

Proceeding Paper

Regression Tree Ensemble to Forecast Thermally-Induced Responses of Long-Span Bridges [†]

Alireza Entezami ^{*}, Bahareh Behkamal, Carlo De Michele and Stefano Mariani

Department of Civil and Environmental Engineering, Politecnico di Milano, Milano, Italy; bahareh.behkamal@polimi.it (B.B.); carlo.demichele@polimi.it (C.D.M.); stefano.mariani@polimi.it (S.M.)

^{*} Correspondence: alireza.entezami@polimi.it

[†] Presented at the 10th International Electronic Conference on Sensors and Applications (ECSA-10), 15–30 November 2023; Available online: <https://ecsa-10.sciforum.net/>.

Abstract: The ambient temperature is a critical factor affecting the deformation of long-span bridges, due to its seasonal fluctuations. Although there exist various sensor technologies and measurement techniques to extract the actual structural response in terms of the displacement field, this is a demanding task in long-term monitoring. To address this challenge, data prediction looks as the best solution. In this paper, the thermally-induced response of a long-span bridge is forecasted with a regression tree ensemble method in conjunction with Bayesian hyperparameter optimization, adopted to tune the proposed regressor. Results testify that the offered method is reliable when there is a linear correlation between the temperature and the induced structural deformation, hence in terms of the thermally-induced displacement field.

Keywords: long-span bridges; supervised learning; regression tree ensemble; temperature effects; remote sensing

1. Introduction

Ambient temperature affects the response of civil structures to the external excitations [1], especially long-span bridges [2]. The critical impact of this environmental factor is related to its seasonal variability [3], regardless of the type, size, and material the structures are made of [4,5]. Notably, temperature variability is an important challenge for long-term structural health monitoring (SHM). Seasonal temperature fluctuations in summer and winter profoundly affect the structural response and the physical properties, leading to changes to the structural characteristics. Several experimental and field monitoring studies demonstrated that varying temperature conditions may have a more significant effect on the structural behavior than operational loadings [6]. This is because spatial and temporal temperature variations cause thermal loads, unpredictable internal stresses and forces in elements, and also changes in the boundary conditions, particularly for large-scale bridges, all of which lead to damages such as cracks in concrete and yielding of steel elements. Therefore, it is essential to regularly monitor structures like long-span bridges under varying ambient temperature effects, to remove such effects [7–9].

The routine and simple monitoring process usually consists in conducting field measurements. Accordingly, structures are equipped with sensors to measure structural and environmental data with accelerometers, strain gauges, thermocouples, anemometers, etc. Recently, the technology of spaceborne remote sensing has helped civil engineers to leverage remote sensors mounted on satellites (like e.g., Sentinel-1, TerraSar-X, COSMO-SkyMed) to capture optical or synthetic aperture radar (SAR) images, and extract structural displacements. The great advantage of this technology is the lack of contact sensors, which represent a considerable cost of the SHM system and may be

Citation: Entezami, A.; Behkamal, B.; De Michele, C.; Mariani, S. Regression Tree Ensemble to Forecast Thermally-Induced Responses of Long-Span Bridges. *Eng. Proc.* **2023**, *56*, x. <https://doi.org/10.3390/xxxxx>

Academic Editor(s): Name

Published: 15 November 2023



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

broken in case of extreme environmental events (e.g., typhoons). Despite the development of new sensing technologies for data measurement, a field measurement process may show several limitations. First, it is not simple to perform a long-term field measurement program. Second, although the spaceborne remote sensing facilitates extracting structural displacement responses, the outcome of this technology (i.e., optical and SAR images) often has huge size (in the unit of GB) leading to limitations linked to memory storage. Third, spaceborne remote sensing cannot supply rich data/features as obtained with SHM via contact-based sensing. In most long-term SHM projects, the monitoring problem is based on small data [10]. However, it is difficult to provide such entire data stream from remote sensors. One of the most appropriate solutions to this challenge is to implement data forecasting or prediction.

According to the concept of data forecasting, it suffices to define the main predictors (input or independent data) and relevant responses (output or dependent data). In health monitoring of civil structures, the environmental conditions (i.e., temperature, humidity, wind, etc.) are related to the predictors, while changes in the structures themselves pertain to the responses. Using such information, supervised regression modeling is the best approach to data forecasting [2,11,12].

The main objective of this paper is to propose a regression ensemble tree (RTE) in conjunction with the Bayesian hyperparameter optimization (BHO), called here RTE-BHO to forecast the thermally-induced responses of long-span bridges. In this method, BHO supports RTE to select the best ensemble learning algorithm between bagging and boosting strategies. To evaluate the performance of the supervised regressor, small sets of displacement responses and ambient temperature records related to a long-span bridge were considered. The displacement histories were extracted from 55 SAR images of TerraSar-X between 2009 and 2010. Results demonstrate that the proposed RTE-BHO provides an automated data forecasting approach, able to predict the displacement responses retrieved from the spaceborne remote sensing technology.

2. Regression Tree Ensemble

Ensemble learning is an advanced machine learning algorithms that can significantly improve the overall performance of any model. In particular, it has been an effective strategy for SHM [13]. For the problem of regression, the RTE method develops a supervised predictive model as a weighted combination of multiple regression trees. Indeed, the RTE benefits the concept of ensemble learning to enhance the decision tree regression that may suffer from high variance, bias, and overfitting. Therefore, multiple regression trees based on the concept of ensemble learning can be combined to make an ensemble of trees and improve the prediction performance.

Bagging and boosting are two popular and tried-and-tested ensemble learning algorithms adopted in RTE. Bagging or bootstrap aggregating utilizes multiple separate sets from the original, randomly generated training data with replacement to train the different regression trees. Let B denotes the number of trees to generate B different bootstrapped training datasets. The predictions of B trees, defined as $\{\hat{y}_1, \dots, \hat{y}_B\}$, can be computed and then averaged to reduce the variance with respect to a single regression tree. The averaging process in the bagging method is given by:

$$\hat{y}_{Bag} = \frac{1}{B} \sum_{i=1}^B \hat{y}_i \quad (1)$$

where \hat{y}_{Bag} denotes the average prediction of all the regression trees. The boosting technique increases the number of trees B sequentially, so that each tree uses a modified version of the whole dataset accounting for information from the previously developed trees.

The fundamental principle of the boosting technique is to increase the number of trees sequentially, wherein each tree uses a modified version of the whole dataset through

information from the previously developed tree. This technique somehow resembles the bagging method except that the trees are grown sequentially, which means that each tree is grown using information from previously developed trees. The boosting technique does not involve bootstrap sampling; instead, each tree is fit to a modified version of the original dataset. The boosting learning process improves the prediction performance of each tree from \hat{y}_1 to \hat{y}_B by updating the weights of the training samples, without the bootstrap sampling. In this regard, the boosted model output \hat{y}_{Boost} is defined as follows:

$$\hat{y}_{Boost} = \sum_{i=1}^B \lambda \hat{y}_i \quad (2)$$

where λ is the shrinkage parameter (i.e., a constant value) that controls the rate of the boosting process for learning.

3. Bayesian Hyperparameter Optimization

In machine learning, a hyperparameter is an unknown element that affects the overall performance of a model [14]. Among the various hyperparameter optimization techniques, BHO is one of the most effective ones. This is because it can tune any type of unknown parameters based on Bayes' theorem. In the forecasting problem by regression modeling, the fundamental principle of BHO is to minimize the following objective function, along with its hyperparameter(s), in a bounded domain:

$$F = \log(1 + E_{MSE}) \quad (3)$$

where E_{MSE} stands for the cross-validation mean-squared-error (MSE) between the observation (i.e., real data) and estimation (i.e., predicted data) values. The function F denotes a logarithmic transformed validation loss, computed for the regressor and hyperparameter values at each iteration. Using this function, BHO considers a probability distribution model for the objective function of interest and the model is then updated at each new evaluation/iteration. In the following, an acquisition function is defined to maximize and also determine the next values(s) of the hyperparameter(s) of the objective function. One can also incorporate a posterior probability distribution for each hyperparameter. If there is any error in the iterative process, one should take more iterations until reaching convergence or consider a stopping condition. By minimizing the objective function F , the best hyperparameter set is that one leading to a good convergence (i.e., a minimum difference) between the observed and estimated data.

Regarding RTE, BHO undertakes the tuning of some important elements, such as the type of ensemble learning algorithm, the minimum number of lead node, the maximum number of decision splits (branch nodes), and the shrinkage parameter in Equation (2).

4. Case Study: The Lupu Bridge

The Lupu Bridge is a steel arch bridge that crosses the Huangpu River in Shanghai, see Figure 1a. This structure features a total length of 750 m, due to a main span of 550 m and two side spans of 100 m. Figure 1b provides the side view of the bridge, along with its main dimensions and structural components. The girder in the side span is characterized by a closed steel box with a width of 41 m and the height of 2.7 m. The box-girder was fixed with the arch ribs, columns, and end cross-beam of the side spans. More structural details can be found in Qin et al. [15].

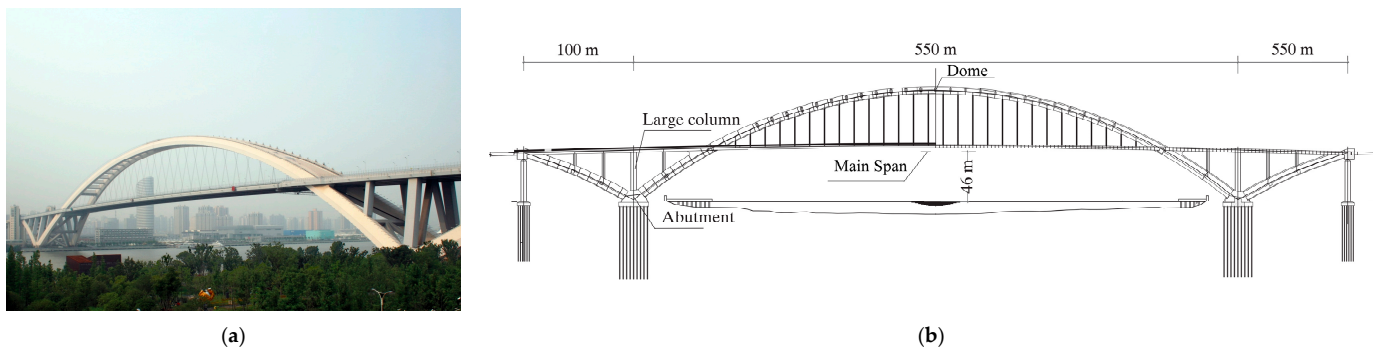


Figure 1. (a) The Lupu Bridge; (b) side view and main dimensions.

4.1. Thermally-Induced Structural Responses

Due to the sensitivity of the Lupu Bridge to geological and environmental conditions, a field measurement based on spaceborne remote sensing and limited SAR images was conducted to obtain the displacement responses of two critical components of the bridge in a long-term monitoring program [15]: the main span, and the dome. Accordingly, 55 SAR images from TerraSar-X between 2009 and 2010 were analyzed to extract the displacement points of the mentioned components. During this field measurement, air temperature data were also recorded, to provide rich information for the prediction problem. Figure 2 shows the time histories of the displacements associated with the dome and the main span, along with the temperature data.

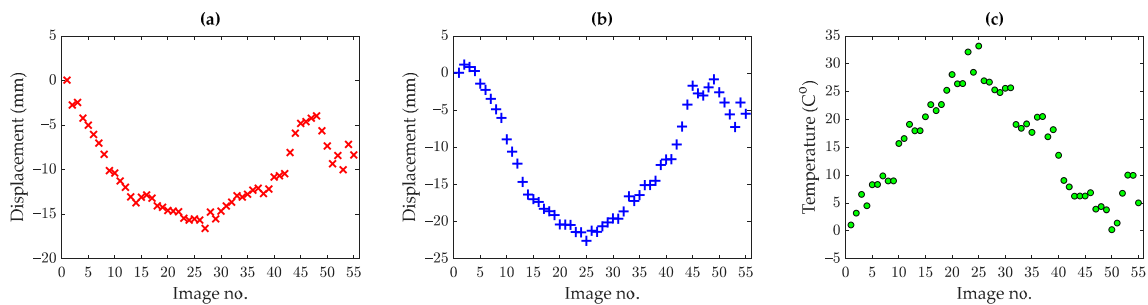


Figure 2. Displacement responses of the Lupu Bridge (a) at the dome, and (b) the main span; (c) relevant temperature data.

4.2. Forecasted Results

The measured temperature and extracted displacement evolutions were collected to be arranged as the training and test datasets. 80% of the temperature and displacement samples were used to generate the training set, while the remaining 20% of the samples were considered as the test set. The number of the training and test points were therefore equal to 44 and 11, respectively. Using BHO with 50 sample iterations, Table 1 reports the main hyperparameters of the RTE models to forecast the displacement responses.

Table 1. Bayesian hyperparameter optimization of the RTE model.

Hyperparameter	Bridge Element	
	Dome	Span
Ensemble learning algorithm	Bagging	Bagging
Number of leaves	5	4
Number of decision nodes	49	34

Using the trained RTE models, Figure 3 compares the predicted and real displacement samples, including those considered in the training process and those

related to the test set. As it can be observed, the predicted and real values are in good agreement with each other. Moreover, Figure 4 displays the regression plot by comparing the actual and predicted data: it can be seen that there are small deviations of all the data points from the straight line providing a perfect fit. To get additional insights, Table 2 gathers the values of the R-squared metric relevant to the RTE-BHO method, in case of different training ratios. The regressors developed for the dome and span have large R-squared values, close to one, when the training ratio is 80%. However, if the training rate is reduced, the R-squared values only slightly reduce.

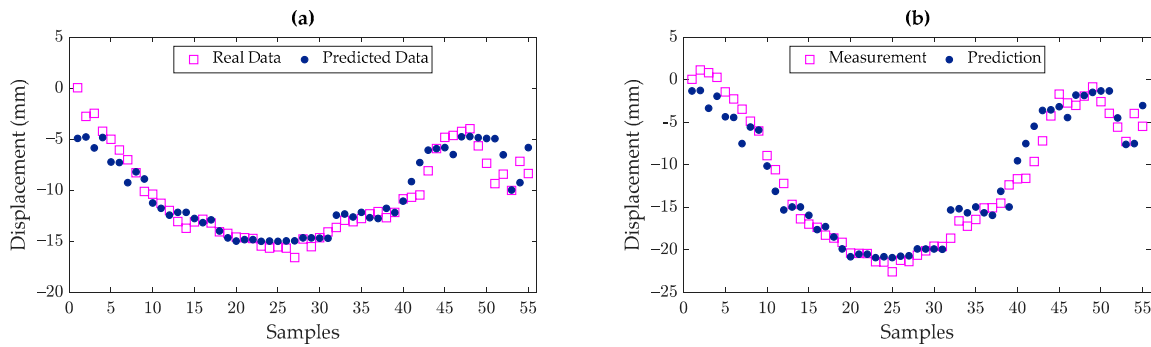


Figure 3. Measured and predicted displacements by the RTE model: (a) dome, and (b) main span.

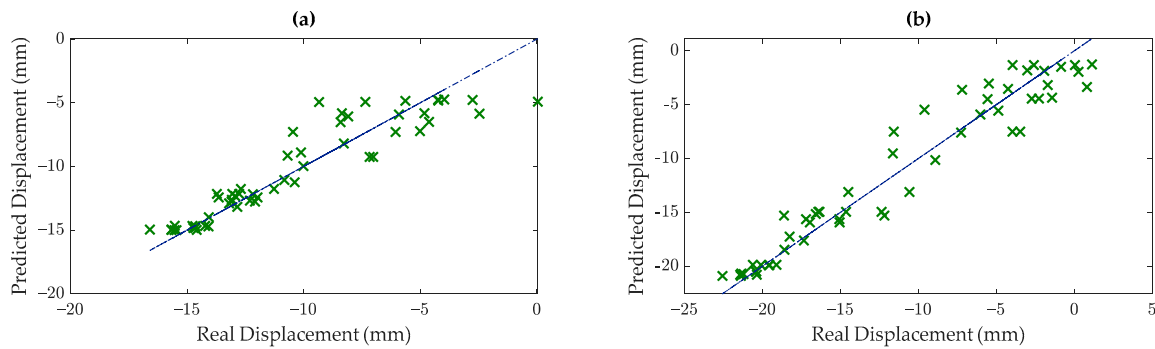


Figure 4. Predicted versus actual displacements, based on the RTE method: (a) dome, and (b) main span.

Table 2. Performance evaluation of the RTE-BHO method, in terms of the R-squared measure under a varying training ratio.

Bridge Element	Training Ratios		
	80%	50%	30%
Dome	0.8625	0.8453	0.7513
Span	0.9365	0.9121	0.8547

5. Conclusions

This paper has addressed the problem of forecasting structural displacements with remote sensing technology. The forecasting problem has been intended to evaluate the effect of ambient temperature on long-span bridges in long-term monitoring programs. A supervised regressor has been developed by combining a regression ensemble tree and Bayesian hyperparameter optimization.

Using the displacement responses at the dome and main span of the Lupu Bridge, the results have shown that the supervised regressor succeeds in forecasting the response data with a high overall performance, even in the case of a small dataset to learn the main features of the said structural response.

Author Contributions: Conceptualization, A.E. and B.B.; methodology, A.E., B.B., C.D.M. and S.M.; software, A.E. and B.B.; validation, A.E. and B.B.; formal analysis, A.E., B.B., C.D.M. and S.M.; investigation, A.E., B.B., C.D.M. and S.M.; resources, A.E. and B.B.; data curation, A.E. and B.B.; writing—original draft preparation, A.E. and B.B.; writing—review and editing, A.E., B.B., C.D.M. and S.M.; visualization, A.E., B.B., C.D.M. and S.M.; supervision, C.D.M. and S.M.; project administration, A.E., B.B., C.D.M. and S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research partially funded by the European Space Agency (ESA) under contract no. 4000132658/20/NL/MH/ac.

Data Availability Statement: Data is unavailable due to privacy or ethical restrictions.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Han, Q.; Ma, Q.; Xu, J.; Liu, M. Structural health monitoring research under varying temperature condition: A review. *J. Civ. Struct. Health Monit.* **2021**, *11*, 149–173. <https://doi.org/10.1007/s13349-020-00444-x>.
2. Behkamal, B.; Entezami, A.; De Michele, C.; Arslan, A.N. Investigation of temperature effects into long-span bridges via hybrid sensing and supervised regression models. *Remote Sens.* **2023**, *15*, 3503. <https://doi.org/10.3390/rs15143503>.
3. Xia, Q.; Zhang, J.; Tian, Y.; Zhang, Y. Experimental study of thermal effects on a long-span suspension bridge. *J. Bridg. Eng.* **2017**, *22*, 04017034. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0001083](https://doi.org/10.1061/(ASCE)BE.1943-5592.0001083).
4. 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. <https://doi.org/10.1016/j.engstruct.2023.115616>.
5. 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. Signal Process.* **2023**, *201*, 110676. <https://doi.org/10.1016/j.ymsp.2023.110676>.
6. Wang, D.; Liu, Y.; Liu, Y. 3D temperature gradient effect on a steel–concrete composite deck in a suspension bridge with field monitoring data. *Struct. Control Health Monit.* **2018**, *25*, e2179. <https://doi.org/10.1002/stc.2179>.
7. Daneshvar, M.H.; Sarmadi, H.; Yuen, K.-V. A locally unsupervised hybrid learning method for removing environmental effects under different measurement periods. *Measurement* **2023**, *208*, 112465. <https://doi.org/10.1016/j.measurement.2023.112465>.
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.* **2023**, *in press*. <https://doi.org/10.1177/14759217231166116>.
9. Entezami, A.; Mariani, S.; Shariatmadar, H. Damage detection in largely unobserved structures under varying environmental conditions: An autoregressive spectrum and multi-level machine learning methodology. *Sensors* **2022**, *22*, 1400. <https://doi.org/10.3390/s22041400>.
10. Entezami, A.; De Michele, C.; Arslan, A.N.; Behkamal, B. Detection of partially structural collapse using long-term small displacement data from satellite images. *Sensors* **2022**, *22*, 4964. <https://doi.org/10.3390/s22134964>.
11. Kromanis, R.; Kripakaran, P. Predicting thermal response of bridges using regression models derived from measurement histories. *Comput. Struct.* **2014**, *136*, 64–77. <https://doi.org/10.1016/j.compstruc.2014.01.026>.
12. Laory, I.; Trinh, T.N.; Smith, I.F.C.; Brownjohn, J.M.W. Methodologies for predicting natural frequency variation of a suspension bridge. *Eng. Struct.* **2014**, *80*, 211–221. <https://doi.org/10.1016/j.engstruct.2014.09.001>.
13. Sarmadi, H.; Entezami, A.; Saeedi Razavi, B.; Yuen, K.-V. Ensemble learning-based structural health monitoring by Mahalanobis distance metrics. *Struct. Control Health Monit.* **2021**, *28*, e2663. <https://doi.org/10.1002/stc.2663>.
14. Yang, L.; Shami, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* **2020**, *415*, 295–316. <https://doi.org/10.1016/j.neucom.2020.07.061>.
15. Qin, X.; Zhang, L.; Yang, M.; Luo, H.; Liao, M.; Ding, X. Mapping surface deformation and thermal dilation of arch bridges by structure-driven multi-temporal DInSAR analysis. *Remote Sens. Environ.* **2018**, *216*, 71–90.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.