13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13
Seoul, South Korea, May 26-30, 2019

Probabilistic Prediction of Vertical Deflection for High-speed Railway Bridges Using a Gaussian Process

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

Kyoung-Chan Lee
Senior Researcher, High-Speed Railroad Systems Research Center, Korea Railroad Research Institute, Uiwang, 16104, Republic of Korea

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

ABSTRACT: Vertical deflection of a high-speed railway bridge is one of the important indicators for managing the safety and running stability of a vehicle. Therefore, efforts have been made to develop sensors for measuring the deflection and predicting its short- and long-term future values. However, the vertical deflection of a railway bridge is stochastic because it involves various sources of uncertainty, which may cause errors in physics-based prediction models. This study proposes a Bayesian approach to build a probabilistic prediction model for the vertical deflection of a railway bridge. For this task, a Gaussian process is introduced to construct a covariance matrix with multiple kernels. Thereafter, actual vision-based measurements, measuring time, and temperature data are used to optimize the hyperparameters of the kernels. As a result, the proposed approach provides a probabilistic prediction interval as well as a predictive mean of the vertical deflections of the bridge. This approach is applied to an actual high-speed railway bridge in the Republic of Korea, and the corresponding analysis results and their performance are discussed.

1. INTRODUCTION

A railway bridge is an important infrastructure for various types of trains, such as a high-speed train, inter-city train, and monorail. In several cases, such trains are used to carry a large amount of freight and numerous passengers; hence, various studies have been carried out to ensure the safety and running serviceability of railway bridges.

One of the important indicators used for inspecting the functionalities of railway bridges is a vertical deflection, and it is especially known to be critical factor for high-speed trains (Guo et al., 2012). For the sake of monitoring the vertical deflection, various researches have been conducted in terms of sensing and predicting its values. Firstly, numerous sensing techniques have been developed: linear variable differential transformers (LVDTs), accelerometer, global positioning systems (GPSs), laser Doppler vibrometers (LDVs), radio detection and ranging (RADAR), and vision-based systems. However, these tend to have measurement errors owing to several factors (Lee et al., 2017).

With regards to predicting, research has mostly been focused on the derivation of physics-based models. For example, the deflection of a prestressed concrete girder which is often used for railway bridges has been predicted in terms of various factors, such as temperature changes, concrete creep and shrinkage, and train loads (Guo et al., 2010). Although these predictions are in agreement with

the actual measurement data, building a well-fitted prediction model for bridges with no prediction error is not an easy task. However, most of the previous studies did not take into account this uncertainty and could not suggest a probabilistic prediction (Beltempo et al., 2018).

As both sensing and predicting the vertical deflection have measurement and prediction errors, it is important to consider these uncertainties using probabilistic concepts. Therefore, this study proposes a probabilistic prediction method for the vertical deflection of a railway bridge employing a Gaussian Process (GP). The vision-based measurement of deflection and temperature data, which were obtained using the techniques of Kim and Kim (2014), are used for training the prediction model. The GP is modeled with multiple kernels and hyperparameters to identify the probabilistic property of time-variant vertical deflection. Once the hyperparameters are optimized with a maximum likelihood concept using the training sensor data, the suggested method provides the prediction intervals (e.g., 95%, 99%) and the predictive mean of the vertical deflection of the railway bridge.

2. PROPOSED METHOD

2.1. Gaussian Process Regression (GPR)

GPR is one of the machine learning-based methods to construct a flexible Bayesian model (Barber, 2012). It aims to build a probabilistic prediction model based on the assumption that all data points are multivariate normally distributed. In other words, any sub-vector of the multivariate data is again a Gaussian random vector (Gubner, 2006).

Let a noisy training dataset $\mathbf{D}$ be from N_D times of measurement, which consists of a training input matrix $\mathbf{X}$ and a training output vector $\mathbf{y}$. Then, the dataset can be expressed as follows:

$$\mathbf{D}=\{(\mathbf{X},\mathbf{y})\}=\left\{(x_{ij},y_i)\middle| i=1,...,N_D; j=1,...,N_x\right\} \tag{1}$$

where x_{ij} is an element of the training input matrix $\mathbf{X}$, which is constructed from N_D times of observation for N_x variables, and y_i is an element in the training output vector $\mathbf{y}$ whose size is N_D by 1. When a test input matrix $\mathbf{X}_*$ is introduced, the multivariate normality assumption is utilized to determine an optimal corresponding output vector $\hat{f}_*(\cdot)$:

$$\begin{aligned}\begin{pmatrix}\mathbf{y}\\ \hat{f}_*\left(\mathbf{X}_*,\hat{\boldsymbol{\theta}}\right)\end{pmatrix} &= N\left(\mathbf{O},\boldsymbol{\Sigma}\right)\\ &= N\left(\mathbf{O},\begin{bmatrix}\mathbf{K}+\sigma_{noise}^2\mathbf{I} & \mathbf{K}_*\\ \mathbf{K}_*^{\mathrm{T}} & \mathbf{K}_{**}+\sigma_{noise}^2\mathbf{I}\end{bmatrix}\right)\end{aligned} \tag{2}$$

where N denotes the multivariate Gaussian distribution, $\mathbf{O}$ is the zero matrix, which is often introduced as a prior mean function for numerical simplicity (Barber, 2012; Rasmussen and Williams, 2006), $\sum$ is the symmetric and positive semidefinite covariance matrix, $\mathbf{K}$ is the covariance matrix of the training input matrix $\mathbf{X}$, $\mathbf{K}_*$ is the covariance matrix between training ($\mathbf{X}$) and test ($\mathbf{X}_*$) inputs, $\mathbf{K}_{**}$ is the covariance matrix of the test input matrix $\mathbf{X}_*$, and σ_{noise}^2 is the variance of noise.

Once a covariance matrix is built for all inputs (e.g., training and test inputs) as Equation (2), the optimal prediction output vector for the given test inputs can be estimated by the property of the multivariate Gaussian distribution, as summarized in Equation (3) (Likar and Kocijan, 2007):

$$\begin{aligned}&\hat{f}_*(\cdot)\middle|\mathbf{y},\mathbf{X},\mathbf{X}_* \sim\\ &N\begin{pmatrix}\mathbf{K}_*^T\left(\mathbf{K}+\sigma_{noise}^2\mathbf{I}\right)^{-1}\mathbf{y},\\ \mathbf{K}_{**}-\mathbf{K}_*^T\left(\mathbf{K}+\sigma_{noise}^2\mathbf{I}\right)^{-1}\mathbf{K}_*+\sigma_{noise}^2\mathbf{I}\end{pmatrix}\end{aligned} \tag{3}$$

Meanwhile, in GP regression, the covariance matrix is modeled by a combination of multiple kernels. A possible choice of kernel considering correlation is the squared exponential (SE) kernel as follows:

$$k_{SE}(\mathbf{x}_a,\mathbf{x}_b) = \sigma_f^2 \exp\left(-\frac{1}{2}\frac{\sqrt{\sum_{i=1}^{N_x}\left[\mathbf{x}_a(1,i)-\mathbf{x}_b(1,i)\right]^2}}{l^2}\right) \quad (4)$$

where σ_f^2 is the variance hyperparameter of inputs which controls the vertical scale of the function change and l is a length-scale hyperparameter that is associated with the horizontal scale of the function change (Murphy, 2014).

However, the covariance matrix $\sum$ often requires additional variance terms for its diagonal terms (i.e., $\sigma_{noise}^2\mathbf{I}$). For this reason, the Kronecker delta function can be introduced.

$$k_{var}(\mathbf{x}_a,\mathbf{x}_b) = \sigma_f^2 \cdot \delta(\mathbf{x}_a,\mathbf{x}_b) \quad (5)$$

In this study, a new kernel was proposed by summing up existing *kernels* without any loss of properties as a single kernel (Rasmussen and Williams, 2006):

$$k(\mathbf{x}_a,\mathbf{x}_b) = \hat{\theta}_0 \exp\left(-\frac{1}{2}\frac{\sqrt{\sum_{i=1}^{N_x}\left[\mathbf{x}_a(1,i)-\mathbf{x}_b(1,i)\right]^2}}{\hat{\theta}_1^2}\right) + \hat{\theta}_2 \cdot \delta(\mathbf{x}_a,\mathbf{x}_b) + \hat{\theta}_3 \quad (6)$$

where $\hat{\theta}_0$ and $\hat{\theta}_1$ are hyperparameters for non-diagonal terms, $\hat{\theta}_2$ is a hyperparameter to control the variance in diagonal terms, and $\hat{\theta}_3$ is a hyperparameter to control the level of the overall values in the covariance matrix. The kernel in Equation (6) has been applied in previous studies (Petelin et al., 2013; Likar and Kocijan, 2007), and it is introduced in this study as well.

Optimizing these hyperparameters in the GPR is a very important task to obtain an accurate prediction model. For optimizing, the concept of maximum likelihood is often introduced (Rasmussen, 2003). Regarding Equation (3), the likelihood of observing the training output vector **y** given the training input **X** can be expressed as a conditional probability using the multivariate normal distribution as follows:

$$p(\mathbf{y}|\boldsymbol{X}) = \frac{1}{(2\pi)^{\frac{N_D}{2}} \cdot |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}\right) \quad (7)$$

For numerical convenience, natural logarithm is introduced for conditional probability and is multiplied by minus one:

$$L = -\ln p(\mathbf{y}|\boldsymbol{X}) = \frac{1}{2}\mathbf{y}^T\left(\mathbf{K}+\sigma_{noise}^2\mathbf{I}\right)^{-1}\mathbf{y} + \frac{1}{2}\ln\left|\mathbf{K}+\sigma_{noise}^2\mathbf{I}\right| + \frac{N_D}{2}\ln(2\pi) \quad (8)$$

where L is the log-likelihood function. Then, the best hyperparameters $\hat{\boldsymbol{\theta}}_{best}$ can be determined through optimization to minimize it:

$$\hat{\boldsymbol{\theta}}_{best} = \arg\min_{\hat{\boldsymbol{\theta}}}(L) \quad (9)$$

2.2. Performance Assessment Measures

There are two assessment aspects for the GPR results; the predictive mean and prediction interval (PI). In terms of the predictive mean, a popular prediction error index is the root-mean-square error (*RMSE*):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_p}(\hat{f}(\mathbf{x}_i) - y(\mathbf{x}_i))^2}{N_p}} \quad (10)$$

where $\hat{f}(\mathbf{x}_i)$ is the predictive mean at the test input $\mathbf{x}_i$, and $y(x_i)$ is the actual measurement. However, if the predictive mean follows the mean trend of the fluctuating dataset, RMSE might not be appropriate. Rather, the mean-error (*ME*) is widely used as a good alternative indicator (Landberg, 2001):

$$ME = \frac{\sum_{i=1}^{N_p}(\hat{f}(\mathbf{x}_i) - y(\mathbf{x}_i))}{N_p} \quad (11)$$

On the other hand, for the performance assessment of a PI, its coverage probability is often assessed (Khosravi et al., 2011). A popular choice is the PI coverage probability (*PICP*):

$$PICP = \frac{1}{N_p}\sum_{i=1}^{N_p} C_i \tag{12}$$

where C_i is the Boolean value which can be evaluated as:

$$C_i = \begin{cases} 1, & y_i \in [L_i, U_i] \\ 0, & y_i \notin [L_i, U_i] \end{cases} \tag{13}$$

where y_i is the i^{th} measurement data, L_i is the lower bound, and U_i is the upper bound of the PI. As the PI is built with the nominal confidence of (1-α) %, which is known as the PI nominal confidence (*PINC*), a good PI will have a similar *PICP* with the given confidence level. Therefore, the average coverage error (*ACE*) can be used with *PICP* and *PINC* as a performance measure of the PI (Shrivastava et al., 2015):

$$ACE = PINC - PICP \tag{14}$$

In this study, *RMSE* and *ME* are introduced to assess the performance of the predictive mean, and *ACE* is used for the PI. The closer these values are to zero, the greater is the prediction model.

3. PROBLEM DESCRIPTION

Eonyang Arch Bridge, which is a railway bridge for high-speed trains in Ulsan, Republic of Korea, was selected as an application example. It was built in 2009, and has two neighbored twin bridges for north- and southbound trains as shown in Figure 1. Among these, the northbound bridge is the target structure of this study.

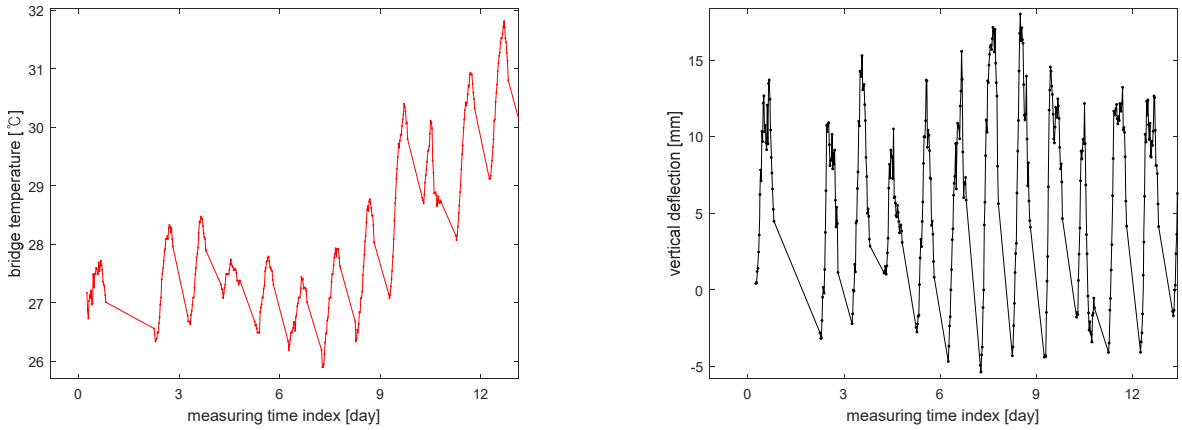
Figure 1: Example bridge: Eonyang Arch Bridge

Two measurement data were obtained: the vertical deflection of the center of the mid-span, and the bridge temperature. The vertical deflection was measured by the vision-based system using cameras installed approximately 100 m from the bridge. The temperature was measure by the resistance temperature detector (RTD) that was installed on the inner side of the center span to avoid direct exposure to sunlight (Lee et al., 2018). Using these sensors, the vertical deflection and bridge temperature were measured once every 30 min for 4.5 months, from July 15 to November 27, 2016, as shown in Table 1.

Table 1: Sample measurement specifications.

Measuring time	*Bridge Temperature [°C]*	*Vertical deflection [mm]*
2016.07.15, 6:00	*21.45*	*1.195*
2016.07.15, 6:30	*23.59*	*2.455*
2016.07.15, 7:00	*25.73*	*3.565*
…	…	…
2016.11.27, 16:00	*9.55*	*-6.127*
2016.11.27, 16:30	*9.26*	*-6.638*
2016.11.27, 17:00	*8.69*	*-7.300*

Here, the positive and negative signs of the vertical deflection denote upward and downward deflection, respectively. A total of 2,292 datasets for 4.5 months were recorded, and it is observed that both the bridge temperature and vertical deflection fluctuate cyclically within one day, as shown in Figure 2.



(a) *(b)*

Figure 2: Fluctuating property of measurement data: (a) bridge temperature and (b) vertical deflection

Figure 3 shows that the bridge temperature and vertical deflection of the example bridge are correlated to each other. According to Nilson (2003), a creep proceeds at a decreasing rate for the first several years, and the American Concrete Institute (2008) states that a major part of the shrinkage effect is manifested in the first year. The example bridge was built in 2009 and the measurement data were obtained in 2016; thus, the bridge temperature can be assumed to have a more significant effect than the creep and shrinkage.

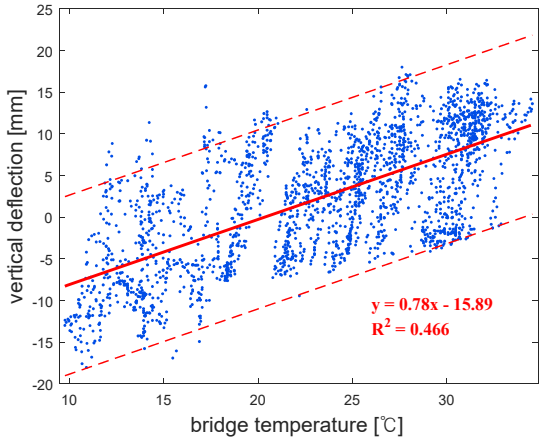


Figure 3: Bridge temperature versus vertical deflection

The example bridge is built for high-speed trains, the Korea Train eXpress (KTX). Therefore, the fluctuating vertical deflection has to be managed for traffic safety and passenger comfort. Indeed, an evaluation presents that the bridge can adversely affect passenger comfort in hot weather exceeding 40 ℃ (Lee et al., 2018). For this reason, the probabilistic prediction of vertical deflection is conducted in this study for its management.

4. ANALYSIS RESULTS

The probabilistic vertical deflection of the example bridge was predicted by utilizing the proposed method. Input (**X**) and output (**y**) training data were designed as the measurements including temperature ($\mathbf{x}_1$), deflection (**y**), and corresponding measuring time ($\mathbf{x}_2$). After optimizing the prediction model using the measured training data, the interested time range ($\mathbf{x}_{2*}$) and corresponding temperature ($\mathbf{x}_{1*}$) were given for obtaining the predictive vertical deflection ($\hat{f}_*(\cdot)$).

Figure 4 shows the results of predictive mean and 95% PI. As shown in the left figure, the actual vertical deflection fluctuates on a daily basis, and the daily fluctuating ranges are approximately up to ±15 mm from the average. This is mainly because the vertical deflection is correlated with the temperature which also changes on a daily basis. In addition, in the figure on the right, another cycle with a period of hundreds of days can be found, which is thought to be due to the seasonal changes in the temperature.

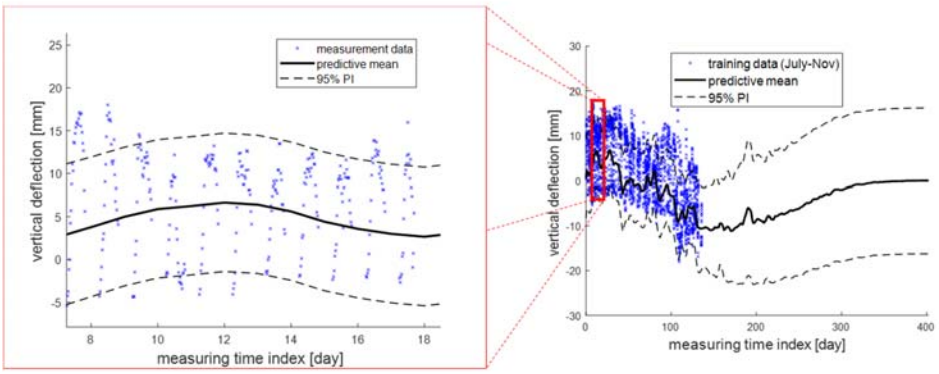


Figure 4: Analysis results of predictive mean and 95% PI from the entire measurement data (right) and its zoom-in view (left)

Meanwhile, the probabilistic prediction model can be updated when additional data is given. For example, if the measurement data is available only from July to August, then the prediction model is constructed as shown in Figure 5. However, when the data of September is given as additional training data, then the prediction model can be updated as shown in Figure 6. Similarly, if the data measured from July to October is used, the probabilistic prediction model is built as shown in Figure 7.

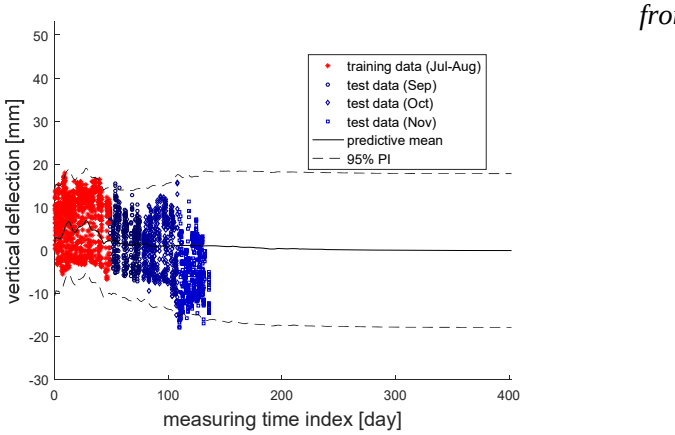


Figure 6: Prediction model based on the datasets from Jul. to Sep.

Figure 5: Prediction model based on the datasets from Jul. to Aug.

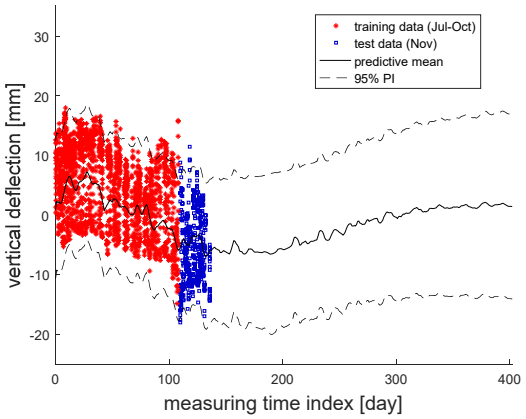


Figure 7: Prediction model based on the datasets from Jul. to Oct.

The prediction models in Figures 5–7 were checked by estimating various performance assessment indexes. First, the *RMSE*s with respect to the datasets of each month were

calculated to check the performance of the predictive mean. Table 2 is composed of two parts: upper and lower part, for index values with respect to the training and test dataset, respectively. The upper part of Table 2 shows that the *RMSE* values are estimated to be approximately 5 mm overall, which is the inherent variability of the vertical deflection to the predictive mean. These *RMSE* values that are approximately 5 mm shows that the constructed models based on different durations of measurement time have a similar level of inherent variability. Likewise, the prediction models have similar *RMSE* values to the test datasets with those of training datasets, summarized in the lower part of Table 2.

Table 2: Root-mean-square error (RMSE) values.

	Model of Jul.-Aug.	*Model of Jul.-Sep.*	*Model of Jul.-Oct.*
Jul.	*4.09 mm*	*4.60 mm*	*4.28 mm*
Aug.	*5.43 mm*	*5.58 mm*	*5.39 mm*
Sep.	*3.98 mm*	*4.78 mm*	*4.24 mm*
Oct.	*5.27 mm*	*4.29 mm*	*5.89 mm*
Nov.	*7.06 mm*	*3.35 mm*	*3.42 mm*

The performance of the predictive mean can be confirmed clearly in Table 3, which shows the results of *ME*. The *ME* index values with respect to the training data are overall close to zero. It is because the inherent variabilities of datasets are canceled out, which means that the prediction models are close to being unbiased. The model, based on the measurement data from July to August, may not predict well the data of November. However, whenever an additional dataset is added to the training data, the absolute *ME* value generally decreases, which means the accuracy of the prediction model improves. For example, the *ME* value of the prediction model constructed using the data from July to October has a small *ME* value of 0.49 mm, when the model is compared with the actual measurement data for November.

Table 3: Mean-error (ME) values.

	Model of Jul.-Aug.	*Model of Jul.-Sep.*	*Model of Jul.-Oct.*
Jul.	*-1.00 mm*	*-1.04 mm*	*-1.09 mm*
Aug.	*0.29 mm*	*-0.67 mm*	*0.56 mm*
Sep.	*1.73 mm*	*-0.82 mm*	*0.59 mm*
Oct.	*0.37 mm*	*-1.47 mm*	*-1.40 mm*
Nov.	*6.22 mm*	*-1.66 mm*	*0.49 mm*

On the other hand, the *ACE* values are calculated to check the performance of the PIs. Table 4 shows the *ACE* values, and it can be seen that the *ACE* values are close to zero throughout. This implies that the PI of each prediction model covers the specified portion (i.e., 95% in this example) of the training and test data. It is also observed that the absolute values of *ACE* decrease when measurement data for a longer duration are used.

Table 4: Average coverage error (ACE) values.

	Model of Jul.-Aug.	*Model of Jul.-Sep.*	*Model of Jul.-Oct.*
Jul.	*0.0170*	*-0.0165*	*0.0125*
Aug.	*-0.0045*	*0.0285*	*-0.0293*
Sep.	*-0.0258*	*0.0104*	*-0.0379*
Oct.	*-0.0397*	*0.0460*	*0.0474*
Nov.	*0.0482*	*-0.0478*	*-0.0284*

5. CONCLUSIONS

This study proposes a probabilistic prediction method of the vertical deflection of a railway bridge based on actual computer vision-measured data. A Gaussian process is employed by modeling multiple kernels and hyperparameters. Once the hyperparameters are optimized using training data, the suggested method builds the predictive mean and 95% PI. This method was applied to an existing railway bridge in the Republic of Korea, and the corresponding analysis results showed that both the predictive mean and PI can be estimated. It further showed that the prediction model can be updated with additional data. The built prediction models agreed with the actual measurement data, and it is expected that the prediction results can

be utilized for decision-making on railway bridge maintenance.

6. ACKNOWLEDGEMENTS

This research was supported by a grant (19SCIP-B138406-04) from Smart Civil Infrastructure Research Program funded by Ministry of Land, Infrastructure and Transport (MOLIT) of Korea government and Korea Agency for Infrastructure Technology Advancement (KAIA). This research was also supported by a grant from R&D Program of the Korean Railroad Research Institute, Republic of Korea.

7. REFERENCES

Barber, D. (2012). Bayesian Reasoning and Machine Learning, Cambridge University Press: Cambridge, UK.

Beltempo, A., Bursi, O. S., Cappello, C., Zonta, D., & Zingales, M. (2018). "A Viscoelastic Model for the Long - Term Deflection of Segmental Prestressed Box Girders." *Computer - Aided Civil and Infrastructure Engineering*, 33(1), 64-78.

Committee ACI. (2008). *Guide for Modeling and Calculating Shrinkage and Creep in Hardened Concrete*, ACI Committee, Farmington Hills, MI, USA.

di Sciascio, F., & Amicarelli, A. N. (2008). "Biomass estimation in batch biotechnological processes by Bayesian Gaussian process regression." *Computers & Chemical Engineering*, 32(12), 3264-3273.

Gubner, J.A. (2006). *Probability and Random Processes for Electrical and Computer Engineers*, Cambridge University Press, Cambridge, MA, USA.

Guo, T., Liu, T., & Li, A. (2012). "Deflection reliability analysis of PSC box-girder bridge under high-speed railway loads." *Advances in Structural Engineering*, 15(11), 2001-2011.

Guo, T., Sause, R., Frangopol, D. M., & Li, A. (2010). "Time-dependent reliability of PSC box-girder bridge considering creep, shrinkage, and corrosion." *Journal of Bridge Engineering*, 16(1), 29-43.

Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011). "Comprehensive review of neural network-based prediction intervals and new advances." *IEEE Transactions on neural networks*, 22(9), 1341-1356.

Kim, S. W., & Kim, N. S. (2013). "Dynamic characteristics of suspension bridge hanger cables using digital image processing." *NDT & E International*, 59, 25-33.

Landberg, L. (2001). "Short-term prediction of local wind conditions." *Journal of Wind Engineering and Industrial Aerodynamics*, 89(3-4), 235-245.

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., Cho, S., & Sim, S. H. (2017). "Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges." *Sensors*, 17(10), 2317.

Likar, B., & Kocijan, J. (2007). "Predictive control of a gas–liquid separation plant based on a Gaussian process model." *Computers & chemical engineering*, 31(3), 142-152.

Muirhead, R.J. (2009). *Aspects of Multivariate Statistical Theory*, John Wiley & Sons: Hoboken, NJ, USA.

Murphy, K.P. (2014). *Machine Learning, A Probabilistic Perspective*, The MIT Press: Cambridge, MA, USA.

Nilson, A. H., Darwin, D., & Dolan, C. W. (2004). *Design of concrete structures 13th ed.* McGraw-Hill Higher Education.

Petelin, D., Grancharova, A., & Kocijan, J. (2013). "Evolving Gaussian process models for prediction of ozone concentration in the air." *Simulation modelling practice and theory*, 33, 68-80.

Rasmussen, B., & HINES, J. W. (2004). "Prediction interval estimation techniques for empirical modeling strategies and their applications to signal validation tasks." *In Applied Computational Intelligence* (pp. 549-556).

Rasmussen, C.E., and Williams, C.K. (2006). *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA, USA.

Shrivastava, N. A., Khosravi, A., & Panigrahi, B. K. (2015). "Prediction interval estimation of electricity prices using PSO-tuned support vector machines." *IEEE Transactions on Industrial Informatics*, 11(2), 322-331.