Bayesian Prediction of Pre-Stressed Concrete Bridge Deflection Using Finite Element Analysis

Jaebeom Lee [1], Kyoung-Chan Lee [2],*, Sung-Han Sim [3], Junhwa Lee [1] and Young-Joo Lee [1],*

[1]   School of Urban and Environmental Engineering, Ulsan National Institute of Science and
     Technology (UNIST), Ulsan 44919, Korea; jblee@unist.ac.kr (J.L.); lee.junhwa@unist.ac.kr (J.L.)
[2]   Advanced Railroad Civil Engineering Division, Korea Railroad Research Institute, Uiwang 16105, Korea
[3]   School of Civil, Architectural Engineering and Landscape Architecture, Sungkyunkwan University, Seoul
     16419, Korea; ssim@skku.edu
*   Correspondence: kclee@krri.re.kr (K.-C.L.); ylee@unist.ac.kr (Y.-J.L.)

Abstract: Vertical deflection has been emphasized as an important safety indicator in the management of railway bridges. Therefore, various standards and studies have suggested physics-based models for predicting the time-dependent deflection of railway bridges. However, these approaches may be limited by model errors caused by uncertainties in various factors, such as material properties, creep coefficient, and temperature. This study proposes a new Bayesian method that employs both a finite element model and actual measurement data. To overcome the limitations of an imperfect finite element model and a shortage of data, Gaussian process regression is introduced and modified to consider both, the finite element analysis results and actual measurement data. In addition, the probabilistic prediction model can be updated whenever additional measurement data is available. In this manner, a probabilistic prediction model, that is customized to a target bridge, can be obtained. The proposed method is applied to a pre-stressed concrete railway bridge in the construction stage in the Republic of Korea, as an example of a bridge for which accurate time-dependent deflection is difficult to predict, and measurement data are insufficient. Probabilistic prediction models are successfully derived by applying the proposed method, and the corresponding prediction results agree with the actual measurements, even though the bridge experienced large downward deflections during the construction stage. In addition, the practical uses of the prediction models are discussed.

Keywords: railway bridge; vertical deflection; probabilistic prediction; Gaussian process regression; finite element

1. Introduction

The vertical deflection of railway bridges is an important safety indicator that is used in the safety management of structures and trains [1,2]. In particular, high-speed trains can be derailed owing to railway defects that can be caused by excessive bridge deflection. Therefore, the monitoring and management of the vertical deflection of railway bridges have been emphasized in several standards, such as the Design Guide for Steel Railway Bridges of the UK (2004), UIC CODE 518 OR (2009), and the Guideline of Track Maintenance of Korea Rail Network Authority (2016), which suggest monitoring for the acceptable vertical deflection at mid-span [3–5]. For example, the Korean railway bridge guideline specifies allowable maximum deflections with respect to the span length (*L*) and train velocity. For instance, for train speeds between 200 km/h and 270 km/h, the allowable threshold of vertical deflection with a span length is shorter than 20 m is L/1300 m [5].

Pre-stressed concrete (PSC) girders have been widely adopted as superstructures for high-speed railway bridges; the long-term deflection of bridges with PSC girders is known to be dependent on

the mechanical responses of concrete, such as creep and shrinkage [6]. Concrete creep and drying shrinkage are time-dependent deformations, caused by sustained loading (e.g., dead load), and loss of moisture, respectively, and these deformations proceed with a decreasing rate for the first several months or years [7]. Considering these factors, a variety of prediction models, such as the CEB-FIP, ACI, KCI, and B3 models, have been developed [8–11]. However, these generalized prediction models are limited by model errors caused by uncertainties in various factors, such as material properties, creep coefficient, and temperature [12]. Kamatchi et al. (2014) showed that prediction models (i.e., CEB-FIP, ACI, KCI, and B3 models) also involve errors with field measurements [13], and Bažant et al. (2012) revealed that existing material models for predicting creep and shrinkage are unsatisfactory [14,15]. In addition, the deflection prediction of a bridge under construction is more challenging, because the structure experiences sudden loading and unloading during the construction stage. Therefore, it is important to consider the physical phenomena associated with these construction events from an early stage to predict the long-term deflection of a bridge under construction accurately.

Meanwhile, data-driven approaches, which utilize actual measurement data, have been studied to construct prediction models as well. Sun and Hao (2011) introduced a time series analysis method utilizing monitoring data and established a prediction model of bridge deflection [16]. Similarly, Xin et al. (2018) suggested a prediction method for bridge deformation, integrating the Kalman filter, the autoregressive integrated moving average model, and a generalized autoregressive conditional heteroskedasticity process, using various measurement data (i.e., from strain gauges, displacement sensors at mid-span, acceleration sensors, cable meters, and temperature sensors) [17]. A wavelet analysis is introduced in various studies to detect the structural damage [18–23], and Lee et al. (2018) introduced a machine-learning-based method to construct a non-parametric prediction model, which showed good agreement with actual measurement data [2].

However, because the amount of measurement data may be insufficient to apply a data-driven approach, a few studies have combined both physics-based models and data-driven approaches. Yang (2007) proposed a method to incorporate measurement data into physics-based prediction models, considering creep and shrinkage to reduce the uncertainty in long-term predictions [24]. Similarly, Bažant et al. (2012) used deflection measurements to update the B3 model [14]. A finite element (FE) model calibration procedure using the monitoring data has often been conducted by researchers studying efficient bridge management [25–30].

Inaccurate prediction of bridge deflection may result in unexpected structural failure (e.g., the Koror-Babeldaob bridge failure in 1996); thus, continuous monitoring and management of vertical deflection are important. This study aims to overcome the limitations of the structural model error and data shortage and provide a probabilistic prediction of bridge deflection. To this end, a new Bayesian method is proposed, utilizing an FE model constructed with physics-based knowledge and actual measurement data. The target bridge of this study is a PSC railway bridge for high-speed trains in the construction stage, to which various construction loads are applied. Therefore, an accurate prediction of time-dependent deflection is difficult to obtain, even using a sophisticated FE analysis tool. Moreover, a sufficient quantity of measurement data for predicting the bridge deflection is unavailable because of the short measurement period, and thus, a data-driven approach may be inaccurate. An FE model is constructed by considering the construction schedule and mechanical responses for predicting vertical deflection as prior knowledge. Then, the prediction model, based on physics-based knowledge, is updated by measurement data utilizing Gaussian process regression (GPR), which is a nonparametric Bayesian method. In this manner, the probabilistic interval (e.g., 95% prediction interval) of the expected vertical deflection can be obtained, which can be used for bridge deflection management in the construction stage.

2. Gaussian Process Regression (GPR)

2.1. Bayesian Inference in GPR

To build a probabilistic prediction model, GPR is used in this study. A Gaussian process (GP) is a stochastic process, by which all data points are assumed to follow a multivariate Gaussian distribution,

and implying that the data points are correlated and normally distributed [31]. In GPR, data points are assumed to originate from a GP, called the GP assumption. It is also known that GPR is a Bayesian method that builds a nonparametric regression model using the GP assumption [32].

Let N_D quantities of noisy training dataset **D**, which consists of a training input matrix **X** and a training output vector **y**, be given as,

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

where x_{ij} is the element of the training input matrix **X**, whose matrix size is N_D by N_x, and y_i is the element in the training output vector **y**, whose size is N_D by 1. The input-output relationship is expressed as,

$$\mathbf{y} = f(\mathbf{X}) + \varepsilon = \phi(\mathbf{X})\mathbf{w} + \varepsilon \tag{2}$$

where $\phi(\cdot)$ is the vector of N_{dim} projection functions (e.g., $\phi(z) = [1, z, z^2, z^3]$ for the three-dimensional polynomial regression of a scalar input z or $\phi(z) = [1, \sin(2\pi z), \cos(2\pi z)]$ for the first-order trigonometric regression), $f(\mathbf{X})$ is the regression function of **X**, which can be calculated as the product of the projected input matrix $\phi(\mathbf{X})$, and the vector of weights **w** (whose sizes are N_D by N_{dim} and N_{dim} by 1, respectively), and ε is the vector of observational errors (whose vector size is N_D by 1). Here, the observational errors are assumed to follow an independent and identically distributed Gaussian distribution with a zero-mean and constant variance of the noise, $\sigma_n{}^2$ for all inputs (i.e., a homoscedastic noise assumption) as:

$$\varepsilon \sim N\left(\mathbf{O}, \sigma_n^2\mathbf{I}\right). \tag{3}$$

To infer the weight vector **w** based on actual measurements in the Bayesian view, the prior assumption on the weight vector is set to be a zero-mean Gaussian with the covariance matrix $\mathbf{\Sigma}_p$:

$$\mathbf{w} \sim N\left(\mathbf{O}, \mathbf{\Sigma}_p\right) \tag{4}$$

which enables the prior prediction on the mean $\mathbf{\mu_{X*}}$ and the standard deviation $\mathbf{\sigma_{X*}}$ based on the given input matrix $\mathbf{X}_*$ as:

$$f(\mathbf{X}_*) = \phi(\mathbf{X}_*)\mathbf{w} \sim N\left(\mathbf{O}, \phi(\mathbf{X}_*)\mathbf{\Sigma}_p\phi(\mathbf{X}_*)^T\right) = N\left(\mathbf{\mu_{X_*}}, \sigma_{\mathbf{X}_*}^2\right). \tag{5}$$

From the independent and identically distributed Gaussian process assumption of observations having homoscedastic noise, the likelihood function can be constructed as the multiplication of normal probability density functions (PDFs) as follows,

$$\begin{aligned}
p(\mathbf{y}|\mathbf{X}, \mathbf{w}) &= \prod_{i=1}^{N_D} P(y_i|\phi(\mathbf{x}_i)\mathbf{w}) = \prod_{i=1}^{N_D} \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{(y_i - \phi(\mathbf{x}_i)\mathbf{w})^2}{2\sigma_n^2}\right) \\
&= \frac{1}{\left(2\pi\sigma_n^2\right)^{\frac{N_D}{2}}} \exp\left(-\frac{1}{2\sigma_n^2}\|\mathbf{y} - \phi(\mathbf{X})\mathbf{w}\|_2^2\right)
\end{aligned} \tag{6}$$

where y_i is an i-th element of the output vector, $\mathbf{x}_i$ is the corresponding input vector, $\|\cdot\|_2$ is an L2 norm which denotes the Euclidean distance. Consequently, Equation (6) indicates the following relation:

$$\mathbf{y}|\mathbf{X}, \mathbf{w} \sim N\left(\phi(\mathbf{X})\mathbf{w}, \sigma_n^2\mathbf{I}\right). \tag{7}$$

Meanwhile, from the Bayes' rule, the posterior probability on the weight vector is expressed as:

$$\begin{aligned}
p\left(\mathbf{w}|\mathbf{X}, \mathbf{y}\right) &= \frac{p(\mathbf{X}, \mathbf{y}|\mathbf{w})p(\mathbf{w})}{p(\mathbf{X}, \mathbf{y})} = \frac{p(\mathbf{y}|\mathbf{X}, \mathbf{w})p(\mathbf{X}, \mathbf{y})}{p(\mathbf{y}|\mathbf{X})} \frac{p(\mathbf{w})}{p(\mathbf{X}, \mathbf{y})} \\
&= \frac{p(\mathbf{y}|\mathbf{X}, \mathbf{w})p(\mathbf{w})}{p(\mathbf{y}|\mathbf{X})} = \frac{p(\mathbf{y}|\mathbf{X}, \mathbf{w})p(\mathbf{w})}{\int p(\mathbf{y}|\mathbf{X}, \mathbf{w})p(\mathbf{w})dw}
\end{aligned} \tag{8}$$

It can be seen from Equation (8) that the posterior probability $p(\mathbf{w}|\mathbf{X},\mathbf{y})$ is proportional to the prior probability $p(\mathbf{w})$ and the likelihood $p(\mathbf{y}|\mathbf{X},\mathbf{w})$, which can be expressed as follows:

$$p\big(\mathbf{w}|\mathbf{X},\mathbf{y}\big) \propto p(\mathbf{y}|\mathbf{X},\mathbf{w})p(\mathbf{w}). \tag{9}$$

From Equations (4) and (6), the posterior probability in Equation (9) is seen to be proportional to the multiplication of two multi-variate Gaussian PDFs:

$$p\big(\mathbf{w}|\mathbf{X},\mathbf{y}\big) \propto \exp\!\left(-\frac{1}{2\sigma_n^2}(\mathbf{y}-\phi(\mathbf{X})\mathbf{w})^T(\mathbf{y}-\phi(\mathbf{X})\mathbf{w})\right)\exp\!\left(-\frac{1}{2}\mathbf{w}^T\Sigma_p^{-1}\mathbf{w}\right). \tag{10}$$

That is [31],

$$p\big(\mathbf{w}|\mathbf{X},\mathbf{y}\big) \propto \exp\!\left(-\frac{1}{2}\big[\mathbf{w}-\bar{\mathbf{w}}\big]^T\big(\sigma_n^{-2}\phi(\mathbf{X})^T\phi(\mathbf{X})+\Sigma_p^{-1}\big)\big[\mathbf{w}-\bar{\mathbf{w}}\big]\right) \tag{11}$$

where

$$\bar{\mathbf{w}} = \sigma_n^{-2}\mathbf{A}^{-1}\phi(\mathbf{X})^T\mathbf{y} \tag{12}$$

and

$$\mathbf{A} = \sigma_n^{-2}\phi(\mathbf{X})^T\phi(\mathbf{X})+\Sigma_p^{-1}. \tag{13}$$

For the given input matrix $\mathbf{X}_*$, the predictive mean $\phi(\mathbf{X}_*)\cdot\mathbf{w}$ is [31],

$$\sigma_n^{-2}\phi(\mathbf{X}_*)\mathbf{A}^{-1}\phi(\mathbf{X})^T\mathbf{y} = \phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X})^T\big(\phi(\mathbf{X})^T\Sigma_p\phi(\mathbf{X})+\sigma_n^2\mathbf{I}\big)^{-1}\mathbf{y} \tag{14}$$

and the predictive variance is [31],

$$\begin{aligned}
\phi(\mathbf{X}_*)\mathbf{A}^{-1}\phi(\mathbf{X}_*)^T &= \phi(\mathbf{X}_*)\big(\sigma_n^{-2}\phi(\mathbf{X})^T\phi(\mathbf{X})+\Sigma_p^{-1}\big)^{-1}\phi(\mathbf{X}_*)^T \\
&= \phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X}_*)^T - \phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X})^T\big(\sigma_n^2+\phi(\mathbf{X})\Sigma_p\phi(\mathbf{X})^T\big)^{-1}\phi(\mathbf{X})\Sigma_p\phi(\mathbf{X}_*)^T
\end{aligned} \tag{15}$$

which is based on the following matric inversion [33]:

$$\big(UWV^T+Z\big)^{-1} = Z^{-1} - Z^{-1}U\big(W^{-1}+V^TZ^{-1}U\big)^{-1}V^TZ^{-1}. \tag{16}$$

In summary, the predictive mean and variance with the posterior inference on the weights vector are:

$$\begin{aligned}
f(\mathbf{X}_*) &= \phi(\mathbf{X}_*)\mathbf{w} \\
&\sim N\!\left(\begin{array}{c}\phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X})^T\big(\phi(\mathbf{X})\Sigma_p\phi(\mathbf{X})^T+\sigma_n^2\mathbf{I}\big)^{-1}\mathbf{y}, \\ \phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X}_*)^T - \phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X})^T\big(\sigma_n^2+\phi(\mathbf{X})\Sigma_p\phi(\mathbf{X})^T\big)^{-1}\phi(\mathbf{X})\Sigma_p\phi(\mathbf{X}_*)^T\end{array}\right) \\
&= N\big(\mu_{\mathbf{X}_*},\sigma_{\mathbf{X}_*}^2\big)
\end{aligned} \tag{17}$$

Here, the term $\phi(\cdot)\Sigma_p\phi(\cdot)^T$ is the covariance matrix of the projected inputs, and when expressed with the capital letter "$\mathbf{K}$," Equation (17) is arranged as:

$$f(\mathbf{X}_*) \sim N\!\left(\mathbf{K}_*\big(\mathbf{K}+\sigma_n^2\mathbf{I}\big)^{-1}\mathbf{y}, \mathbf{K}_{**}-\mathbf{K}_*\big(\sigma_n^2+\mathbf{K}\big)^{-1}\mathbf{K}_*\right) \tag{18}$$

where $\mathbf{K}=\phi(\mathbf{X})\Sigma_p\phi(\mathbf{X})^T$, $\mathbf{K}_*=\phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X})^T$, and $\mathbf{K}_{**}=\phi(\mathbf{X}_*)\Sigma_p\phi(\mathbf{X}_*)^T$.

2.2. Covariance Matrix Design Using Kernel Functions

In GPR, the covariance matrix of the projected inputs is numerically estimated using kernel functions. That is, the covariance of two input vectors $\mathbf{X}_a$ and $\mathbf{X}_b$ is defined by the multiplication of

two standard deviations and one correlation coefficient, and is expressed with a kernel function k as follows,

$$\sigma_{\mathbf{x}_a}\sigma_{\mathbf{x}_b}\rho_{\mathbf{x}_a\mathbf{x}_b} = k(\mathbf{x}_a, \mathbf{x}_b) \tag{19}$$

where $\sigma_{\mathbf{x}_a}$ and $\sigma_{\mathbf{x}_b}$ are the standard deviations of $\mathbf{X}_a$, and $\mathbf{X}_b$, respectively, and $\rho_{\mathbf{x}_a\mathbf{x}_b}$ is the correlation coefficient of $\mathbf{X}_a$ and $\mathbf{X}_b$. The covariance matrix, based on kernel functions, enables the GP model to fit data even though the data exhibits nonlinearity [31].

If data have the same variance, then the input data are homoscedastic, and the standard deviation must be a constant that is the square root of the homoscedastic variance. Meanwhile, a correlation coefficient between the two stationary data is generally expressed to be inversely proportional to the Euclidean distance for stationary data, and one representative is the squared exponential kernel (k_{SE}) [31],

$$k_{SE}(\mathbf{x}_a, \mathbf{x}_b) = \theta_1^2 \exp\left(-\frac{1}{2}\frac{\|\mathbf{x}_a(1,i), \mathbf{x}_b(1,i)\|_2^2}{\theta_2^2}\right) \tag{20}$$

where $\theta_1{}^2$ is the variance hyperparameter that controls the vertical scale of the function change, and $\theta_2{}^2$ is the length-scale hyperparameter that controls the horizontal scale of the function change [31].

The variance of noise ($\sigma_n{}^2$) in Equation (2) is often expressed by a kernel function (k_{var}) using the Kronecker delta function,

$$k_{var}(\mathbf{x}_a, \mathbf{x}_b) = \theta_3^2 \cdot \delta(\mathbf{x}_a, \mathbf{x}_b) \tag{21}$$

where $\theta_3{}^2$ is the variance hyperparameter for the input noise. In addition, a bias is introduced using a constant kernel (k_c) with the hyperparameter θ_4:

$$k_c(\mathbf{x}_a, \mathbf{x}_b) = \theta_4. \tag{22}$$

In this study, a new kernel is suggested by combining the following kernels:

$$k(\mathbf{x}_a, \mathbf{x}_b) = \theta_1^2 \exp\left(-\frac{1}{2}\frac{\|\mathbf{x}_a(1,i), \mathbf{x}_b(1,i)\|_2^2}{\theta_2^2}\right) + \theta_3^2 \cdot \delta(\mathbf{x}_a, \mathbf{x}_b) + \theta_4. \tag{23}$$

These hyperparameters strongly affect the performance of GPR, and their optimal values are estimated based on the given observation data and updated whenever additional data is available. For this purpose, the marginal likelihood $p(\mathbf{y}|\mathbf{X})$, which means the probability to observe the output $\mathbf{y}$ given the input $\mathbf{X}$, is introduced. The marginal likelihood of the observations with the prior inference on the weights vector is expressed as the probability density function of the multivariate Gaussian distribution as follows:

$$p(\mathbf{y}|X) = \frac{1}{(2\pi)^{\frac{N_D}{2}} \cdot |\mathbf{K} + \sigma_n^2\mathbf{I}|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}\mathbf{y}^T\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{y}\right). \tag{24}$$

Then, the marginal likelihood will be maximized with the best combination of the hyperparameters, which yields an optimization problem. In this study, for numerical convenience, the minus log-marginal likelihood is utilized as the objective function (L):

$$L = -\ln p(\mathbf{y}|X) = \frac{1}{2}\mathbf{y}^T\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{y} + \frac{1}{2}\ln|\mathbf{K} + \sigma_n^2\mathbf{I}| + \frac{N_D}{2}\ln(2\pi). \tag{25}$$

The best hyperparameters (θ_{best}) can be determined through optimization to minimize the objective function:

$$\theta_{best} = \underset{\theta}{\mathrm{argmin}}(L). \tag{26}$$

2.3. GPR with FE Analysis Results

In several previous studies, the prior mean function in GPR is set to a zero matrix, as described in Equation (4). However, the results of FE analysis can be introduced as the prior mean $m(\cdot)$, also called the explicit basis function, and the input-output relationship is expressed as:

$$\mathbf{y} = f(\mathbf{X}) + \varepsilon = \{m(\mathbf{X}) + \phi(\mathbf{X})\mathbf{w}\} + \varepsilon. \tag{27}$$

Compared with Equation (2), Equation (27) is seen to include the prior mean based on FE analysis, and this can help to overcome the limitation on predictions of the vertical deflection of a bridge under construction, when sufficient data is not available.

For computational simplicity, the residual differences between the output values and the explicit basis function values are considered as new outputs as follows,

$$\mathbf{y}' = \mathbf{y} - m(\mathbf{X}) = f'(\mathbf{X}) + \varepsilon = \phi(\mathbf{X})\mathbf{w} + \varepsilon \tag{28}$$

where $\mathbf{y}'$ is the set of new training outputs, and $f'(\cdot)$ is the new regression function. Then, with the new training outputs ($\mathbf{y}'$), we have an input-output relationship with a zero prior mean again, which was discussed in Section 2.1. The marginal likelihood can now be expressed as the conditional probability of the new training outputs given the training input matrix ($\mathbf{X}$) as follows:

$$p(\mathbf{y}'|\mathbf{X}) = \frac{1}{(2\pi)^{\frac{N_D}{2}} \cdot \left|\mathbf{K} + \sigma_n^2\mathbf{I}\right|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}\mathbf{y}'^T\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{y}'\right). \tag{29}$$

Once the hyperparameters are optimized, the predictive outputs can be estimated as follows,

$$f'(\mathbf{X}_*) \sim N\left(\mathbf{K}_*^{\mathbf{T}}\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{y}', \ \mathbf{K}_{**} - \mathbf{K}_*^{\mathbf{T}}\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{K}_*\right) \tag{30}$$

which is equivalent to:

$$f(\mathbf{X}_*) \sim N\left(m(\mathbf{X}_*) + \mathbf{K}_*^{\mathbf{T}}\left(\mathbf{K} + \sigma_{ne}^2\mathbf{I}\right)^{-1}(\mathbf{y} - m(\mathbf{X})), \ \mathbf{K}_{**} - \mathbf{K}_*^{\mathbf{T}}\left(\mathbf{K} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{K}_*\right). \tag{31}$$

2.4. Approximation Algorithm for Reducing Computational Cost

One of the limitations involved in the application of the GPR is the computational cost. GPR requires the calculation of the inverse matrix of the covariance matrix, as explained in Section 2.1; hence, the cost of GPR is sensitive to the size of the input and output datasets. To address this issue, various approximation methods were developed, which use a partial dataset to derive performance that is comparable to a standard GPR, where a full dataset is applied [31,34]. Among these methods, the Subset of Data (SoD) approximation, which is simple and known to have a better speed-accuracy trade-off than other approximations [35,36], is introduced for this study.

The key idea of the SoD approximation is to select the subset that provides a reasonable hyperparameter optimization, compared with the regression, using the whole dataset. Utilizing only the M_D number among the N_D quantities of the full dataset ($M_D < N_D$) naturally decreases the computational cost. The important task here is to select subset points, which facilitate an acceptable regression performance for building the posterior distribution, called the active set. Data selection algorithms, which use an active learning criterion (e.g., differential entropy [37], matching pursuit [38], or information gain [39]) have been used as more effective approaches because randomly selected subset points generally do not perform well [31,36]. This study introduces the differential entropy score suggested by Lawrence et al. (2003) as the criterion, which is maximized at the point with the largest variance [37]. Using the selection criterion, a portion of the whole dataset is selected as the active set ($\mathbf{D}_A$),

$$\mathbf{D}_A = \left\{(\mathbf{X}_A, \mathbf{y}_A)\right\} = \left\{(x_{ij}, y_i) \middle| i = 1, \ldots, M_D; j = 1, \ldots, N_x\right\} \tag{32}$$

where $\mathbf{X}_A$ and $\mathbf{y}_A$ are the selected training input matrix, and output vector (i.e., active set), respectively, whose sizes are M_D by N_x and M_D by 1. Once the active set is selected, the hyperparameters are optimized by utilizing the maximum likelihood concept described in Equations (29)–(31) with the active set. The marginal likelihood of the training output vector given the training input matrix of the active set can be described as follows,

$$p(\mathbf{y}_A | \mathbf{X}_A) = \frac{1}{(2\pi)^{\frac{M_D}{2}} \cdot \left|\mathbf{K}_A + \sigma_n^2 \mathbf{I}\right|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2} \mathbf{y}_A^T \left(\mathbf{K}_A + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y}_A\right) \tag{33}$$

where $\mathbf{K}_A$ is the covariance matrix of the training input matrix of the active set $\mathbf{X}_A$, whose size is M_D by M_D. The computational cost for calculating the inverse matrix in the likelihood function is much lower because of the small size of the covariance matrix. After optimizing the hyperparameters, the predictive outputs can be estimated:

$$f(\mathbf{X}_*) \sim N\left(\mathbf{K}_{A*}\left(\mathbf{K}_A + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y}, \mathbf{K}_{A**} - \mathbf{K}_{A*}\left(\sigma_n^2 + \mathbf{K}_A\right)^{-1} \mathbf{K}_{A*}\right) \tag{34}$$

where $\mathbf{K}_{A*}$ is the covariance matrix between the selected training ($\mathbf{X}_A$) inputs and the N_P number of prediction ($\mathbf{X}_*$) inputs, and $\mathbf{K}_{A**}$ is the covariance matrix of the prediction input matrix $\mathbf{X}_*$, whose sizes are M_D by N_P and N_P by N_P, respectively. When the explicit basis function is incorporated as the prior mean, the new outputs are calculated as:

$$f(\mathbf{X}_*) \sim N\left(\begin{array}{c} m(\mathbf{X}_*) + \mathbf{K}_{A*}^{\mathrm{T}}\left(\mathbf{K}_A + \sigma_n^2 \mathbf{I}\right)^{-1}(\mathbf{y}_A - m(\mathbf{X}_A)), \\ \mathbf{K}_{A**} - \mathbf{K}_{A*}^{\mathrm{T}}\left(\mathbf{K}_A + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{K}_{A*} \end{array} \right). \tag{35}$$

3. Application Example

3.1. Example Bridge Description

To test the proposed method, it was applied to a full-scale railway bridge under construction. The bridge consists of four PSC girders (G1, G2, G3, and G4, as shown in Figure 1a), and each girder has a trivet-shaped cross section, as shown in Figure 1b, and a width, height, and length of 2.0 m, 2.2 m, and 39.94 m, respectively. Among these four girders, G1 was selected as the target, for which actual measurements were taken, by considering various physical properties, such as vertical deflection and temperature.
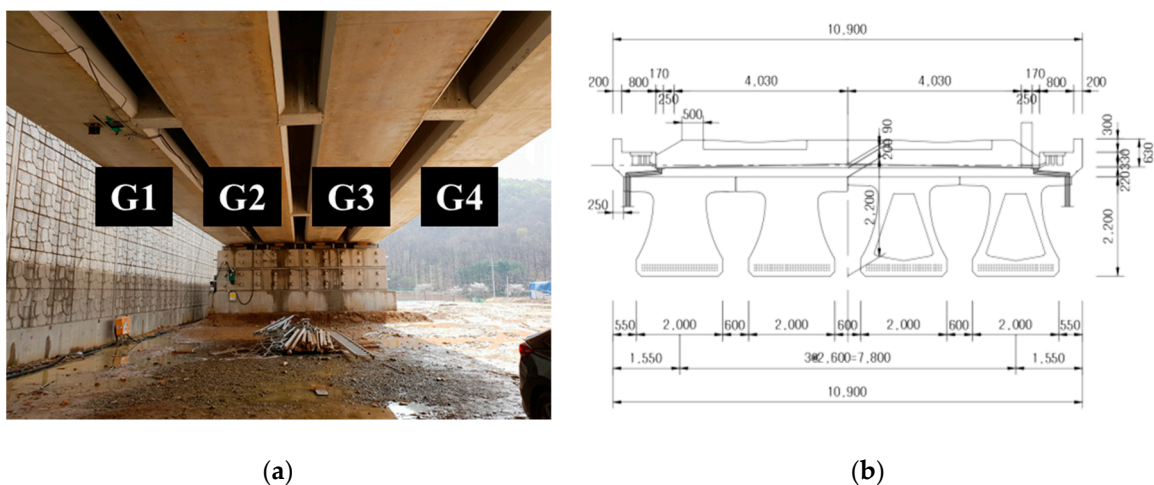

(a)   (b)

Figure 1. (**a**) Four pre-stressed concrete (PSC) girders of the example bridge; (**b**) a cross-sectional drawing.

By following the schedule listed in Table 1, the girder experienced construction loads by the tensioning and detensioning of tendons and substructure placements. The bridge construction started with girder concrete casting and tendon tensioning, which caused an upward deflection on 18 September 2017, defined as day index 1. Tendon releasing and girder erection proceeded in sequence on 22 September (day index 5), and 6 November (day index 50), respectively. Figure 2a shows the slab placement, which was performed from 25 January to 8 February 2018, and it caused a downward deflection. Figure 2b shows the placement of protective walls, which was performed from 14 February to 16 March 2018, and it caused additional downward deflection of the target bridge girder.

Table 1. Construction schedule of the example bridge.

Date	Day Index	Construction Schedule
18 September 2017	1	Girder concrete casting and tendon tensioning
22 September 2017	5	Tendon release
6 November 2017	50	Girder erection
8–9 February 2018	144–145	Slab placement
14 February 2018–16 March 2018	150–180	Protective wall placement

(**a**)   (**b**)

Figure 2. Bridge photos in the construction stages of (**a**) slab placement; and (**b**) protective wall placement.

3.2. Prior prediction based on FE analysis

For the prior prediction of the time-dependent vertical deflection of the target bridge, in this study, the three-dimensional FE model, established by Lee et al. (2018) is introduced [40]. The FE model was constructed using commercial FE analysis software, MIDAS/CIVIL 2017, which supports the consideration of a construction schedule and the time-dependent structural responses of creep and shrinkage [40,41]. The FE model has 212 nodes and 208 elements, and the constructed model is shown in Figure 3. In addition, the material properties of concrete and steel strands are summarized in Tables 2 and 3, respectively. The concrete for the slab, cross beam, and girder were modeled using the *C30*, *C30*, and *C49* models of MIDAS/CIVIL 2017, respectively, and the *SWPC7B 15.2 mm, Low Relaxation* model of MIDAS/CIVIL 2017 was used for strands. The construction stages and the associated schedule in Table 1 were considered by conducting a *construction stage analysis* in the software, and the KCI-USD12 code [10] was utilized to estimate the time-dependent deflection of the example bridge. The FE analysis result for the vertical deflection of the bridge, during the construction stages, was introduced as the explicit basis function in the GPR, and the prior prediction was updated using actual measurement data.
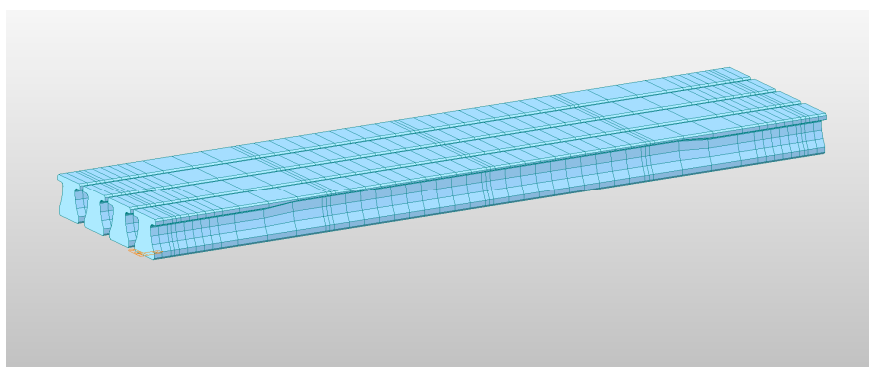
Figure 3. Finite element (FE) model of the example bridge for MIDAS/CIVIL 2017.

Table 2. Material properties of concrete.

Classification	Concrete		
	Slab	**Cross Beam**	**Girder**
Material code in MIDAS	*C30*	*C30*	*C49*
Compressive strength	30 MPa	30 MPa	49 MPa
Modulus of elasticity	27.52 GPa	27.52 GPa	32.08 GPa
Unit weight	2.5 ton/m^3	2.5 ton/m^3	2.5 ton/m^3
Poisson's ratio	0.18	0.18	0.18

Table 3. Material properties of strand.

Classification	Strand
Material code in MIDAS	*SWPC7B 15.2 mm, Low Relaxation*
Modulus of elasticity	200,000 MPa
Unit weight	7.85 ton/m^3
Poisson's ratio	0.3
Total tendon area	15.2 mm × 90 EA
Ultimate strength	1900 MPa
Yield strength	1600 MPa

3.3. Measurement Data

The vertical deflection was measured using the computer vision system proposed by Lee et al. [42]. A vision system with the hardware configurations shown in Figure 4a was deployed to monitor the displacement of the main target. In this dual-camera system, the displacement of the main target can be measured, regardless of camera movement, by placing a sub-target displacement at a stationary reference point. As shown in Figure 4b, the main target was deployed at the mid-span of the girder at which displacement was monitored. With the specifications of the camera and the lens listed in Table 4, the employed computer vision system can measure displacement with a resolution of 0.2 mm at a distance of 20 m.

Table 4. Hardware specifications.

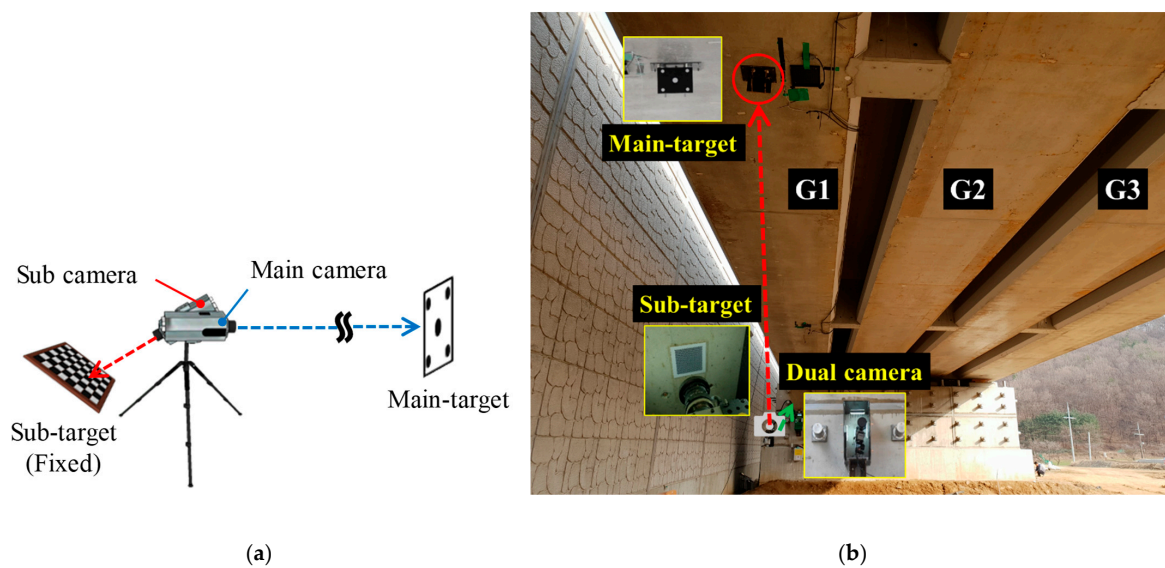| Item	Specifications
Main camera	Telescope lens (focal length: 800 mm) 1920 × 1200 resolution
Sub-camera	Wide-angle lens (focal length: 6 mm) 1920 × 1200 resolution
Main target	Five circular dots 15 × 11 cm^2
Sub-target	32 × 31 checkerboard

Figure 4. Experimental setup: (**a**) Dual camera configuration; (**b**) deployment of the computer vision system and the thermometer.

A total of 34,468 measurements were obtained for about nine months, beginning on 4 January 2018. Figure 5a shows the measurement data and the FE analysis results for the vertical deflection. Due to a malfunction in the dual-camera system, sensor data were not obtained from 10 May (day index 235) to 25 June (day index 281) 2018. The FE analysis results (represented by the green line) considers the construction schedule and the associated physical phenomena, and thus, it follows the trend of the measurements by the dual-camera system represented by black dots. However, the FE analysis can provide only a long-term deterministic prediction, without daily fluctuations, as shown in Figure 5b, which is a magnification of a portion of the plot in Figure 5a. Moreover, it is observed that the FE analysis results have a bias. That is, the FE analysis result cannot exactly predict the vertical deflection alone due to its model error.
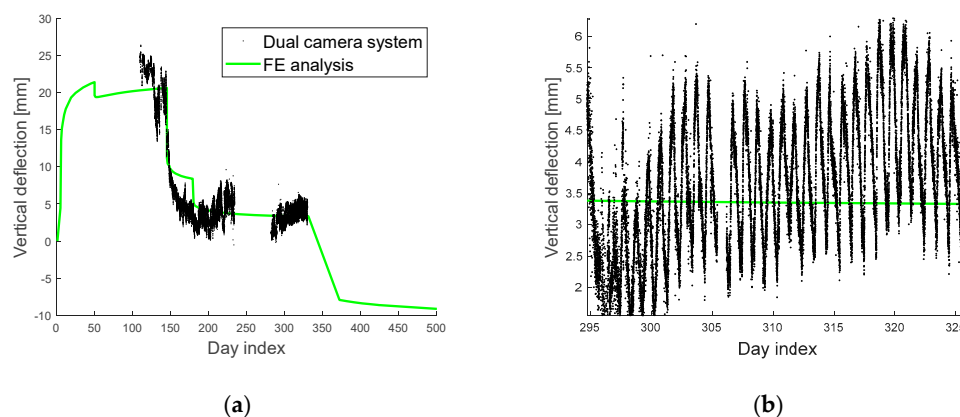

Figure 5. Plots of (**a**) sensor data of vertical deflection versus the prior prediction from the FE analysis; (**b**) a magnified view of a portion of the plot in Figure 5(**a**).

3.4. Probabilistic Prediction Results

The proposed method is applied to build probabilistic prediction models, and the flow chart, shown in Figure 6, summarizes the procedure. First, data are processed with the measurement data according to the FE model. Second, the SoD approximation is applied to reduce the computational cost. In this application example, a convergence test of the SoD approximation was conducted for the full set of 1690 data, and Figure 7 shows the results. It is observed from the figure that the root

mean square errors (RMSEs) of predictive mean and variance decrease with the increasing amount of considered data. Based on these results, the minimum number of SoD was selected to be 400. Next, kernel hyperparameters were determined through an optimization procedure using SoD, and the predictive mean and variance were calculated.
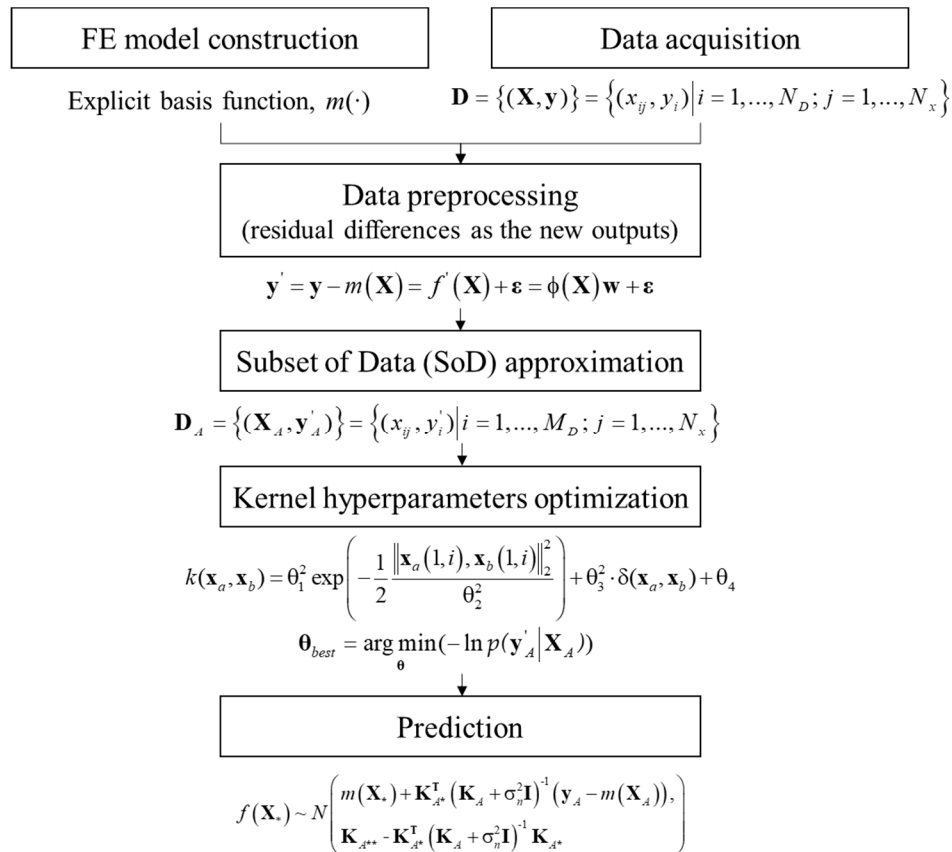

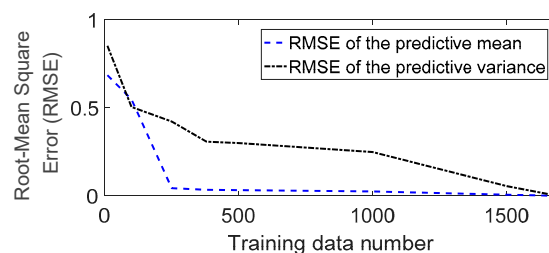**Figure 6.** Flow chart of the probabilistic prediction procedure.

Figure 7. Convergence test result for the training data number using the SoD approximation.

The measurement data from day index 109 to 140 is used, and the prediction results with, and without, the use of the FE analysis results are presented in Figure 8a,b respectively. It is seen in both of the figures that, until day 145, the actual measurement data on the vertical deflection are primarily located within the 99% prediction interval (PI). In Figure 8a, however, the measurement data are outside the prediction interval on day 145, when the slab was placed. This is because the slab placement produced a large downward deflection on the target girder, and it could not be considered without the FE analysis. As a result, the predictive mean and standard deviation of the vertical deflection are estimated to be 20.849, and 2.099 mm, respectively. However, the actual measurement on day 145 was 12.64 mm. Using the normal distribution assumption on the vertical deflection in GPR, the probability of observing at least the level of deviation from the mean is calculated to be $2 \times \Phi(-|12.64 - 20.849|/2.099)$ = 9.19×10^{-5}, where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF), which

indicates a very rare event. On the other hand, when the FE analysis result is used for the probability prediction, the predictive mean and standard deviation are estimated to be 12.639 and 2.364 mm for day 145, and the probability of observing at least the level of deviation from the mean is calculated to be $2 \times \Phi(-|12.64-12.639|/2.364) = 0.9996$, as shown in Figure 8b, which indicates a very probable event. These probabilities can be used for effective bridge management.
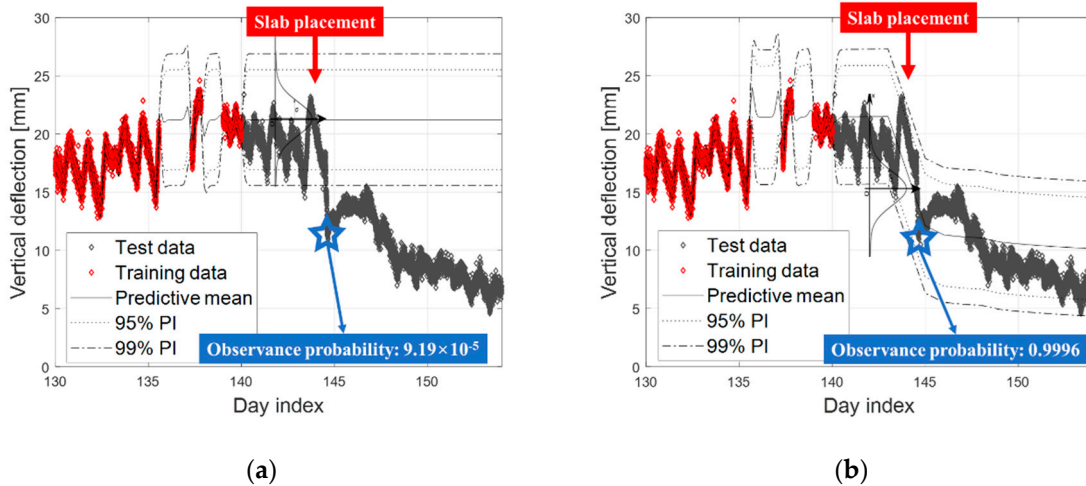

(a) (b)

Figure 8. Probabilistic prediction model trained using the data before the slab placement (**a**) without consideration of the FE analysis result; and (**b**) with the result.

Similarly, predictions are made using measurement data for longer periods, as described in Figure 9, by employing both, the measurement data and the prior prediction from the FE analysis. According to a previous study [2], measurement data, obtained for approximately three months, were necessary for an accurate prediction on bridge deflection when considering environmental effects. If sufficient training data are not available, environmental effects can be considered in the prediction through their introduction into the FE analysis. However, this is not within the scope of this study. The prediction model in Figure 10a is constructed using the measurement data between the day indices 109 and 180, and the model in Figure 10b utilized the data between the day indices 109 and 287. First, Figure 10a, which describes the deflection after the placement of the protective walls shows that the vertical deflection after the protective wall placement is within the expected range. Similarly, Figure 10b shows that the deflections are within the acceptable ranges, which means that the construction has proceeded as expected. Finally, the long-term deflection prediction model was constructed using the entire measurements from day indices 109 to 330, as shown in Figure 10c.
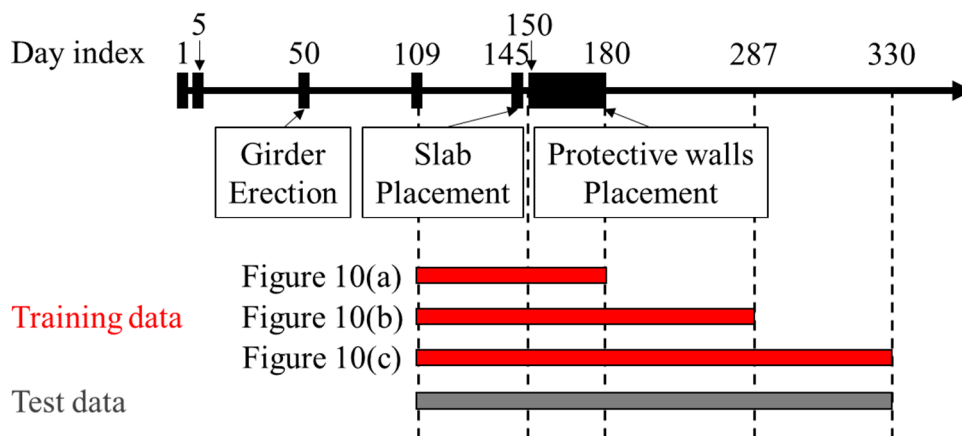

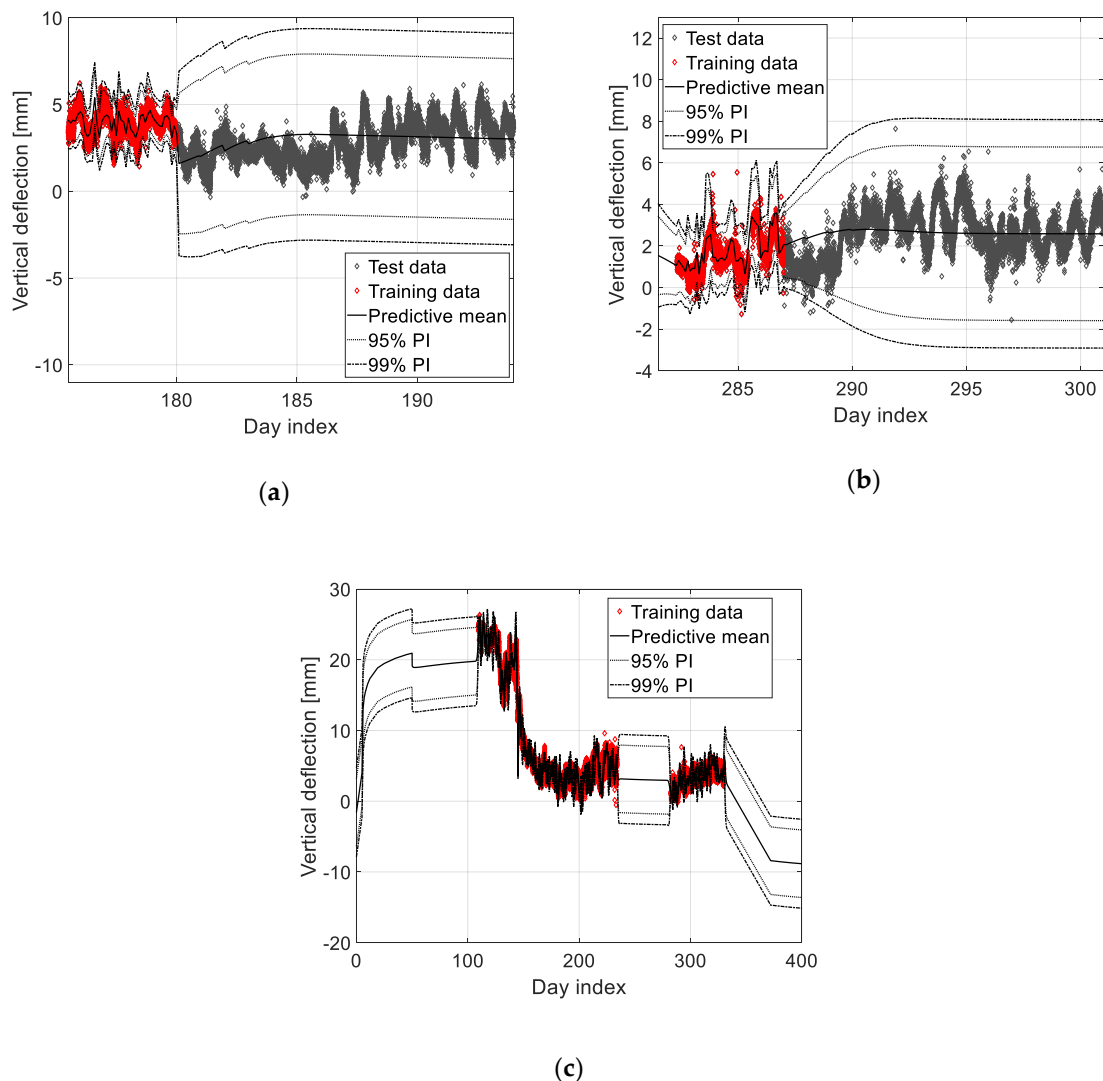**Figure 9.** Construction schedule and the training and test data of each prediction model.

(**a**)

(**b**)

(**c**)

Figure 10. Probabilistic prediction model trained using the data between day indices (**a**) 109 and 180; (**b**) 109 and 287; (**c**) 109 and 330.

As described above, the proposed method can be utilized to calculate the observance probability of an actual measurement. If the observance probability is estimated to be low, it may indicate that there is a serious problem with the bridge, and the bridge manager may consider shutting down the bridge immediately. On the other hand, if the observance probability is estimated to be high, despite the large deflection, the bridge manager can be confident that there is no problem with the bridge. In this manner, the proposed method is expected to support the efficient monitoring and management of bridges.

4. Conclusions

This study suggested a new nonparametric Bayesian method, which employs both an FE model and actual measurement data to construct the probabilistic prediction model of bridge deflection. To overcome the limitations of an imperfect FE model and a shortage of data, the GPR was introduced and modified to consider both, the FE analysis results and actual measurement data. In addition, the probabilistic prediction model can be updated whenever additional measurement data is available. The proposed method was applied to an actual railway bridge in the Republic of Korea, which is a PSC bridge under construction, and the probabilistic prediction results agreed with the measurements. In

addition, the proposed method calculates the observance probability for a given measurement, which can be useful for efficient monitoring and management of bridges.

References

1. Guo, T.; Sause, R.; Frangopol, D.M.; Li, A. Time-dependent reliability of PSC box-girder bridge considering creep, shrinkage, and corrosion. *J. Bridge Eng.* **2010**, *16*, 29–43. [CrossRef]

2. Lee, J.; Lee, K.C.; Lee, Y.J. Long-Term Deflection Prediction from Computer Vision-Measured Data History for High-Speed Railway Bridge. *Sensors* **2018**, *18*, 1488. [CrossRef] [PubMed]

3. International Union of Railways (UIC). *Testing and Approval of Railway Vehicle from the Point of View of Their Dynamic Behavior-Safety-Track Fatigue-Running behavior*; UIC Code 518; UIC: Paris, France, 2009.

4. Iles, D.C. *Design Guide for Steel Railway Bridges*; Steel Construction Institute: Berkshire, UK, 2004.

5. Korea Rail Network Authority. *Guideline of Track Maintenance*; Korea Rail Network Authority: Daejeon, Korea, 2016.

6. Guo, T.; Liu, T.; Li, A. Deflection reliability analysis of PSC box-girder bridge under high-speed railway loads. *Adv. Struct. Eng.* **2012**, *15*, 2001–2011. [CrossRef]

7. Nilson, A.H.; Winter, G.; Urquhart, L.C.; Charles Edward, O.R. *Design of Concrete Structures*; McGraw-Hill: New York, NY, USA, 1991; Volume 49.

8. Fédération Internationale Du Béton. *Structural Concrete: Textbook on Behaviour, Design and Performance*; Updated Knowledge of the CEB/FIP Model Code 1990; Bulletin No. 2; FIB: Lausanne, Switzerland, 1999.

9. Videla, C.; Carreira, D.J.; Garner, N. Guide for modeling and calculating shrinkage and creep in hardened concrete. *ACI Rep.* **2008**, 209.2R–2–209.2R–44.

10. KCI Committee. *Concrete Design Code and Commentary*; KCI Committee: Seoul, Korea, 2012.

11. Bažant, Z.P.; Baweja, S. Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3. *ACI Spec. Publ.* **2000**, *194*, 1–84.

12. Bažant, Z.P.; Yu, Q.; Li, G.H.; Klein, G.J.; Kristek, V. Excessive deflections of record-span prestressed box girder. *Concr. Int.* **2010**, *32*, 44–52.

13. Kamatchi, P.; Rao, K.B.; Dhayalini, B.; Saibabu, S.; Parivallal, S.; Ravisankar, K.; Iyer, N.R. Long-term prestress loss and camber of box-girder bridge. *ACI Struct. J.* **2014**, *111*, 1297. [CrossRef]

14. Bažant, Z.P.; Yu, Q.; Li, G.H. Excessive long-time deflections of prestressed box girders. I: Record-span bridge in Palau and other paradigms. *J. Struct. Eng.* **2012**, *138*, 676–686. [CrossRef]

15. Bažant, Z.P.; Yu, Q.; Li, G.H. Excessive long-time deflections of prestressed box girders. II: Numerical analysis and lessons learned. *J. Struct. Eng.* **2012**, *138*, 687–696. [CrossRef]

16. Sun, L.; Hao, X.W. Analysis of Bridge Deflection Based on Time Series. *Appl. Mech. Mater.* **2011**, *71*, 4545–4548. [CrossRef]

17. Xin, J.; Zhou, J.; Yang, S.; Li, X.; Wang, Y. Bridge structure deformation prediction based on GNSS data using Kalman-ARIMA-GARCH model. *Sensors* **2018**, *18*, 298. [CrossRef] [PubMed]

18. Pnevmatikos, N.G.; Hatzigeorgiou, G.D. Damage detection of framed structures subjected to earthquake excitation using discrete wavelet analysis. *Bull. Earthq. Eng.* **2017**, *15*, 227–248. [CrossRef]

19. Pnevmatikos, N.G.; Blachowski, B.; Hatzigeorgiou, G.D.; Swiercz, A. Wavelet analysis based damage localization in steel frames with bolted connections. *Smart Struct. Syst.* **2016**, *18*, 1189–1202. [CrossRef]

20. Liu, J.L.; Wang, Z.C.; Ren, W.X.; Li, X.X. Structural time-varying damage detection using synchrosqueezing wavelet transform. *Smart Struct. Syst.* **2015**, *15*, 119–133. [CrossRef]

21. Wang, C.; Ren, W.X.; Wang, Z.C.; Zhu, H.P. Time-varying physical parameter identification of shear type structures based on discrete wavelet transform. *Smart Struct. Syst.* **2014**, *14*, 831–845. [CrossRef]

22. Law, S.S.; Zhu, X.Q.; Tian, Y.J.; Li, X.Y.; Wu, S.Q. Statistical damage classification method based on wavelet packet analysis. *Struct. Eng. Mech.* **2013**, *46*, 459–486. [CrossRef]

23. Fan, Z.; Feng, X.; Zhou, J. A novel transmissibility concept based on wavelet transform for structural damage detection. *Smart Struct. Syst.* **2013**, *12*, 291–308. [CrossRef]

24. Yang, I.H. Prediction of time-dependent effects in concrete structures using early measurement data. *Eng. Struct.* **2007**, *29*, 2701–2710. [CrossRef]

25. Guo, T.; Chen, Z. Deflection control of long-span PSC box-girder bridge based on field monitoring and probabilistic FEA. *J. Perform. Constr. Facil.* **2016**, *30*, 04016053. [CrossRef]

26. Lee, Y.J.; Kim, R.; Suh, W.; Park, K. Probabilistic fatigue life updating for railway bridges based on local inspection and repair. *Sensors* **2017**, *17*, 936. [CrossRef]

27. Kim, R.E.; Moreu, F.; Spencer, B.F. System identification of an in-service railroad bridge using wireless smart sensors. *Smart Struct. Syst.* **2015**, *15*, 683–698. [CrossRef]

28. Moreu, F.; Kim, R.E.; Spencer, B.F., Jr. Railroad bridge monitoring using wireless smart sensors. *Struct. Control Health Monit.* **2017**, *24*, e1863. [CrossRef]

29. Kim, R.E.; Moreu, F.; Spencer, B.F., Jr. Hybrid model for railroad bridge dynamics. *J. Struct. Eng.* **2016**, *142*, 04016066. [CrossRef]

30. He, W.Y.; Zhu, S. Adaptive-scale damage detection strategy for plate structures based on wavelet finite element model. *Struct. Eng. Mech.* **2015**, *45*, 239–256. [CrossRef]

31. Rasmussen, C.E.; Williams, C.K. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.

32. Barber, D. *Bayesian Reasoning and Machine Learning*; Cambridge University Press: Cambridge, UK, 2012.

33. Hager, W.W. Updating the inverse of a matrix. *SIAM Rev.* **1989**, *31*, 221–239. [CrossRef]

34. Snelson, E.L. Flexible and Efficient Gaussian Process Models for Machine Learning. Ph.D. Thesis, University College London (UCL), London, UK, 2007.

35. Chalupka, K.; Williams, C.K.; Murray, I. A framework for evaluating approximation methods for Gaussian process regression. *J. Mach. Learn. Res.* **2013**, *14*, 333–350.

36. Liu, H.; Ong, Y.S.; Shen, X.; Cai, J. When Gaussian process meets big data: A review of scalable GPs. *arXiv* **2018**, arXiv:1807.01065.

37. Herbrich, R.; Lawrence, N.D.; Seeger, M. Fast sparse Gaussian process methods: The informative vector machine. In *Proceedings of the 15th International Conference on Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2002; pp. 625–632.

38. Keerthi, S.; Chu, W. A matching pursuit approach to sparse gaussian process regression. In *Proceedings of the 18th International Conference on Neural Information Processing Systems*; MIT Press: Vancouver, BC, Canada, 2006; pp. 643–650.

39. Seeger, M. *Bayesian Gaussian Process Models: PAC-Bayesian Generalisation Error Bounds and Sparse Approximations*; No. THESIS_LIB; University of Edinburgh: Edinburgh, UK, 2003.

40. Lee, J.; Lee, K.C.; Kwon, H.C.; Lim, K.M.; Min, K.H. Long-term Behavior of Pretension Girder Bridges by Construction Phase Time. *J. Korean Soc. Hazard Mitig.* **2018**, *18*, 51–56. [CrossRef]

41. *Midas Civil 2017 [Computer Software]*; MIDAS Engineering Software: Seongnam, Korea, 2017.

42. Lee, J.; Lee, K.C.; Jeong, S.; Lee, Y.J.; Sim, S.H. Long-term displacement of full-scale bridges using camera ego-motion compensation. Mechanical Systems and Signal Processing, Under Review. 2019.