Integrating Remote Sensing and on-site data to derive historical time series for Long-Term structural monitoring

Integrating Remote Sensing and On-Site Data to Derive Historical Time Series for Long-Term Structural Monitoring

Stefania COCCIMIGLIO [1], Gaetano MIRAGLIA [1,2], Cristian CAPODICASA [1,*], Alessandro MARZANI [3], Rosario CERAVOLO [1,2]
[1] Politecnico di Torino, Italy
[2] R3C – Responsible Risk Resilience Centre, Politecnico di Torino
[3] Alma Mater Studiorum - Università di Bologna

*cristian.capodicasa@polito.it

Abstract. In the long-term monitoring of full-scale structures, it is essential to analyse the temporal evolution of specific parameters to accurately evaluate the health condition of the structure. A comprehensive understanding requires sufficient amounts of heterogeneous data describing different aspects of structural behaviour. In this context, the integration of remotely acquired data, such as satellite data, into monitoring protocols could be strategic for filling gaps in on-site measurements, offering extended and continuous time series.

Considering the structure, the environment, and the soil as an integrated system, this study proposes a method for reconstructing historical time series of soil parameters that could influence structural response. The model used is based on sparse Bayesian learning (SBL) and combines remotely sensed data with local measurements, achieving an accurate reconstruction of in situ parameters. This approach allows for the simulation of soil and structural behaviour even in the absence of direct data, improving the ability to distinguish between physiological variations and real signs of damage. This is essential to avoid interpretative errors that could compromise the effectiveness of monitoring and the safety of structures.

Keywords: SHM, satellite data, on-site monitoring, long-term monitoring

Introduction

The progressive ageing of civil structures, combined with continuous exposure to environmental and operational actions, has made structural health monitoring (SHM) a key tool for ensuring structural safety and durability. SHM systems enable tracking of structural behaviour through data acquisition and analysis, supporting performance assessment under both operational conditions and extreme events. In this context, data-driven approaches have gained increasing relevance, as they allow the exploitation of monitoring data to describe the temporal evolution of structural parameters.

A fundamental requirement for effective SHM is the availability of long-term and continuous datasets. However, permanent monitoring systems are currently limited to a relatively small number of structures, typically those of high strategic or cultural importance [1], [2]. Their installation and maintenance involve significant costs and technical complexity, often resulting in datasets characterised by limited temporal coverage. This limitation affects not only structural measurements but also environmental and soil-related variables, which are essential for a comprehensive

https://doi.org/10.58286/33796

interpretation of structural response. Environmental and Operational Variations (EOVs) are known to significantly influence structural behaviour [3], [4], inducing reversible changes that are not associated with damage. These effects complicate data interpretation and may lead to misclassification of structural conditions, particularly in data-driven frameworks [5]. Distinguishing between physiological variations and damage-related changes therefore remains a critical challenge.

To address data scarcity and improve temporal coverage, remote sensing techniques have emerged as a valuable complementary source of information. This study proposes a sparse Bayesian learning-based (SBL) framework to integrate remote sensing and on-site data, enabling the reconstruction of historical time series of soil-related parameters and supporting a more reliable interpretation of long-term structural behaviour.

1. Research Background

A fundamental aspect of SHM is the analysis of the temporal evolution of structural response, which provides insight into the health state of a structure over time. Long-term monitoring systems are particularly valuable in this context, as they enable the identification of trends, anomalies, and changes in structural behaviour. Nevertheless, the relative scarcity of permanent monitoring installations often results in fragmented datasets, hindering the development of robust models for structural assessment. One of the primary challenges in interpreting monitoring data lies in the influence of EOVs. Structural parameters, both static and dynamic, are affected by external factors such as temperature, humidity, and soil conditions, which can induce significant yet reversible changes in structural response. These variations introduce uncertainty into data analysis and may compromise the reliability of damage detection strategies, especially when data-driven methods are employed. Consequently, the integration of environmental information has become a key aspect of modern SHM approaches, allowing for a more accurate interpretation of structural behaviour and a reduction of false alarms.

In this context, recent research has explored the use of remote sensing technologies to complement traditional monitoring systems. Satellite-based techniques, particularly interferometric methods, have been widely adopted for the measurement of ground displacements at the millimetric scale. These approaches have demonstrated their effectiveness across a range of applications, including individual structures , infrastructure networks [6], and large urban areas [7]. In addition to displacement measurements, satellite data can also provide geophysical and environmental information [8], offering new opportunities for integrating structural and environmental datasets.

Despite these advantages, the use of remote sensing data in SHM presents several challenges. In particular, the physical interpretation of remotely sensed variables is often not straightforward and requires validation through comparison with in situ measurements. Moreover, the integration of heterogeneous data sources demands advanced modelling techniques capable of capturing complex and potentially non-linear relationships between environmental inputs and structural response. Among the factors influencing structural behaviour, soil-structure interactions (SSIs) play a crucial role, yet they are often insufficiently characterised due to the scarcity of soil-related measurements. This limitation is particularly relevant for complex or monumental structures, where soil conditions may significantly affect both static and dynamic responses. Recent studies have addressed this issue by investigating SSI and structure-soil-structure interactions (SSSIs), highlighting the dependence of soil mechanical properties on environmental conditions [9], [10].

To address these challenges, data-driven approaches capable of handling limited and heterogeneous datasets are required. SBL represents a promising solution in this regard, as it enables the modelling of complex relationships while promoting sparsity and robustness. Although SBL has been primarily applied to damage detection problems in SHM [11], its extension to multi-output regression offers the potential to capture hidden dependencies between environmental variables and structural response. This capability is particularly relevant for reconstructing historical time series and inferring unobserved parameters, thereby enhancing the understanding of long-term structural behaviour within a coupled structure-soil-environment system.

2. Data Integration for Multi-Output Regression Using Sparse Bayesian Learning

Understanding the evolution of structural response under environmental and soil variations is crucial in structural monitoring, particularly for data-driven approaches. The effectiveness of SHM depends heavily on the type, quality, and quantity of data collected. Integrating heterogeneous data sources, including structural, soil-related, and environmental variables, can provide a more comprehensive understanding of structural behaviour, even when long-term monitoring records are limited. Satellite-based observations play a critical role in this context, enabling the reconstruction of historical time series for periods where in situ data are sparse or unavailable, thereby filling essential gaps in the monitoring record.

This study proposes a flexible multi-input multi-output regression methodology within the SBL paradigm to model the influence of soil and environmental factors on structural response. Specifically, relevance vector machines (RVMs) are employed to integrate multiple explanatory variables through custom basis functions, allowing relevant relationships to be identified without prior knowledge of the underlying interactions.

A preliminary linear correlation analysis using Pearson's coefficient r [12], can help identify candidate input variables. However, relationships between structural responses and environmental inputs are rarely purely linear. The SBL-RVM framework addresses this by performing a non-linear, multidimensional correlation analysis. Sparsity-promoting priors allow the model to discard basis functions that contribute negligibly, effectively selecting the most informative input-output relationships and quantifying their influence on the target outputs. Model performance is evaluated using the normalized mean absolute error (NMAE) [13], which provides a scale-independent metric for assessing predictive accuracy.

2.1 Sparse Bayesian Learning Methodology

To handle multi-input, multi-output non-linear regression problems, the proposed methodology builds upon the concept of a dictionary, widely employed in compressive sensing [14], [15]. The dictionary matrix $\mathbf{D}$ collects candidate basis functions representing the input signals, while regression weights $\boldsymbol{\beta}$ are estimated to reconstruct the output $\mathbf{x}$ according to

$$\mathbf{x} = \mathbf{D}\boldsymbol{\beta} . \quad (1)$$

The weights are modelled as Gaussian-distributed variables, with hyperparameters controlling sparsity (α) and observation noise precision ($\rho = \sigma^{-2}$), which are iteratively optimised via type-II maximum likelihood. This process automatically prunes irrelevant basis functions and adaptively estimates noise levels, enhancing the robustness of the model.

A key advantage of the RVM-SBL framework is its flexibility in designing basis functions tailored to the problem. Series expansions, such as polynomial or Fourier series expansions, can be incorporated into the dictionary, with latent parameters defining the expansion order. The algorithm iterates across latent parameter sets, selecting the optimal configuration by minimising a cost function J_c that balances training and validation errors, thereby preventing overfitting while capturing complex non-linear dependencies in multi-variable systems. The resulting posterior distributions over weights and predictions yield multivariate Gaussian outputs, providing mean estimates $\mathbf{y}_*$ and predictive variances V_* for each output variable at each time instant. This probabilistic formulation enables uncertainty quantification, which is crucial for distinguishing between normal environmental variations and potential damage-related anomalies.

By integrating satellite-derived environmental data with local in situ measurements, the methodology facilitates the reconstruction of historical time series of soil and environmental variables, allowing the estimation of previously unobserved states. This approach captures the complex interactions between structure, soil, and environment, even under non-linear and multivariate dependencies, and extends the temporal coverage of monitoring records. Overall, the SBL-based framework enables efficient exploitation of heterogeneous datasets in SHM, improving the interpretability of structural responses influenced by soil conditions and supporting more reliable long-term structural health assessment.

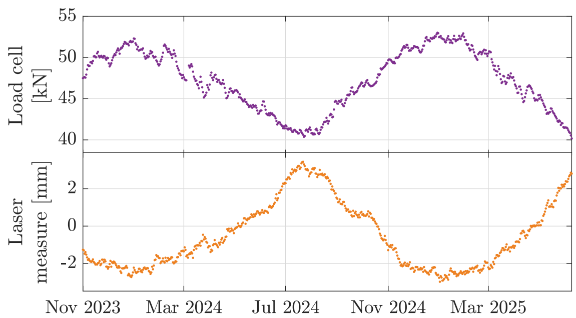
3. Application to a Real Monitored Structure

Fig. 1. The Sanctuary of Vicoforte.

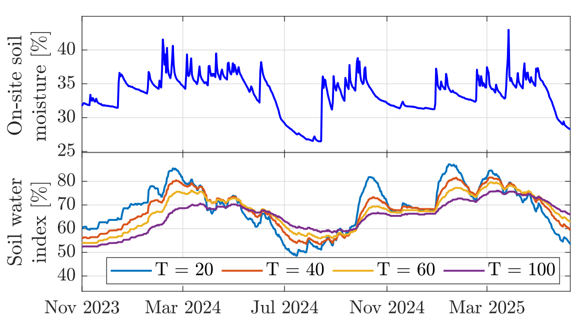
Investigating structural behaviour under environmental variations is often limited by the lack of diverse and simultaneous datasets. The Sanctuary of Vicoforte (Fig. 1) was selected as a case study due to its extensive monitoring system, comprising over 120 sensors, including load cells, laser measures, thermometers, thermo-hygrometers, accelerometers, piezometers, and soil sensors. This system enables continuous assessment of structural, environmental, and soil conditions, making the Sanctuary one of the most extensively monitored masonry structures worldwide. Its massive oval dome (major axis 37.23 m, minor axis 24.89 m) is the largest among oval masonry domes and the fifth-largest globally [16]. Structural monitoring began in the 1980s, with a 1987 reinforcement system using 56 post-tensioned Dywidag steel tie rods. A low-rate monitoring system was added in 2004 and restored in 2024, tracking stresses, crack widths, temperature, humidity, and dome displacements. A dynamic monitoring system installed in 2015 generates historical time series of natural frequencies [17]. The north-west bell tower is also monitored by mixed-rate sensors.

In this study, data processing integrates in situ structural measurements with environmental (air temperature, humidity, rainfall) and satellite-derived soil moisture (SWI) data. Structural data focus on load cells and laser measures (Fig. 2, a), particularly near the main crack above the west dome ring, showing seasonal cyclicity influenced by differing thermal expansion of masonry and steel [18]. Soil data (Fig. 2, b), including in situ thermo-hygrometer measurements and SWI from Sentinel-1 and MetOp satellites, demonstrate trends consistent with ground measurements, though with some amplitude and mean discrepancies [19], [20], [21]. Environmental data (Fig. 2, c) from ARPA Piemonte (Mondovì station) confirm seasonal patterns in air temperature and rainfall, correlating with soil moisture variations [4], [22].

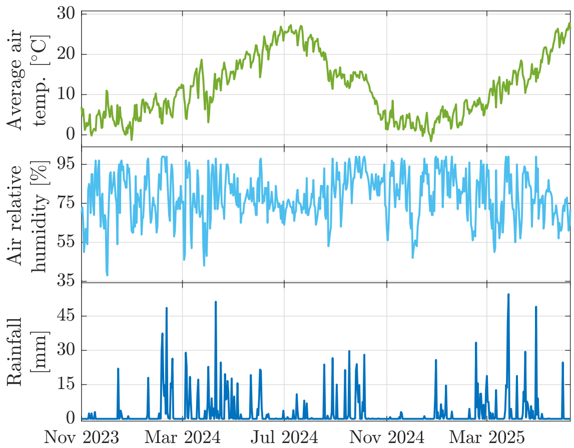
This integrated approach allows a comprehensive analysis of structural, environmental, and geotechnical interactions, providing valuable insights for validating numerical models under varying environmental conditions.



(a)



(b)



(c)

Fig. 2. Time series of data acquired on the Sanctuary of Vicoforte: (a) structural data, (b) soil data, and (c) environmental data.

4. Correlation Analysis

Following data collection, all datasets were pre-processed to remove outliers and standardise sampling. Initial scatter plots revealed potential non-linear behaviours, prompting Pearson's correlation analysis. Structural monitoring focused on load cell LC 15 and laser measure LM 2, which exhibited a strong negative correlation ($r = -0.953$) with seasonal patterns, largely influenced by differing thermal expansion between steel and masonry. Both parameters also correlated significantly with air temperature (load cell $r = 0.932$, laser measure $r = 0.887$). The feasibility of substituting on-site soil moisture (SM) with satellite-derived SWI was examined. SWI(20) showed the highest correlation with SM ($r = 0.622$), with correlations decreasing for higher T-values (SWI(40) $r = 0.572$, SWI(60) $r = 0.490$, SWI(100) $r = 0.325$). Structural correlations with soil moisture were moderate for SM (LC $r = 0.350$, LM $r = -0.424$) but stronger for SWI(20) (LC $r = 0.668$, LM $r = -0.709$). Air temperature correlated more strongly with SWI(20) ($r = -0.533$) than with SM ($r = -0.284$), suggesting indirect temperature effects on structural response.

Overall, relationships between structural behaviour, soil moisture, and environmental parameters are predominantly non-linear. This highlights the limitations of simple correlation analyses and the need for a more comprehensive statistical framework to capture complex interactions and reliably integrate satellite-derived soil data for soil-structure interaction studies.

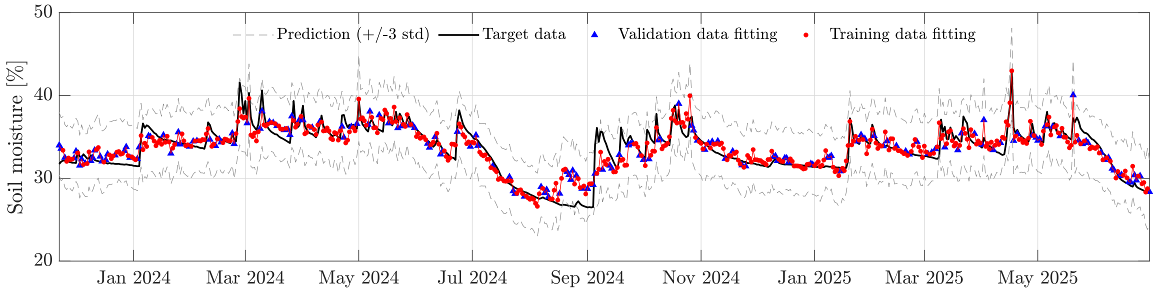
5. Predicting Soil Moisture Time Series Using SBL Regression

The SBL regression model was applied to predict soil moisture (SM) dynamics and to generate time series preceding in situ monitoring. Basis functions, including polynomial and Fourier series, were carefully selected, and predictors were chosen based on prior correlation studies. The proposed model allows the automatic selection of the expansion order of polynomial and Fourier components, thus reducing the need for manual hyperparameter tuning. The model was trained and validated on the period from 22 September 2023 to 30 June 2025, using 70% of data for training and 30% for validation. Initially, SWI(20) was used as the sole predictor due to its strong correlation with on-site SM, followed by incremental inclusion of all SWI series ($T = 20 - 100$) and environmental variables (air relative humidity RH_{air}, rainfall R, air temperature T_{air}). Latent parameters were optimised using a cost function assigning 40% weight to training error and 60% to the validation-training error difference.

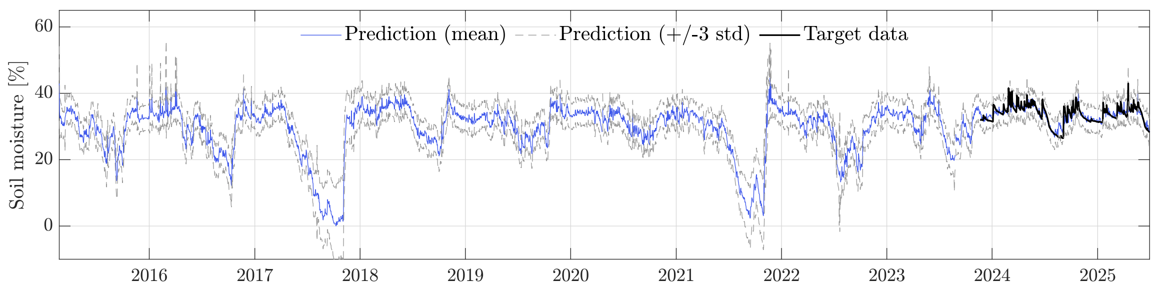
Table 1. Goodness-of-fit metrics of the regression model for different predictors

Predictors of SM	NMAE [%]		Overfitting ratio
	Training	Validation	
SWI(20)	4.592	4.548	0.990
SWI(20-100)	3.076	3.253	1.058
SWI(20-100), RH_{air}	2.776	3.063	1.103
SWI(20-100), RH_{air}, R	2.510	2.912	1.160
SWI(20-100), RH_{air}, R, T_{air}	2.598	2.951	1.139

Goodness-of-fit metrics demonstrate substantial improvement with additional predictors (Table 1). The final model achieved NMAE ≈ 2-3%, while overfitting ratios remained close to 1.2, confirming strong generalisation. Such results are promising in the perspective of integration of satellite data for SHM purposes, as also shown in Fig. 3, where the model is able to capture both overall trend and short-term fluctuations of soil moisture.



(a)



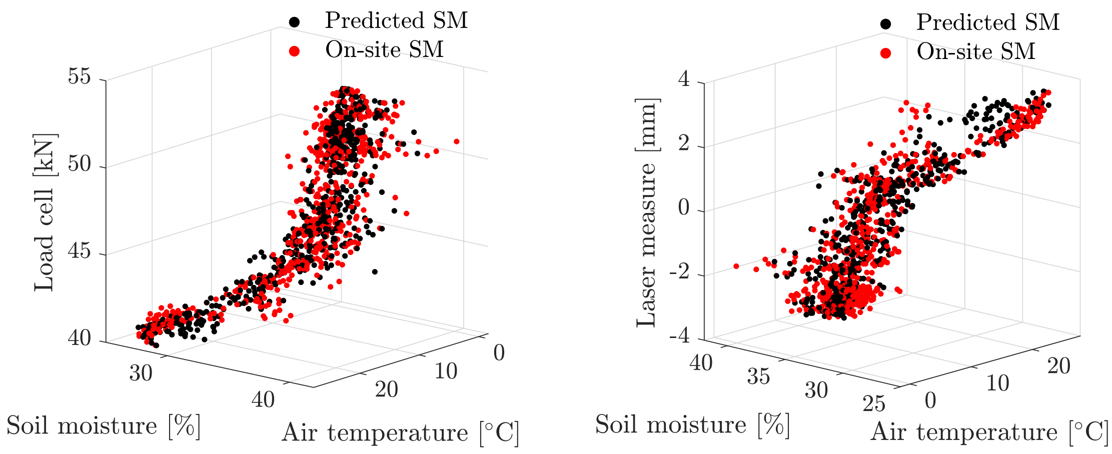
(b)

Fig. 3. Soil moisture regression model: (a) training and validation datasets, (b) testing dataset.

The good quality of the regression model is also confirmed by Fig. 4, where scatter plots of the structural parameters and both the predicted soil moisture and the measured target values are compared. This is further confirmed by the relatively small normalised absolute differences of the correlation coefficients in Table 2.

Table 2. Correlation coefficient comparison of structural parameters with on-site and predicted soil moisture

r	Source SM	Pred. SM	Normalised abs. diff. [%]
LC15	0.3503	0.3738	6.287
LM2	-0.4243	-0.4767	10.997



(a) (b)

Fig. 4. 3D scatterplot of environmental variables with: (a) load cell LC15, (b) laser measure LM2.

6. Conclusions

This study presents a robust framework for integrating environmental and structural monitoring data in the context of soil-structure interaction. By applying an SBL-based regression approach to the extensively monitored Sanctuary of Vicoforte, the methodology accurately captured soil moisture dynamics and their influence on structural behaviour. The correlation analysis revealed strong relationships between structural parameters and environmental variables, particularly soil moisture and air temperature, while highlighting non-linear effects around ~31% soil moisture that limit linear modelling approaches.

The SBL regression model successfully predicted on-site soil moisture with a normalised mean absolute error below 3% and a normalised absolute difference under 10% compared with measured data, preserving key dual-behaviour patterns and demonstrating robust generalisation to historical periods lacking measurements. By integrating satellite-derived SWI with additional environmental variables, the approach provides a flexible, computationally efficient, and reproducible framework for reconstructing missing environmental data and supporting structural health monitoring.

Looking forward, extending this methodology to model soil mechanical properties and coupling it with advanced finite element modelling techniques could enable a non-invasive, predictive link between soil states and structural modal parameters, enhancing the understanding of environment-structure interactions within complex SSI scenarios.

References

[1] C. Gentile, A. Ruccolo, and F. Canali, ‘Continuous monitoring of the Milan Cathedral: dynamic characteristics and vibration-based SHM’, *J Civil Struct Health Monit*, vol. 9, no. 5, pp. 671–688, Nov. 2019, doi: 10.1007/s13349-019-00361-8.

[2] F. Lorenzoni, F. Casarin, C. Modena, M. Caldon, K. Islami, and F. da Porto, ‘Structural health monitoring of the Roman Arena of Verona, Italy’, *Journal of Civil Structural Health Monitoring*, vol. 3, pp. 227–246, 2013.

[3] R. Ceravolo, G. Coletta, G. Miraglia, and F. Palma, ‘Statistical correlation between environmental time series and data from long-term monitoring of buildings’, *Mechanical Systems and Signal Processing*, vol. 152, p. 107460, 2021.

[4] H. Sohn, ‘Effects of environmental and operational variability on structural health monitoring’, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1851, pp. 539–560, 2007.

[5] Y. Huang, ‘A state-of-the-art review on Bayesian inference in structural system identification and damage assessment’, *Structure and Infrastructure Engineering*, 2019.

[6] M. Lazecky, D. Perissin, M. Bakon, J. M. de Sousa, I. Hlavacova, and N. Real, ‘Potential of satellite InSAR techniques for monitoring of bridge deformations’, in *2015 Joint Urban Remote Sensing Event (JURSE)*, 2015, pp. 1–4. doi: 10.1109/JURSE.2015.7120497.

[7] M. Bonano, M. Manunta, A. Pepe, L. Paglia, and R. Lanari, ‘From previous C-band to new X-band SAR systems: Assessment of the DInSAR mapping improvement for deformation time-series retrieval in urban areas’, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 1973–1984, 2013.

[8] S. Coccimiglio, G. Coletta, E. Lenticchia, G. Miraglia, and R. Ceravolo, ‘Combining satellite geophysical data with continuous on-site measurements for monitoring the dynamic parameters of civil structures’, *Scientific Reports*, vol. 12, no. 1, p. 2275, 2022.

[9] G. Mylonakis and G. Gazetas, ‘SEISMIC SOIL-STRUCTURE INTERACTION: BENEFICIAL OR DETRIMENTAL?’, *Journal of Earthquake Engineering*, vol. 4, no. 3, pp. 277–301, Jul. 2000, doi: 10.1080/13632460009350372.

[10] M. Lou, H. Wang, X. Chen, and Y. Zhai, ‘Structure–soil–structure interaction: Literature review’, *Soil Dynamics and Earthquake Engineering*, vol. 31, no. 12, pp. 1724–1731, Dec. 2011, doi: 10.1016/j.soildyn.2011.07.008.

[11] R. Hou, X. Wang, Q. Xia, and Y. Xia, ‘Sparse Bayesian learning for structural damage detection under varying temperature conditions’, *Mechanical Systems and Signal Processing*, vol. 145, p. 106965, 2020.

[12] M. Ezekiel, *Methods of correlation analysis*. New York, USA: Wiley, 1930.

[13] D. A. Freedman, *Statistical models: theory and practice*. cambridge university press, 2009.

[14] R. Fuentes, N. Dervilis, K. Worden, and E. J. Cross, ‘Efficient parameter identification and model selection in nonlinear dynamical systems via sparse Bayesian learning’, in *Journal of Physics: Conference Series*, IOP Publishing, 2019, p. 012050.

[15] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, ‘Compressive sensing: From theory to applications, a survey’, *Journal of Communications and networks*, vol. 15, no. 5, pp. 443–456, 2013.

[16] M. L. Pecorelli, R. Ceravolo, G. De Lucia, and R. Epicoco, ‘A vibration-based health monitoring program for a large and seismically vulnerable masonry dome’, in *Journal of Physics: Conference Series*, IOP Publishing, 2017, p. 012009.

[17] S. Coccimiglio, G. Miraglia, V. Cavanni, A. Crocetti, and R. Ceravolo, ‘Automated Mode Tracking via Supervised Classification and Adaptive Parameter Calibration for Seismic Monitoring with Sparse Sensors’, *Bulletin of Earthquake Engineering*, In press.

[18] R. Ceravolo, A. De Marinis, M. L. Pecorelli, and L. Zanotti Fragonara, ‘Monitoring of masonry historical constructions: 10 years of static monitoring of the world’s largest oval dome’, *Struct. Control Health Monit.*, vol. 24, no. 10, p. e1988, Oct. 2017, doi: 10.1002/stc.1988.

[19] R. Boulkhiout, R. Boulkhiout, S. Messast, and S. Messast, ‘Effect of Soil on the Seismic Response of Structures Taking Into Consideration Soil-Structure Interaction’, *International Journal of Geotechnical Earthquake Engineering*, 2020, doi: 10.4018/ijgee.2020070104.

[20] ‘S1 Mission’. Accessed: Dec. 13, 2024. [Online]. Available: https://sentiwiki.copernicus.eu/web/s1-mission

[21] ‘MetOp - Earth Online’. Accessed: Dec. 13, 2024. [Online]. Available: https://earth.esa.int/eogateway/missions/metop

[22] H. Sohn, K. Worden, and C. R. Farrar, ‘Statistical damage classification under changing environmental and operational conditions’, *Journal of Intelligent Material Systems and Structures*, vol. 13, no. 9, pp. 561–574, 2002, doi: 10.1106/104538902028492.