**Cold Climate Effects on Bridge Dynamics: A Hybrid Unsupervised Learning Approach for Robust Structural Health Monitoring**

Soheila Goli, Alireza Entezami, Alberto Corigliano

Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, Italy

soheila.goli@mail.polimi.it

**Abstract.** Temperature fluctuations in cold climates can significantly distort the dynamic characteristics of bridges, leading to misinterpretations in structural health monitoring (SHM) systems. This study proposes a novel unsupervised learning framework designed to robustly mitigate the influence of freezing temperatures on bridge eigenfrequencies, thereby enhancing the reliability of SHM. The approach integrates Gaussian Mixture Modeling (GMM) to probabilistically cluster structural response data into local environmental regimes, combined with Local Principal Component Analysis (LPCA) to reconstruct and normalize eigenfrequency variations within each cluster. By addressing the nonlinear effects of cold climates, the framework effectively isolates and removes environmental biases from modal properties without relying on external temperature measurements. The method is validated using long-term monitoring data from the Z24 Bridge in Switzerland, which was subjected to severe freezing conditions. Results demonstrate that the proposed hybrid learning approach successfully eliminates abrupt frequency shifts during freezing periods and maintains consistent dynamic behavior across varying operational scenarios. The integration of probabilistic clustering and localized dimensionality reduction significantly enhances the robustness of SHM systems under environmental variability. This research represents an important advancement for SHM in cold regions, offering a practical, scalable, and data-driven solution for more accurate structural condition assessment under extreme weather effects.

**Keywords:** Structural Health Monitoring (SHM), Bridge Dynamics, Cold Climate, Modal Frequency, Unsupervised Learning

1. Introduction

The practice of health monitoring of bridges through advanced sensing technologies has become indispensable for modern infrastructure management, serving as both an early warning system and a decision-support tool for maintenance planning. The core of structural health monitoring (SHM) systems lies the integration of non-destructive evaluation methods with dense sensor arrays that capture structural responses across varying operational conditions [1]. Where traditional visual inspections fall short, these automated monitoring solutions offer continuous, data-rich insights into structural behavior through two analytical paradigms. Physics-based approaches construct detailed numerical models rooted in mechanical theory and material properties, demanding significant expertise in finite element analysis and model calibration. Moreover, purely data-driven strategies bypass complex modeling requirements by applying machine learning algorithms directly to sensor measurements, extracting damage-sensitive features through statistical pattern recognition. The latter approach has gained particular traction due to its inherent adaptability to unique structural configurations and ability to process streaming sensor data in real-time. As bridge networks age and environmental pressures intensify, the evolution of these monitoring methodologies continues to push the boundaries of predictive maintenance and risk mitigation in civil infrastructure systems.

The application of vibration-based SHM to bridges in cold climate regions faces significant challenges due to pronounced fluctuations in eigenfrequency measurements. These variations stem primarily from temperature-induced stiffening of structural components during freezing conditions, which systematically elevates overall bridge stiffness and consequently modifies modal parameters [2]. Such environmentally-driven frequency shifts create substantial complications for conventional anomaly detection algorithms. When processing these temperature-distorted dynamic characteristics, the algorithms frequently generate false-positive indications of structural abnormalities [3]. This leads to two critical failure modes: erroneous damage alerts in intact structures and the potential masking of genuine structural anomalies [4], both of which compromise the reliability of monitoring systems in cold weather conditions.

In SHM systems, data normalization serves as a fundamental approach to mitigate the influence of environmental and operational variations on measured structural responses. Two distinct methodological frameworks exist for implementing this normalization, differentiated by their data requirements and modeling approaches. The supervised paradigm utilizes measured environmental/operational parameters as input variables to predict structural responses, with the resulting residuals between observed and predicted values providing normalized features. Conversely, the unsupervised framework operates directly on structural response measurements, employing reconstruction techniques to generate residuals that isolate environmental effects. These approaches are respectively implemented through supervised regression models [5-7] and unsupervised reconstruction algorithms, each offering distinct advantages depending on data availability and monitoring objectives [8-10].

Even though conventional data normalization methods have demonstrated effectiveness in addressing environmental and operational variability in modal frequencies [11], further research is required to specifically counteract the pronounced effects of cold climates on dynamic structural responses. Advanced machine learning techniques, particularly statistical and hybrid unsupervised learning approaches, offer promising solutions to this challenge. Statistical unsupervised learning enables the extraction of inherent data patterns without relying on pre-labeled outcomes, making it particularly valuable for analyzing unclassified monitoring data and revealing their intrinsic characteristics data [12, 13]. Hybrid unsupervised learning builds upon this foundation by strategically combining multiple unsupervised techniques, thereby enhancing model accuracy and robustness beyond what single-method approaches can achieve [9, 10]. These advanced methodologies present significant potential for developing more resilient normalization techniques capable of handling the unique challenges posed by extreme cold weather conditions in SHM applications [11-13].

1. Proposed Unsupervised Learning Approach
   1. Local Data Preparation

GMM is a probabilistic technique for data clustering that assumes all data points are generated from a mixture of several Gaussian distributions with unknown parameters. GMM is highly suitable for clustering applications because it provides a methodological way to describe complex data distributions through a combination of simpler Gaussian processes, each characterized by a mean vector and a covariance matrix based on unlabeled training data [17].

Given a dataset where each is a -dimensional feature vector and is the number of training instances, a GMM aims to model this dataset as a mixture of different Gaussian distributions. The probability density function of under a GMM can be expressed as:

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where is the mixing coefficient for the -th component, with and ; is the probability density function of the Gaussian distribution for the -th component with mean vector and covariance matrix ; and denotes the parameters of the GMM. The parameters of the GMM are typically estimated using the Expectation-Maximization (EM) algorithm, which iteratively optimizes the likelihood function. The EM algorithm consists of the expectation step, which calculates the expected value of the latent variables given the current parameter estimates and maximization step, which updates the GMM parameters to maximize the expected log-likelihood obtained in the previous step.

Although the GMM parameters are estimated within its algorithm, the number of components () is an important hyperparameter of this technique. Accordingly, it is necessary to use supplementary approaches to determine this hyperparameter before data clustering. Akaike information criterion (AIC) is a well-known approach:

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

where is the likelihood of the model given the best-fit parameters. Using some sample components, the sample with the minimum AIC value is the optimal choice for . The likelihood function is derived as:

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

Once the GMM parameters are estimated, it separates the global training data into clusters.

* 1. Unsupervised Data Normalization

PCA is an unsupervised technique for dimensionality reduction, feature extraction, and data normalization. One of the noticeable characteristics of PCA relates to its data reconstruction [15]. In this context, it is able to receive the original data and reconstruct it. For the problem of data normalization, such a reconstruction characteristic is valuable that makes PCA a popular approach to removing the environmental and operational variations. However, when the severity of variability is substantial and nonlinear, PCA fails in properly normalizing response data [18-20]. To circumvent these disadvantages, the proposed integrated method leverages the local data provided by the GMM clustering to derive LPCA models for data normalization.

Given the original training data into clusters, one initially needs to estimate the covariance matrix of each cluster . Each covariance matrix is decomposed by the eigenvalue decomposition technique to derive the eigenvectors and eigenvalues:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

where contains the eigenvectors of the -th cluster, , and is a diagonal matrix with the eigenvalues for the -th cluster. In the following, the top eigenvectors are extracted:

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

The projection of clustered data based on the principles of PCA is performed to obtain:

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

where are the scores of the principal components for the -th cluster. Eventually, the residual data of each cluster is computed as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

where represents the residual matrix of the -th cluster. Finally, it is necessary rearrange the residual subsets to align with the original order of samples in the training matrix and obtain the final residual matrix of the training data . This process ensures that each residual sample is correctly mapped back to its original position.

To determine the residual data for a test sample using the original training data and the clusters obtained through LPCA, one needs to follow a process that includes assigning the test sample to the appropriate cluster, projecting it into the subspace defined by that cluster principal components, and then calculating the residual data. Initially, one should assign the test sample to the most appropriate cluster. This strategy can be performed using the nearest centroid approach as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

where denotes the mean vector of cluster and is the index of the cluster closest to the test sample **.** After assigning this sample to its cluster , it should be projected onto the subspace spanned by the principal components of the cluster of interest:

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

where is the matrix containing the principal components (eigenvectors) of . Accordingly, the reconstruction version of the test sample is given by:

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

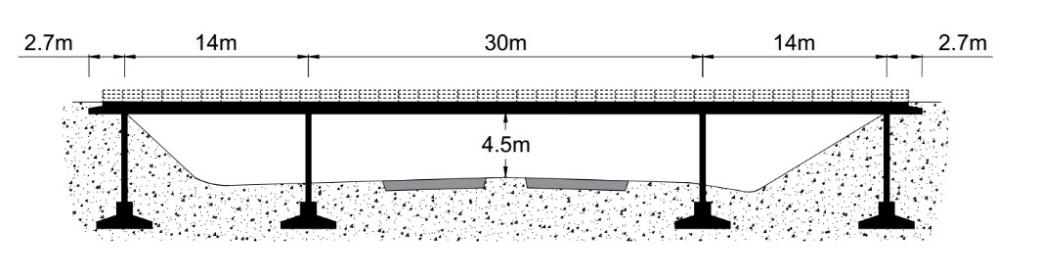
Finally, the residual between the original and reconstructed test sample is simply obtained as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (11) |

It should be explained that because only one test point is considered during the testing (inspection) phase, one can only determine its residual vector. Therefore, the collection of more than one test sample makes the residual matrix of the test data .

1. Case Study

The case study for validation of the proposed method is the Z24 bridge, which was built in 1963 and demolished in 1998 for constructing a new bridge structure with a larger side span [21]. **Fig. 1** shows an elevation view of this bridge along with the main dimensions of the three spans. The Z24 bridge was located in in Switzerland. During 1997-1998, a continuous monitoring program was performed to measure structural responses (acceleration time histories) and environmental factors such as air temperature, humidity, rainfall, and wind speed and direction [21]. An operational modal analysis was incorporated into acceleration responses to identify modal data, particularly the bridge eigenfrequencies as shown in **Fig. 2**.



#### **Fig. 1.** An elevation view of the well-known Z24 bridge

|  |
| --- |
|  |
|  |

#### **Fig. 2.** The eigenfrequencies of the Z24 bridge related to the first four modes (a)-(d)

The analysis of structural eigenfrequencies in the Z24 bridge reveals significant variability induced by freezing weather conditions, as evidenced in **Fig. 2**. During periods of sub-zero temperatures (samples 400-1600), pronounced fluctuations occur in the undamaged frequencies, despite the absence of structural damage. As demonstrated by Entezami et al. [3], these temperature-driven variations can substantially compromise monitoring accuracy, potentially leading to false structural assessments. Such findings underscore the critical need for advanced machine learning approaches capable of robustly isolating and eliminating freezing effects from dynamic response data to ensure reliable structural health evaluation.

Based on the proposed method, the bridge eigenfrequencies are divided into training and test datasets with the ratio of 80%-20%. Accordingly, the training data includes 2500 frequency samples, while the remaining features serve as the test points. Having considered the training data, the proposed method begins by determining the optimal number of components using the AIC function. **Fig. 3** indicates the evolution of the AIC values under 19 sample components varying from 2 to 20. As can be observed, the optimal is identical to 9. Therefore, the training data is grouped into 9 clusters .

**Fig. 4** demonstrates the effectiveness of the proposed normalization method in mitigating cold climate effects on the Z24 bridge eigenfrequencies. The processed results show successful elimination of the previously observed abrupt frequency variations in the sub-zero temperature range (samples 400-1600), with the normalized frequencies exhibiting stable behavior across this critical interval. Furthermore, the method maintains consistent performance across different operational conditions, as evidenced by the uniform variability range observed in test points (samples 2501-3127). These results collectively validate the robustness and reliability of the integrated approach in compensating for freezing weather effects on structural dynamic characteristics. The normalized output demonstrates significant improvement over raw frequency measurements, confirming the method's practical utility for structural health monitoring in cold climate regions.



#### **Fig. 3.** Determination of the optimal component of the GMM model for clustering the training data

|  |
| --- |
|  |
|  |

#### **Fig. 4.** The normalized eigenfrequencies of the Z24 bridge related to the first four modes (a)-(d) by using the proposed method

1. Conclusions

This study proposes a novel statistical hybrid learning approach integrating data clustering and unsupervised normalization to address cold climate effects in structural health monitoring. The methodology comprises two sequential phases: (1) GMM-based clustering of global training data into localized subsets capturing distinct environmental regimes, followed by (2) LPCA applied independently to each cluster for residual extraction. The LPCA phase computes residual matrices as differences between original observations and their cluster-specific reconstructions, effectively isolating temperature-induced variations. Validation using the benchmark Z24 bridge dataset demonstrates that this combined clustering-reconstruction framework successfully mitigates cold climate influences on modal frequencies. The results establish that the synergistic integration of probabilistic clustering (i.e., GMM) with localized dimensionality reduction (i.e., LPCA) provides an effective solution for environmental variability compensation in operational modal analysis.

References

1. Sarmadi, H., Entezami, A., Yuen, K.‑V., Behkamal, B.: Review on smartphone sensing technology for structural health monitoring. Measurement 223, 113716 (2023).
2. Kita, A., Cavalagli, N., Ubertini, F.: Temperature effects on static and dynamic behavior of Consoli Palace in Gubbio, Italy. Mech. Syst. Sig. Process. 120, 180–202 (2019).
3. Entezami, A., Sarmadi, H., Behkamal, B.: Removal of freezing effects from modal frequencies of civil structures for structural health monitoring. Engineering Structures 319, 118722 (2024).
4. Sarmadi, H., Yuen, K.‑V.: Structural health monitoring by a novel probabilistic machine‑learning method based on extreme value theory and mixture quantile modeling. Mech. Syst. Sig. Process. 173, 109049 (2022).
5. Behkamal, B., Entezami, A., De Michele, C., Arslan, A.N.: Elimination of thermal effects from limited structural displacements based on remote sensing by machine‑learning techniques. Remote Sensing 15, 3095 (2023).
6. García‑Macías, E., Ubertini, F.: Least Angle Regression for early‑stage identification of earthquake‑induced damage in a monumental masonry palace: Palazzo dei Consoli. Engineering Structures 259, 114119 (2022).
7. Roberts, C., Cava, D.G., Avendaño‑Valencia, L.D.: Addressing practicalities in multivariate nonlinear regression for mitigating environmental and operational variations. Struct. Health Monit. 22, 1237–1255 (2022).
8. Sarmadi, H., Entezami, A., Magalhães, F.: Unsupervised data normalization for continuous dynamic monitoring by an innovative hybrid feature weighting‑selection algorithm and natural nearest‑neighbor searching. Struct. Health Monit. 22, 4005–4026 (2023).
9. Daneshvar, M.H., Sarmadi, H., Yuen, K.‑V.: A locally unsupervised hybrid learning method for removing environmental effects under different measurement periods. Measurement 208, 112465 (2023).
10. Entezami, A., Sarmadi, H., Behkamal, B.: A novel double‑hybrid learning method for modal‑frequency‑based damage assessment of bridge structures under different environmental variation patterns. Mech. Syst. Sig. Process. 201, 110676 (2023).
11. Entezami, A., Shariatmadar, H., Karamodin, A.: Improving feature extraction via time series modeling for structural health monitoring based on unsupervised learning methods. Scientia Iranica 27(3): 1001–1018 (2020).
12. Rezaiee-Pajand, M., Entezami, A., Shariatmadar, H.: An iterative order determination method for time-series modeling in structural health monitoring. Advances in Structural Engineering 21(2): 300–314 (2018).
13. Entezami, A.: Structural health monitoring by time series analysis and statistical distance measures. Springer, Berlin/Heidelberg, Germany (2021).
14. Wang, Z., Yang, D.‑H., Yi, T.‑H., Zhang, G.‑H., Han, J.‑G.: Eliminating environmental and operational effects on structural modal frequency: A comprehensive review. Struct. Contr. Health Monit. 29, e3073 (2022).
15. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edn. Springer, New York (2013).
16. Eltouny, K., Gomaa, M., Liang, X.: Unsupervised learning methods for data‑driven vibration‑based structural health monitoring: A review. Sensors 23 (2023).
17. Aggarwal, C.C., Reddy, C.K.: Data Clustering: Algorithms and Applications. CRC Press, Boca Raton (2016).
18. Gharibnezhad, F., Mujica, L.E., Rodellar, J.: Applying robust variant of principal component analysis as a damage detector in the presence of outliers. Mech. Syst. Sig. Process. 50–51, 467–479 (2015).
19. Giglioni, V., García‑Macías, E., Venanzi, I., Ierimonti, L., Ubertini, F.: The use of receiver‑operating‑characteristic curves and precision‑versus‑recall curves as performance metrics in unsupervised structural damage classification under changing environment. Engineering Structures 246, 113029 (2021).
20. Sarmadi, H., Entezami, A., Saeedi Razavi, B., Yuen, K.‑V.: Ensemble‑learning‑based structural health monitoring by Mahalanobis distance metrics. Struct. Contr. Health Monit. 28, e2663 (2021).
21. Peeters, B., De Roeck, G.: One‑year monitoring of the Z24‑Bridge: Environmental effects versus damage events. Earthquake Eng. Struct. Dyn. 30, 149–171 (2001).