14th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP14
Dublin, Ireland, July 9-13, 2023

Probabilistic prediction of structural response of cable bridges based on structural health monitoring data

Minsun Kim
Graduate Student, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Seungjun Lee
Graduate Student, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Jingoo Lee
Graduate Student, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Young-Joo Lee
Associate Professor, Department of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

ABSTRACT: In recent years, various studies have attempted to predict the structural response of a cable bridge based on collected structural health monitoring data. However, it is still a challenging task to predict the structural response and assess the structural condition of a cable bridge because it involves various sources of uncertainty. In addition, because the number of sensors attached to a cable bridge is generally large, it is computationally expensive and impractical to use all of the sensor data for the response prediction. Therefore, this study proposes a new method to probabilistically predict the structural response of a cable bridge based on structural health monitoring data obtained from various sensors. To select meaningful sensor data for response prediction, a new index is developed based on their correlation coefficients. The proposed method employs Gaussian process regression (GPR), a nonparametric Bayesian method, to build a probabilistic prediction model based on the selected measurement data. The proposed method is tested by predicting the cable tension forces of an actual cable-stayed bridge in the Republic of Korea. Among the various sensors deployed on the target bridge, the new correlation-coefficient-based index allows us to select only the sensors that provide meaningful measurement data, which significantly increases the computational efficiency of the prediction. In addition, the prediction results show good agreement with the actual measurement results, thereby verifying that the proposed method can be used for the structural health monitoring of cable bridges.

1. INTRODUCTION

Since wireless smart sensor technology has been developed in previous studies (Jang et al., 2010; Sim et al., 2013; Park et al., 2014), structural health monitoring techniques have been widely applied to assess and monitor the structural condition of a cable bridge. For this purpose, several sensors are often deployed to provide various information such as displacement, strain, tension force, and temperature. Particularly, previous studies have attempted to predict the structural response of cable bridges based on the collected measurement data.

However, there are various sources of uncertainty involving with the response prediction of a bridge, owing to model misspecification, limited data size, etc., and one needs to consider such uncertainty (Lee et al., 2018). Previous studies have attempted to consider the uncertainty of the measurement data by employing probabilistic methods (Lee et al., 2018; Lee et al., 2019).

Because the number of sensors deployed on a cable bridge is generally large, it may cause low accuracy in predictive models and the prediction can be computationally expensive. Saunders et al. (2006) proposed an approach to feature selection to reduce the dimensionality of data and improve prediction accuracy and computational efficiency, which shows that it is necessary to select meaningful input data from various measurement data (Cai et al., 2018).

This study proposes a new method to probabilistically predict the structural response of a cable bridge based on structural health monitoring data obtained from various sensors. The proposed method selects relevant sensor data to improve the accuracy of the predictive model using a new correlation-coefficient-based index. In addition, the proposed method employs GPR to obtain a probabilistic prediction of the target response.

2. PROPOSED METHOD

In this study, a new probabilistic method is proposed for an optimal prediction model. The proposed method improves the accuracy and efficiency of the prediction model while dealing with the multiple cables of a cable bridge. In this regard, it involves two steps as shown in Figure 1. First, to improve a predictive model, meaningful input data is selected among the entire measurement data. Second, the selected input data is applied to a probabilistic predictive model employing GPR.

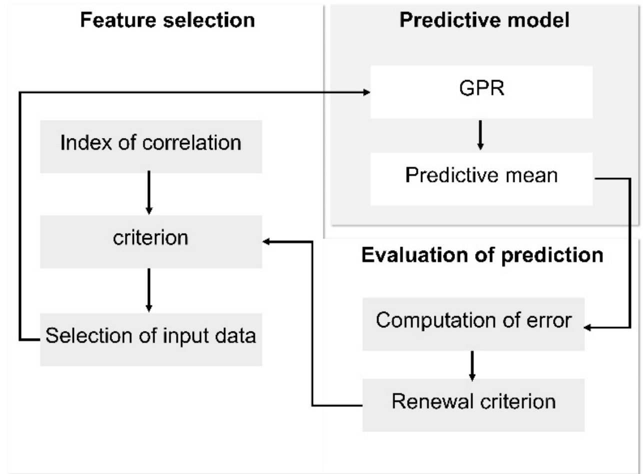

Figure 1: Framework for proposed method.

2.1. Index for feature selection

Although multiple measurement data are obtained by various sensors deployed on a cable bridge, it is not rational to construct a predictive model using all input data, because the corresponding high-dimensional data analysis leads to the complexity of a predictive model and can decrease learning accuracy. Therefore, to remove such irrelevant and redundant input data, Cai et al. (2018) suggested an evaluation measure for feature selection based on the Pearson correlation coefficient.

However, in the case of dealing with multiple tension forces (i.e., multiple output data), it is difficult to set a specific criterion based on correlation coefficients. To resolve this issue, in this study, the Pearson correlation coefficients among data sets are normalized as follows:

$$Z_j = \frac{|\rho_{i,j}| - |\rho_j|_{mean}}{|\rho_j|_{std}} \quad (1)$$

where Z_j is the normalized correlation coefficient (termed the Z-score) of the j-th output, $|\rho_{i,j}|$ is the absolute value of the Pearson correlation coefficient between the i-th input and j-th output, and $|\rho_j|_{mean}$ and $|\rho_j|_{std}$ are the mean and standard deviation of the j-th output, respectively. Figures 2 and 3 show the maximum and minimum values of the Pearson correlation coefficients and Z-scores of output datasets, and Figure 3 is observed the Z-score has a more uniform distribution, which helps establish a more consistent selection criterion.

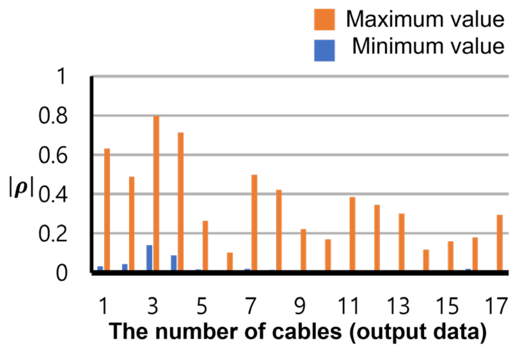

Figure 2: The maximum and minimum values of the Pearson correlation coefficients.

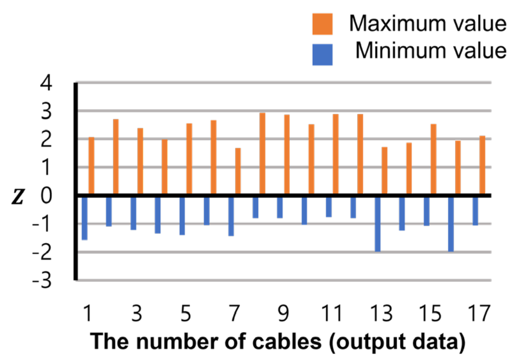

Figure 3: The maximum and minimum values of Z-scores.

2.2. Gaussian process regression

This study suggests a new probabilistic method for the prediction of tension force using GPR that is a machine learning-based non-parametric regression method and provides standard deviations as well as mean values in prediction. This section briefly describes GPR; the detailed theory of GPR can be found in Rasmussen (2003) and Lee et al. (2018).

GPR is described in relation of the input and output data by multivariate normal distribution as Equation 2, given as:

$$\begin{bmatrix} f(\mathbf{X}) \\ f_*(\mathbf{X}_*) \end{bmatrix} = N\left(0, \begin{bmatrix} \mathrm{K} + \sigma_n^2 \mathrm{I} & \mathrm{K}_* \\ \mathrm{K}_*^T & \mathrm{K}_{**} + \sigma_n^2 \mathrm{I} \end{bmatrix}\right), \quad (2)$$

where $\mathbf{X}, \mathbf{X}_*$ are respectively the training and test input matrices, $f(\cdot)$ is the prior function of $\mathbf{X}$, and $f_*(\cdot)$ is the posterior function of $\mathbf{X}_*$. Additionally, $N(\cdot)$ is the multivariate normal distribution, 0 is the zeros matrix, σ_n^2 is constant variance of the noise, and $\mathrm{K}, \mathrm{K}_*, \mathrm{K}_{**}$ are covariance matrix with embodied the kernel function that indicate the relationship of data.

3. APPLICATION EXAMPLE

To test the proposed method, it is used to predict the cable tension forces of an actual cable-stayed bridge in the Republic of Korea, the 2nd Jin-do bridge. Figure 4 briefly shows the target bridge and the locations of sensors. The input dataset comprises information on the temperature, wind speed, and the direction of wind. In addition, the tension force is predicted by the learning model. For all datasets, five-week measurement results were used: for four weeks as training data and for a week as test data.

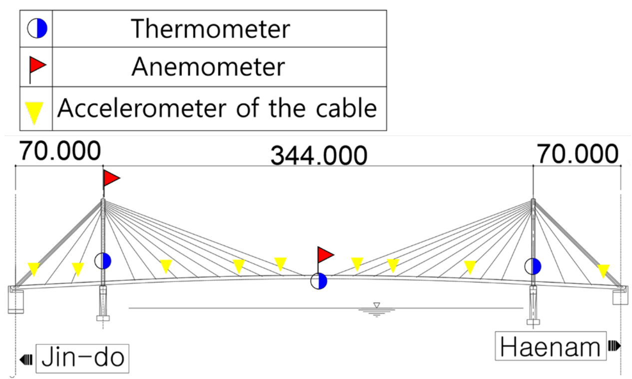

Figure 4: Location of sensors at the 2nd Jin-do bridge.

To evaluate the prediction performance, in this study, the prediction error is defined as the symmetric mean absolute percentage of error (sMAPE) (Hyndman and Koehler, 2006) which is a method based on a ratio, which calculates the quantity of the error between the mean prediction and measurement data. In addition, for a comparison purpose, two cases of selecting datasets based on the Pearson correlation coefficients and Z-scores are addressed.

3.1. Composition of Criterion: Pearson correlation coefficient

In this case, the criterion of feature selection is established as an absolute value of the Pearson correlation coefficient. That is, after the absolute value of the correlation coefficient between each output and input data is calculated, the input data is selected as a specific criterion.

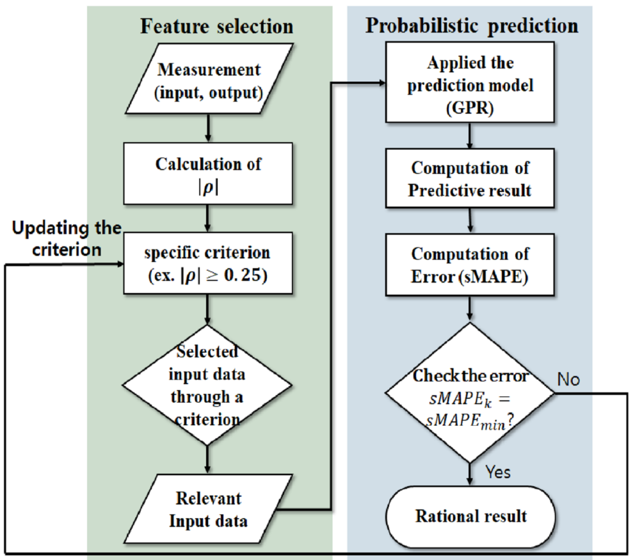

Figure 5: The flowchart of feature selection and probabilistic prediction using Pearson correlation coefficients.

Figure 5 shows the flowchart of feature selection and probabilistic prediction using Pearson correlation coefficients, and Figure 6 shows the analysis results. In the figure, the selection criterion constitutes the interval of correlation coefficient with 0.05 for the prediction accuracy, and it was observed that improved accuracy can be achieved in prediction with a certain selection criterion for each cable. However, the optimal selection criteria vary from cable to cable, which makes it difficult to establish a consistent selection criterion.

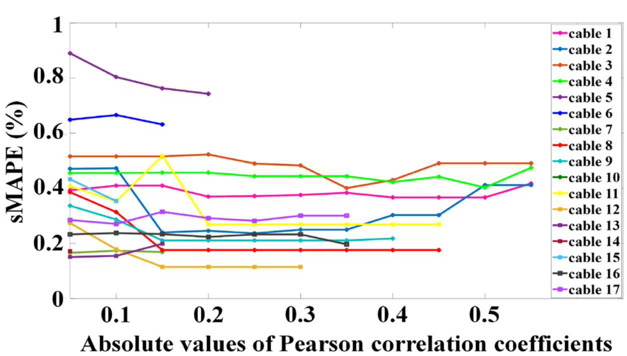

Figure 6: All cable's error of prediction means based on absolute values of Pearson correlation coefficients.

3.2. Composition of criterion: normalization of correlation coefficient (Z-score)

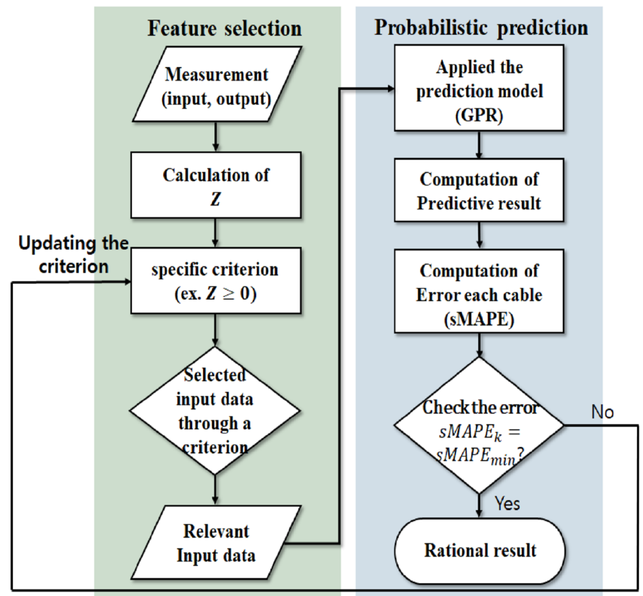

Figure 7: The flowchart of feature selection and probabilistic prediction using Z-scores.

In this case, the criterion of feature selection is established based on the Z-score described in Equation 1, and Figure 7 shows the flowchart of feature selection.

To investigate the performance of the Z-score-based selection criterion, it was tested with the interval of Z-score with 0.25. Table 1 shows the Z-score-based selection criterion and the number of minimum prediction errors for target cables, and Figure 8 shows the analysis results. 0.5 and 0.75 shows better performance than the others as a threshold value. In addition, because the error sum was slightly smaller with 0.5, in this example, it was decided that 0.5 was the best threshold value for the Z-score-based selection criterion.

Table 1: Z-score-based selection criterion and the corresponding number of minimum prediction errors.

Z-score-based criterion	The number of minimum prediction errors
$z \geq -1$	3

$z \geq -0.75$	2
$z \geq -0.5$	0
$z \geq -0.25$	3
$z \geq 0$	1
$z \geq 0.25$	2
$z \geq 0.5$	7
$z \geq 0.75$	7
$z \geq 1$	6

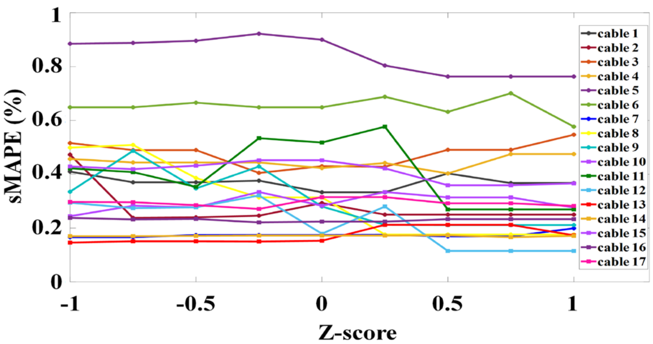

Figure 8: All cable's error of prediction means based on Z-scores.

4. CONCLUSIONS

This study proposes a new data-driven method to probabilistically predict the structural response of a cable bridge based on structural health monitoring data obtained from various sensors. The proposed method used a feature selection technique based on the Z-score to improve the accuracy of prediction models. In addition, to consider the uncertainty of measurement data, the proposed method introduces the GPR. Finally, it was applied to an application example of an actual cable-stayed bridge in the Republic of Korea, and it was demonstrated that the proposed method could provide rational prediction results.

5. ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2022-00144434). This work is also supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 23RMPP-C163162-03).

6. REFERENCES

Cai, J., Luo, J., Wang, S., and Yang, S. (2018). "Feature selection in machine learning: A new perspective." *Neurocomputing*, 300, 70-79

Hyndman, R. J., & Koehler, A. B. (2006). "Another look at measures of forecast accuracy." *International journal of forecasting*, *22*(4), 679-688.

Jang, S., Jo, H., Cho, S., Mechitov, K., Rice, J. A., Sim, S. H., ... & Agha, G. (2010). "Structural health monitoring of a cable-stayed bridge using smart sensor technology: deployment and evaluation." *Smart Structures and Systems*, *6*(5_6), 439-459.

Saunders, C., Grobelnik, M., Gunn, S., and Shawe-Taylor, J. (Eds.). (2006). Subspace, Latent Structure and Feature Selection: Statistical and Optimization Perspectives Workshop, SLSFS 2005 Bohinj, Slovenia, February 23-25, 2005, Revised Selected Papers (Vol. 3940), Springer, Berlin, Heidelberg, 1-207.

Sim, S. H., Li, J., Jo, H., Park, J. W., Cho, S., Spencer Jr, B. F., & Jung, H. J. (2013). "A wireless smart sensor network for automated monitoring of cable tension." *Smart Materials and Structures*, 23(2), 025006.

Lee, J., Lee, K. C., & Lee, Y. -J. (2018). "Long-term deflection prediction from computer vision-measured data history for high-speed railway bridges." *Sensors*, 18(5), 1488

Lee, J., Lee, K. C., Sim, S. H., Lee, J., & Lee, Y. -J. (2019). "Bayesian Prediction of Pre-Stressed Concrete Bridge Deflection Using Finite Element Analysis." *Sensors*, 19(22), 4956.

Park, J. W., Sim, S. H., & Jung, H. J. (2014). "Wireless displacement sensing system for bridges using multi-sensor fusion." *Smart Materials and Structures*, *23*(4), 045022.

Rasmussen, C. E. (2003, February). Gaussian processes in machine learning. In Summer school on machine learning, Springer, Berlin, Heidelberg, 63-71.