health Monit 2025;24:14759217241310525. http://dx.doi.org/10.1177/14759217241310525.

[27] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2010;22:1345–59. http://dx.doi.org/10.1109/TKDE.2009.191.

[28] Capes R, Passera E. Product description and format specification. Technical report EGMS-D6-PDD-SC1-2.0-009, EGMS-SC1 ORIGINAL Consortium; 2023, URL: https://land.copernicus.eu/pan-european/european-ground-motion-service, Document Owner: Renalt Capes. Reviewed and approved by Alessandro Ferretti and Lorenzo Solari (EEA).

[29] Ferretti A, Passera E, Capes R. End-to-end implementation and operation of the European ground motion service (EGMS): Algorithm theoretical basis document. Technical report EGMS-D3-ALG-SC1-2.0-006, 2023.

[30] Kotzerke P, Siegmund R, Langenwalter J. End-to-end implementation and operation of the European ground motion service (EGMS): Product user manual. Technical report EGMS-D4-PUM-SC1-2.0-007, 2022.

[31] ESA. Sentinel-1B in-flight anomaly summary report. Technical report, 2022.

[32] Pepe A, Bonano M, Zhao Q, Yang T, Wang H. The use of C-/X-band time-gapped SAR data and geotechnical models for the study of Shanghai's ocean-reclaimed lands through the SBAS-DInSAR technique. Remote Sens 2016;8. http://dx.doi.org/10.3390/rs8110911.

[33] Muñoz Sabater J, Dutra E, Agustí-Panareda A, Albergel C, Arduini G, Balsamo G, Boussetta S, Choulga M, Harrigan S, Hersbach H, Martens B, Miralles DG, Piles M, Rodríguez-Fernández NJ, Zsoter E, Buontempo C, Thépaut JN. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst Sci Data 2021;13:4349–83. http://dx.doi.org/10.5194/essd-13-4349-2021.

[34] Zippenfenig P. Open-meteo.com weather api. 2023, http://dx.doi.org/10.5281/zenodo.7970649, URL: https://open-meteo.com/.

[35] OMIRL. Osservatorio meteo idrologico della regione liguria. 2026, https://omirl.regione.liguria.it/#/map. [Accessed 03 June 2026].

[36] Fernando B, Habrard A, Sebban M, Tuytelaars T. Unsupervised visual domain adaptation using subspace alignment. In: Proceedings of the IEEE international conference on computer vision. ICCV, 2013, p. 2960–7. http://dx.doi.org/10.1109/ICCV.2013.368.

[37] Alahmad W, Quqa S, Ubertini F, Gentilini C. Anomaly identification in historical modular structures using satellite data: Methodology and application to Portico di San Luca, Bologna. J Build Eng 2025;114893. http://dx.doi.org/10.1016/j.jobe.2025.114893.

[38] Alahmad W, Quqa S, Ugolini A, Gentilini C. Simplified parametric structural model for masonry porticoes and bridges. In: Castellazzi G, Gentilini C, editors. Mechanics of masonry structures strengthened with composite materials. Cham: Springer Nature Switzerland; 2025b, p. 475–84.

[39] Google LLC. Google earth pro (version 7.3.6). 2024, https://www.google.com/earth/versions/#earth-pro. [Accessed 14 April 2025].

[40] ISPRA. Italian landslide inventory (iffi project). 2024, https://idrogeo.isprambiente.it/app/page/open-data. [Accessed 06 February 2026].

[41] Poole J, Gardner P, Dervilis N, Bull L, Worden K. On statistic alignment for domain adaptation in structural health monitoring. Struct Health Monit 2023;22:1581–600. http://dx.doi.org/10.1177/14759217221110441.