Smart Structures and Systems, *Vol. 33, No. 1 (2024) 17-26*
https://doi.org/10.12989/sss.2024.33.1.017

A long-term tunnel settlement prediction model based on BO-GPBE with SHM data

Yang Ding*[1,2,3], Yu-Jun Wei [4a], Pei-Sen Xi [5b], Peng-Peng Ang [5c] and Zhen Han [6d]

[1] *Department of Civil Engineering, Hangzhou City University, Hangzhou 310015, China*
[2] *Zhejiang Engineering Research Center of Intelligent Urban Infrastructure, Hangzhou City University, 310015, China*
[3] *Key Laboratory of Safe Construction and Intelligent Maintenance for Urban Shield Tunnels of Zhejiang Province, Hangzhou City University, Hangzhou, 310015, China*
[4] *Department of Civil Engineering, Zhejiang University, Hangzhou 310058, China*
[5] *Zhejiang Engineering Research Center of Smart Rail Transportation, Power China Huadong Engineering Corporation Limited, Hangzhou 311100, China*
[6] *Nanjing Metro Operation Co., Ltd., Nanjing, Jiangsu, 210012, China*

(*Received July 29, 2023, Revised November 23, 2023, Accepted December 1, 2023*)

Abstract. The new metro crossing the existing metro will cause the settlement or floating of the existing structures, which will have safety problems for the operation of the existing metro and the construction of the new metro. Therefore, it is necessary to monitor and predict the settlement of the existing metro caused by the construction of the new metro in real time. Considering the complexity and uncertainty of metro settlement, a Gaussian Prior Bayesian Emulator (GPBE) probability prediction model based on Bayesian optimization (BO) is proposed, that is, BO-GPBE. Firstly, the settlement monitoring data are analyzed to get the influence of the new metro on the settlement of the existing metro. Then, five different acquisition functions, that is, expected improvement (EI), expected improvement per second (EIPS), expected improvement per second plus (EIPSP), lower confidence bound (LCB), probability of improvement (PI) are selected to construct BO model, and then BO-GPBE model is established. Finally, three years settlement monitoring data were collected by structural health monitoring (SHM) system installed on Nanjing Metro Line 10 are employed to demonstrate the effectiveness of BO-GPBE for forecasting the settlement.

Keywords: Bayesian emulator; Bayesian optimization; Gaussian prior; settlement probability prediction; structural health monitoring

1. Introduction

In recent years, the rapid urbanization has resulted in a significant increase in urban traffic demand, while the limited traffic supply capacity has failed to keep up with the growth (Ding *et al*. 2023a, b). This has led to severe traffic congestion in big cities, making the metro system the only viable solution to alleviate the problem (Ding *et al*. 2023c, Moaveni and Najafi 2017, Li *et al*. 2020). For instance, as of December 30, 2021, Shanghai has 20 metro lines, Beijing has 27 metro lines, and Nanjing has 11 metro lines. However, the construction of new metro lines can cause settlement or floating of existing structures, including existing metros, buildings, among others, which can result in safety damage accidents (Li and Wang 2019, Qiu *et al*. 2020). Therefore, it is necessary to strengthen the existing structures and monitor/predict their settlement to prevent accidents from happening (Liu *et al*. 2020a, Mu *et al*. 2021, Qu *et al*. 2023).

In addition, settlement monitoring and prediction play a crucial role in determining the appropriate location for metro reinforcement (Ding *et al*. 2023d, Miliziano and de Lillis 2019). Therefore, settlement monitoring and prediction are essential measures during metro construction and operation (Ding *et al*. 2023e, Tu *et al*. 2020). Nowadays, prediction methods can be divided into two categories: deterministic and uncertainty methods (Ding *et al*. 2023f, Ye *et al*. 2019, 2020). For deterministic prediction methods, Deng *et al*. (2022) derived a prediction formula for surface subsidence using the mirror theory and Mindlin solution and built a finite difference model based on the Changsha power tunnel project. The results of theoretical prediction and numerical simulation matched well with the field monitoring data, demonstrating the accuracy of the ground loss model and the theoretical prediction formula. Zhang *et al*. (2020a) showcased the application of machine learning (ML) algorithms to predict settlement caused by tunnel excavation. The results showed that the backpropagation neural network (BPNN) has strong extrapolation ability, making it suitable for establishing settlement prediction models under conditions with a small existing dataset. Shi *et al*. (2019) applied the support vector machine (SVM) information granulation method to predict

*Corresponding author, Ph.D., Lecturer,
E-mail: ceyangding@zju.edu.cn
[a] Ph.D. Student, E-mail: ceyjwei@zju.edu.cn
[b] Engineer, E-mail: cepsxi@zju.edu.cn
[c] Engineer, E-mail: cepengpengang@zju.edu.cn
[d] Engineer, E-mail: 605460326@qq.com

ISSN: 1738-1584 (Print), 1738-1991 (Online)

the deformation of surrounding rocks. The results indicated that the SVM information granulation method was effective in predicting the shallow tunnel surrounding rock deformation in Panlongshan Tunnel of Qinglan Expressway in China, demonstrating relatively high accuracy. Overall, the deterministic prediction methods have their advantages, but the uncertainty methods, such as the Bayesian optimization-based GPBE model proposed by the authors, can better quantify the uncertainty arising from the complex and uncertain metro settlement process. These methods can provide reliable and accurate probabilistic predictions of settlement, which is essential for ensuring the safety and efficiency of metro construction and operation.

However, the metro construction process involves various uncertain factors, and the deterministic prediction method may not effectively quantify the uncertainty (Zhang *et al*. 2019, 2023, Fang *et al*. 2020). Therefore, uncertainty prediction methods are becoming more popular in settlement prediction. For instance, Bullock *et al*. (2019) proposed a model for predicting the settlement of shallow foundation structures on liquefied foundations during earthquakes. They modeled the estimated uncertainty of the model using a lognormal distribution, and the total uncertainty of their prediction had strict characteristics. Gong *et al*. (2014) used the maximum likelihood method to estimate statistical data of soil parameters from hypothetical field exploration at different levels, analyzed land subsidence caused by tunnel excavation using the probability method, and assessed the effect of field exploration. Wang *et al*. (2013) used a smooth correlation vector machine with wavelet kernel (wsRVM) to study the development of surface subsidence caused by tunnel excavation. The results showed that the prediction model performed well, and the forecast distribution could provide a measure of forecast uncertainty. Overall, uncertainty prediction methods can better handle the uncertainty arising from the complex and uncertain metro settlement process, and provide more reliable and accurate probabilistic predictions of settlement, which is essential for ensuring the safety and efficiency of metro construction and operation.

In uncertainty prediction methods, there are many unknown parameters known as hyperparameters that require optimization algorithms to automatically identify them (Ni *et al*. 2023, Zhang *et al*. 2020b, Luat *et al*. 2020). Hyperparameters also have uncertainty characteristics that need to be optimized by the optimization algorithm based on uncertainty theory (Han *et al*. 2022). For example, Chen and Sun (2021) developed a Bayesian learning approach to update the complete model and quantify uncertainty to determine the baseline of damage detection. Huang *et al*. (2021) introduced the Bayesian optimization algorithm to optimize the hyperparameters of the gradient-enhanced regression tree to obtain more satisfactory prediction accuracy and calculation cost. Liu *et al*. (2020b) used DEPSO to introduce differential evolution (DE) operators into the elementary particle swarm optimization (PSO) algorithm to estimate the hyperparameters of the Gaussian process regression model. Madra *et al*. (2019) calibrated the model's hyperparameters using a customized multi-objective evolutionary algorithm. Tan *et al*. (2016) developed a fast variational approximation scheme for Gaussian processes that enables the uncertainty of the hyperparameters of the covariance function to be handled without the use of Monte Carlo methods and is robust to overfitting. Overall, optimizing hyperparameters is crucial in uncertainty prediction methods to improve the accuracy and reliability of the prediction. The optimization algorithms can be customized and tailored to the specific characteristics of the model and the dataset, and can be used to handle the uncertainty of hyperparameters, which is essential for effective metro settlement prediction.

In this paper, the improved Gaussian Prior Bayesian Emulator (GPBE) was established based on Bayesian optimization (BO), that is, BO-GPBE. The present BO-GPBE model enables analytical expressions of posterior predictive distribution, which overcomes the issue of the high computational cost. The application of the BO-GPBE to probabilistic predicts of the settlement is verified based on SHM data collected from Nanjing Metro Line 10. The influence of five different acquisition functions in BO, which include expected improvement (EI), expected improvement per second (EIPS), expected improvement per second plus (EIPSP), lower confidence bound (LCB), probability of improvement (PI) on prediction performance of the settlement is investigated.

2. Probability prediction model based on BO-GPBE

2.1 Bayesian theory

Bayes theorem was developed by Thomas Bayes to describe the relationship between two conditional probabilities. In other words, the Bayesian theory is a probability model, which can be expressed by (Ni *et al*. 2020)

$$p(A|B) = \frac{p(A,B)}{p(B)} = \frac{p(A)p(B|A)}{p(B)} \propto p(A)p(B|A) \tag{1}$$

where $p(A|B)$ is the a posteriori probability of event A; $p(B|A)$ is the likelihood function, which is the modification of prior probability after considering new evidence, and new evidence can change our estimate of probability by affecting likelihood; $p(A)$ is the a priori probability of event A, which is a preliminary estimate of the probability before considering new evidence, and the prior probability is usually based on previous experience, knowledge or assumptions; $p(B)$ is a constant.

It can be seen that if we want to know the probability of parameter A when parameter B is known, we can calculate it by likelihood function and prior function. Specifically, in the process of Bayesian update, we first calculate the posterior probability according to the prior probability and likelihood, and then this posterior probability becomes a new prior probability for the next Bayesian update. This process can be repeated to constantly update the probability estimates based on new evidence. Furthermore, when there

are n random events, that is, A_1,…, A_n, , which can be expressed by (Ding *et al*. 2023f).

$$p(\mathrm{A}_i|\mathrm{B}) = \frac{p(\mathrm{A}_i\mathrm{B})}{p(\mathrm{B})} = \frac{p(\mathrm{A}_i)p(\mathrm{B}|\mathrm{A}_i)}{\sum_{i=1}^{n} p(\mathrm{A}_i)p(\mathrm{B}|\mathrm{A}_i)} \tag{2}$$

2.2 Gaussian prior Bayesian emulator

Gaussian process is a random process, that is, a set of random variables indexed by time or space. The finite random variables in this set form a multi-dimensional Gaussian distribution. Gaussian process is the joint distribution of all random variables in this set. Suppose there are n data sets $\mathcal{D} = \{(x_i, y_i)|_{i=1}^{n}\}$, x_i is the input value and y_i is the output value. Based on the Gaussian process, the following relationship exists between the input value and the output value, that is, (Wan and Ni 2018)

$$y = f(x) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_{\mathrm{n}}^2) \tag{3}$$

where $f(x)$ represents the implicit function; ε is noise; σ_{n}^2 is the variance of the error value. In the assumption of Gaussian process, the mean function $m(x)$ and covariance function $C(x, x')$ of implicit function can be expressed by (Wan and Ni 2018)

$$\begin{aligned} m(x) &= \mathbb{E}[f(x)] \\ C(x, x') &= \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))] \end{aligned} \tag{4}$$

Furthermore, the joint Gaussian distribution can be expressed as (Wan and Ni 2018)

$$\begin{aligned} p(\mathbf{f}, f_*) &= \mathcal{N}\left(\begin{bmatrix}\mathbf{0}\\ 0\end{bmatrix}, \begin{bmatrix}\mathbf{C} & \mathbf{C}_* \\ \mathbf{C}_*^{\mathrm{T}} & \tilde{C}\end{bmatrix}\right) \\ \mathbf{C} &= C(\mathbf{X}, \mathbf{X}) \\ \mathbf{C}_* &= C(x_*, \mathbf{X}) \\ \tilde{C} &= C(x_*, x_*) \end{aligned} \tag{5}$$

where $X = \{x_i\}|_{i=1}^{n}$, $Y = \{y_i\}|_{i=1}^{n}$, $\mathbf{f} = f(\mathbf{X})$, $f_* = f(x_*)$.

Based on the Bayesian theorem, the posterior distribution of the predicted value can be expressed by

$$\begin{aligned} p(\mathbf{f}, f_*|\mathbf{Y}) &= \frac{p(\mathbf{f}, f_*)p(\mathbf{Y}|\mathbf{f})}{p(\mathbf{Y})} \\ p(\mathbf{Y}|\mathbf{f}) &= \mathcal{N}(\mathbf{f}, \sigma_{\mathrm{n}}^2 \mathrm{I}_n) \end{aligned} \tag{6}$$

where $\mathbf{I}_n$ is an identity matrix of $n \times n$.

Based on the definition of conditional probability distribution, when the value of observation output $\mathbf{Y}$ is given, the posterior distribution of prediction output f_* can be expressed as

$$\begin{aligned} p(f_*|\mathbf{Y}) &= \int p(\mathbf{f}, f_*|\mathbf{Y}) d\mathbf{f} \\ &= \frac{1}{p(\mathbf{Y})} \int p(\mathbf{f}, f_*) p(\mathbf{Y}|f) d\mathbf{f} \end{aligned} \tag{7}$$

When all parameters in Eq. (7) can be expressed by Gaussian distribution, the posterior distribution of prediction output can also be expressed by Gaussian distribution, that is, Gaussian Prior Bayesian Emulator (GPBE) (Ye *et al*. 2021)

$$\begin{aligned} p(f_*|\mathbf{Y}) &= \mathcal{N}\left(\mu_{f_*}, \sigma_{f_*}^2\right) \\ \mu_{f_*} &= \mathbf{C}_* \mathbf{K}^{-1} \mathbf{Y} \\ \sigma_{f_*}^2 &= \tilde{C} - \mathbf{C}_*^{\mathrm{T}} \mathbf{K}^{-1} \mathbf{C}_* \end{aligned} \tag{8}$$

where $\mathbf{K} = \mathbf{C} + \sigma_{\mathrm{n}}^2 \mathbf{I}_n$.

Obviously, the mean function and covariance function are important functions in the Bayesian prediction model based on Gaussian process. Generally, the mean function is assumed to be 0 (Wan and Ni 2019), and the covariance function is squared exponential (SE), which can be expressed as

$$C_{\mathrm{SE}}(x, x') = \eta^2 \exp\left[-\frac{(x - x')^2}{2\ell^2}\right] \tag{9}$$

where η^2 is the signal variance and ℓ is the characteristic length scale. Specifically, η^2 represents the random noise in the model. Larger signal variance means larger noise of observation data, while smaller signal variance means smaller noise of observation data. By adjusting the signal variance, the fitting degree of the Gaussian process model to the observed data and the prediction uncertainty of the unknown data can be controlled. ℓ determines how fast the function changes in the input space. A larger length scale means that the function changes slowly, while a smaller length scale means that the function changes violently. By adjusting the length scale, the smoothness and fluctuation of Gaussian process model can be controlled.

2.3 Evaluation of hyperparameters with the Bayesian optimization

In machine learning literature, the parameters of GPBE are commonly termed as hyperparameters, that is, the signal variance η^2 and the characteristic length scale ℓ. Bayesian optimization (BO) is an application method of the Bayesian theorem, which incorporates prior distribution and updates the prior with samples drawn from to obtain a posterior that accurately approximates (Brochu *et al*. 2010). Thus, the hyperparameters problems in the GPBE can be solved by BO, which can be cast as an optimization problem by maximizing or minimizing an objective function with some unknown parameters (Wang *et al*. 2019).

In general, the objective function $J(\theta)$ can be sampled at $\theta_t = \underset{\theta}{\mathrm{argmax}} u(\theta|\mathcal{D}_{1:t-1})$, where u is the acquisition function and $\mathcal{D}_{1:t-1} = \{(\theta_1, J(\theta_1)), \ldots, (\theta_{t-1}, J(\theta_{t-1}))\}$ are the t-1 samples drawn from the objective function so far (Vereecken *et al*. 2020). Generally, the acquisition function has many forms (Jones *et al*. 1998, Bull 2011, Snoek *et al*. 2012, Gelbart *et al*. 2014), such as expected improvement (EI), expected improvement per second (EIPS), expected improvement per second plus (EIPSP), lower confidence bound (LCB), probability of improvement (PI), and etc., which can be expressed separately as the Eq. (10) to Eq. (14).

$$\mathrm{EI}(\theta) = E \max(J(\theta_{best}) - J(\theta), 0) \tag{10}$$

$$\mathrm{EIPS}(\theta) = \frac{\mathrm{EI}(\theta)}{\mu(\theta)} \tag{11}$$

$$\mathrm{EIPSP}(\theta) = \begin{cases} (\mu(\theta) - J(\theta_{best}) - \xi)\Phi(Z) + \sigma(x)\phi(Z) & \text{if } \sigma(x) > 0 \\ 0 & \text{if } \sigma(x) = 0 \end{cases}$$
$$Z = \begin{cases} (\mu(\theta) - J(\theta_{best}) - \xi)/\sigma(\theta) & \text{if } \sigma(\theta) > 0 \\ 0 & \text{if } \sigma(\theta) = 0 \end{cases} \quad (12)$$

$$\mathrm{LCB}(\theta) = \mu(\theta) - \beta\sigma(\theta) \quad (13)$$

$$\mathrm{PI}(\theta) = \mu(\theta) - J(\theta_{best}) \quad (14)$$

where $\theta = \{\eta^2, \ell\}$ in the GP model; $J(\theta_{best})$ is the value of the best sample so far, and θ_{best} is the location of that sample, i.e., $\theta_{best} = \underset{\theta \in \{\theta_1, \theta_2 \ldots, \theta_t\}}{\operatorname{argmin}} J(\theta_i)$; $\mu(\theta)$ and $\sigma(\theta)$ are the mean and standard deviation of the GP posterior predictive at θ, respectively; $\Phi(Z)$ and $\phi(Z)$ are the cumulative distribution function and probability density function of the standard normal distribution, respectively; and ξ determines the amount of exploration during optimization, where higher values lead to further exploration.

3. Engineering application

Zhongsheng station of line 7 is located at the intersection of Taishan road and Hexi street, which is planned to open to traffic in 2023 year. It is arranged along the north-south direction of Taishan Road and intersects with Zhongsheng station of line 10. At the intersection, Zhongsheng station of line 7 crosses Zhongsheng station of line 10. The station scale is 270 m × 21.9 m × 21.06 m (long × wide × deep). The Zhongsheng station of the existing line 10 is arranged in the east-west direction along the south side of Hexi street, which can be seen as Fig. 1 (Ding *et al*. 2023g, h).

The existing Zhongsheng Station on Line 10 has a narrow transfer space and a side platform with a width of only 3.75 m, which includes one bay in the paid area and two bays in the non-paid area. Due to this, there is severe passenger flow congestion during operation. In contrast, the newly built Zhongsheng Station on Line 7 is an island type station with a rectangular frame structure, consisting of two floors underground (partially three floors) and a column spacing of 9 m. It was constructed using the open excavation method (partially concealed excavation method).

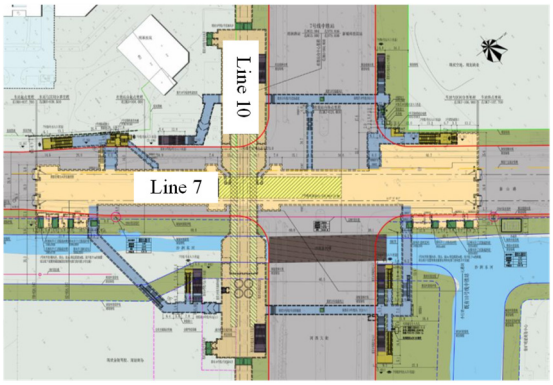
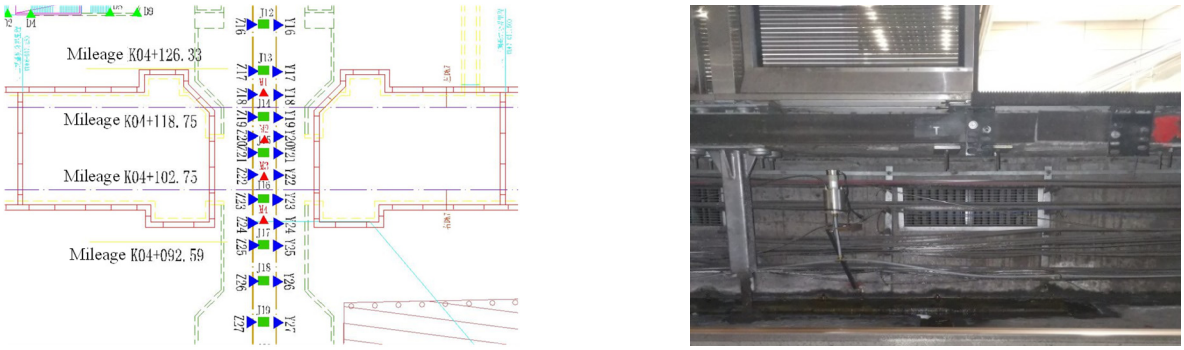


Fig. 1 General layout of Zhongsheng station

The station hall is located on the first floor underground, and the platform is on the second floor underground, with an effective platform width of 13 m. The total length of the station outer covering is 270.0 m, the outer covering width of the standard section is 21.9 m, and the thickness of the roof covering soil is about 3.2 m. The buried depth of the bottom plate of the standard section is 20.76 m, and the buried depth of the bottom plate of the left and right end wells is 22.15 m.

Zhongsheng station of existing Line 10 was completed in 2004, and initial value monitoring during the operation period was completed in August 2005. The construction of the underground excavation and underpass section of the new Line 7 station may have several impacts on the Zhongsheng station of existing Line 10, including: (1) Improper opening of MJS, horizontal freezing holes, and deep hole grouting may result in water and sand gushing, posing a potential safety hazard to the settlement of operating subway stations. (2) MJS is adjacent to the existing line pile site for shotcrete construction, and the pressure in the formation is controlled too high, which may cause uplift to the existing line station. (3) The freezing construction may have significant frost heave effects due to uneven solid addition and local water systems, which may cause uplift effects on existing lines. (4) The MJS horizontal reinforcement area outside the excavation section of the concealed excavation section is located in the silty fine sand layer and silty soil layer, which is a confined aquifer with poor self-stability. During the construction process, the hole wall may collapse, leading to accidents such as buried drilling and hole mouth surge, causing disturbance to the soil under the bottom plate of the Zhongsheng station of existing Line 10, affecting the stability of the existing Line 10. (5) At the end of freezing and thawing, the thawing settlement of the soil layer has a significant impact on the bottom slab of the upper station, resulting in an impact on the structural stability of the Zhongsheng station of existing Line 10. (6) During the excavation of the underpass section, the disturbance of soil excavation and the inadequate support of the cross-section are likely to cause settlement impact on the Zhongsheng station of existing Line 10. (7) The freezing water stop effect is not ideal, and water and sand gushing occur during the excavation process, affecting the operation safety of the existing line. (8) After excavation, improper support removal or failure of poured concrete to meet strength requirements may affect the stability of the existing line. (9) After the construction of the secondary lining is completed, gaps behind the structure may have a safety impact on the long-term operation of the existing line. (10) During the excavation process, the MJS reinforced solids and frozen curtains need to bear the water and soil pressure of the external confined water. If the reinforcement strength is insufficient, it may cause sudden surges in the reinforcement area at the excavation bottom, which may affect the



(a) General layout of monitoring points of line 10 (b) Measuring point and monitoring system

Fig. 2 General layout of Zhongsheng station

stability of the existing line.

The SHM system deployed on Zhongsheng Station is shown in Fig. 2, consisting of 31 settlement monitoring points before the construction of new Metro Line 7, denoted as J1 - J31, and four settlement monitoring points after the construction of new Metro Line 7, denoted as M1 - M4. Specifically, the letter J represents the existing column, and the letter M represents the reinforced column. The distance between existing columns is 10m, and the distance between existing columns and reinforced columns is 5m. However, due to the extensive monitoring layout point map, some monitoring layout points have been cut off, mainly including the monitoring layout points of new Metro Line 7 that crosses the existing Line 10, denoted as J12 - J19.

The settlement monitoring data from November 23, 2018 to November 23, 2021 were used to demonstrate the presented BO-GPBE prediction model. As shown in Figs. 3-7, during 2018 and 2019 year, the settlement of the existing tunnel structure was caused by the deformation of the structure during the excavation process of construction, resulting in settlement. During 2019 and 2020 year, the settlement and deformation of the existing tunnel structure were stable because, after the completion of construction, without external interference, the tunnel structure would not experience significant deformation, and uniform settlement would be generated under its own load. During 2020-2021 year, the construction of new tunnels caused the existing tunnel structure to rise. The management unit adopted MJS to reinforce the existing structure to ensure that the deformation of the existing tunnel structure did not exceed the specified value. Therefore, it is apparent that the excavation of the new tunnel had an impact on the existing structure.

To overcome the computational intensity of predicting time series as more data becomes available, the moving window strategy is adopted to reduce the size of the training data, effectively alleviating the high computational cost. As the window slides along the data, a new process model is generated by including the newest sample and excluding the oldest one. With the moving window strategy, a small set of training data, closest to the prediction point and with a fixed size, is adopted for constructing BO-GPBE. In this paper, the moving window size is set to 10. Specifically, the next settlement value is only related to its previous ten historical

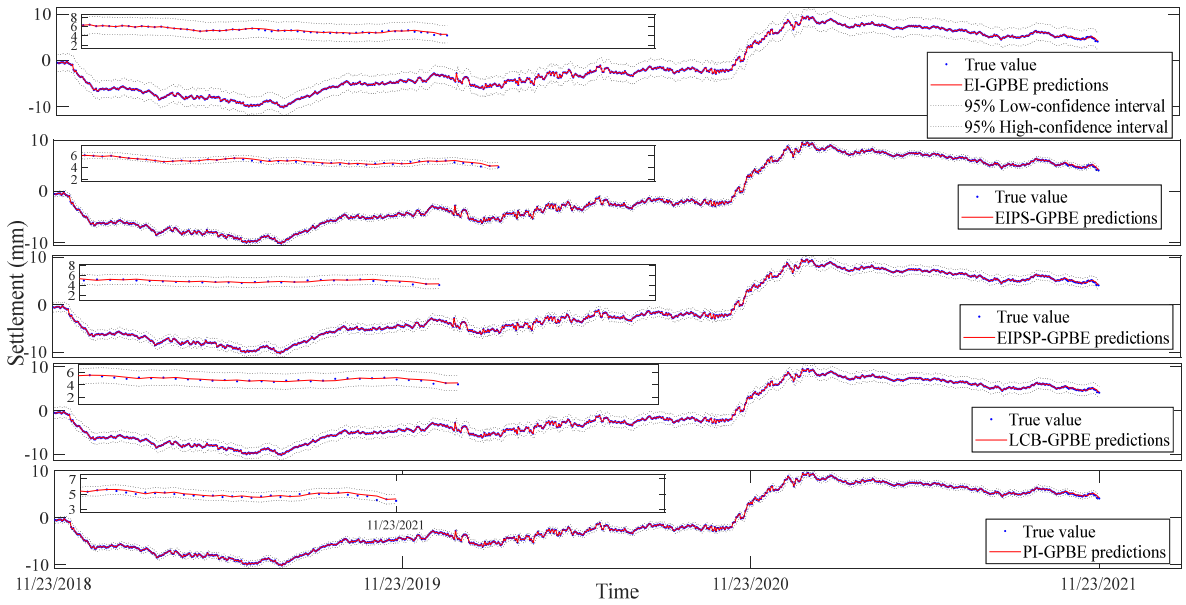


Fig. 3 Prediction results of point J13

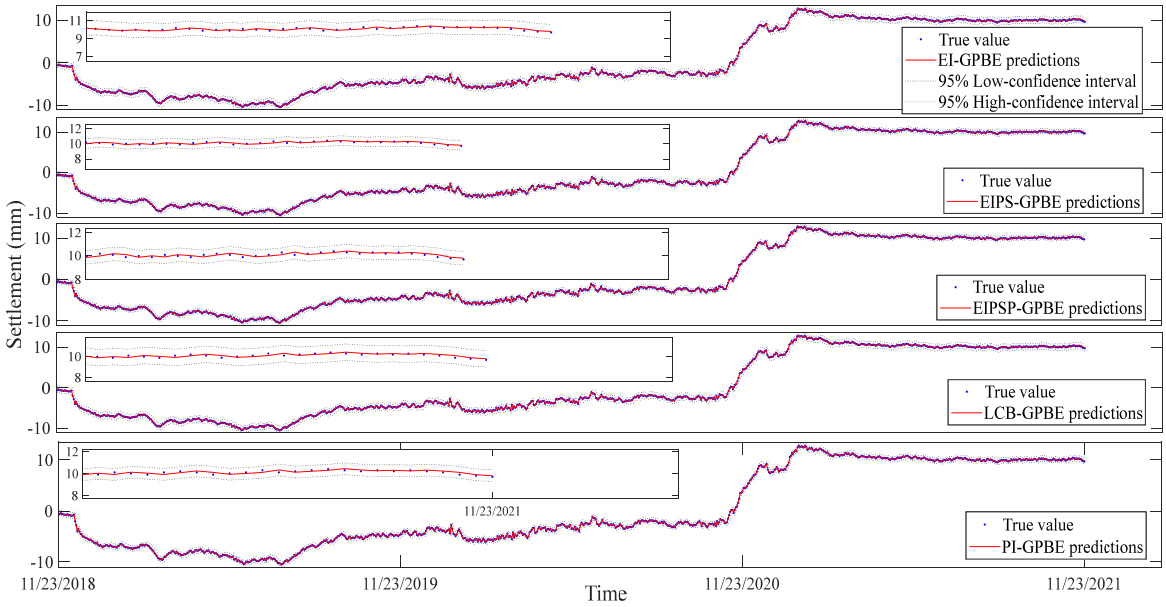


Fig. 4 Prediction results of point J14

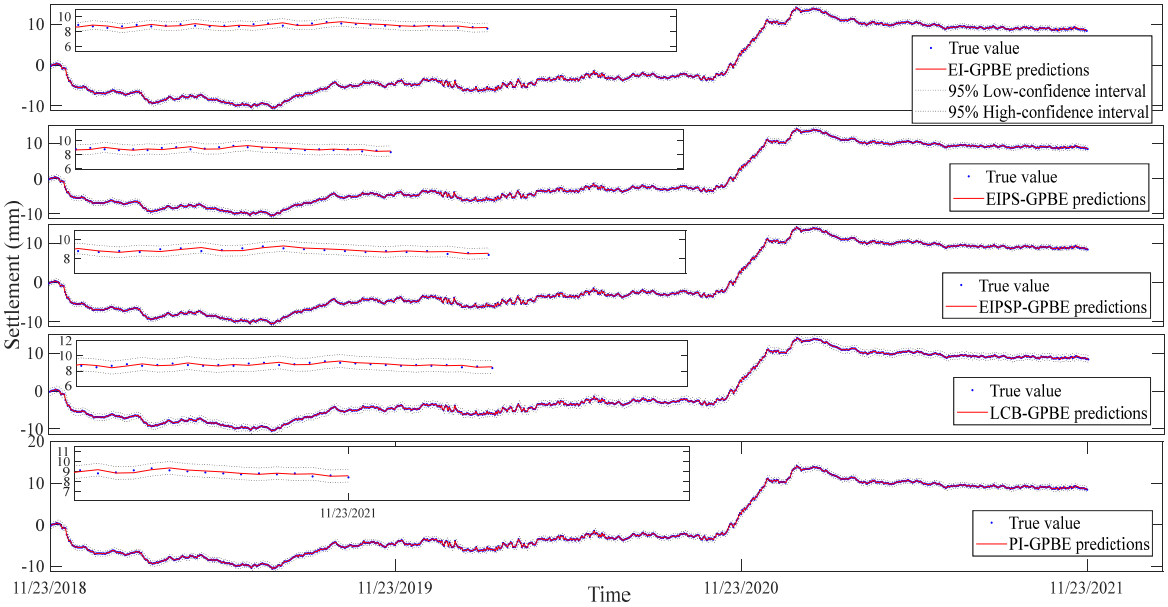


Fig. 5 Prediction results of point J15

settlement values. The predicting results of monitoring points J13-J17 of Line 10 based on different acquisition functions, i.e., EI, EIPS, EIPSP, LCB, and PI, are shown in Figs. 3-7. As depicted in Figs. 3-7, the proposed model with different acquisition functions (EI, EIPS, EIPSP, LCB, and PI) can effectively predict the deformation patterns (settlement and float) of existing tunnel structures and reflect the trend of deformation. Furthermore, the proposed model can fully characterize the uncertainty of deformation data by using the 95% confidence interval to represent the uncertainty of predicted data. Specifically, during the training process, the training values based on the BO-GPBE model are in good agreement with the measured settlement values. During the prediction process, the predicted values based on the BO-GPBE model are within the upper and lower ranges of the measured settlement values, that is, 95% confidence interval.

It can be seen from the Figs. 3-7, the settlement value of each monitoring point decreases (sinking), then increases (floating), and finally tends to be stable. Specially, the settlement has increased (floating) significantly during 2019 to 2020 year, which is caused by the construction of new Metro Line 7. In other words, when the new subway crosses the existing subway, the original stress changes due to the disturbance of the surrounding rock and soil mass, resulting in sinking or floating. Furthermore, the root mean square error (RMSE) is used to assess the performance of different acquisition functions in settlement predict quantitatively,

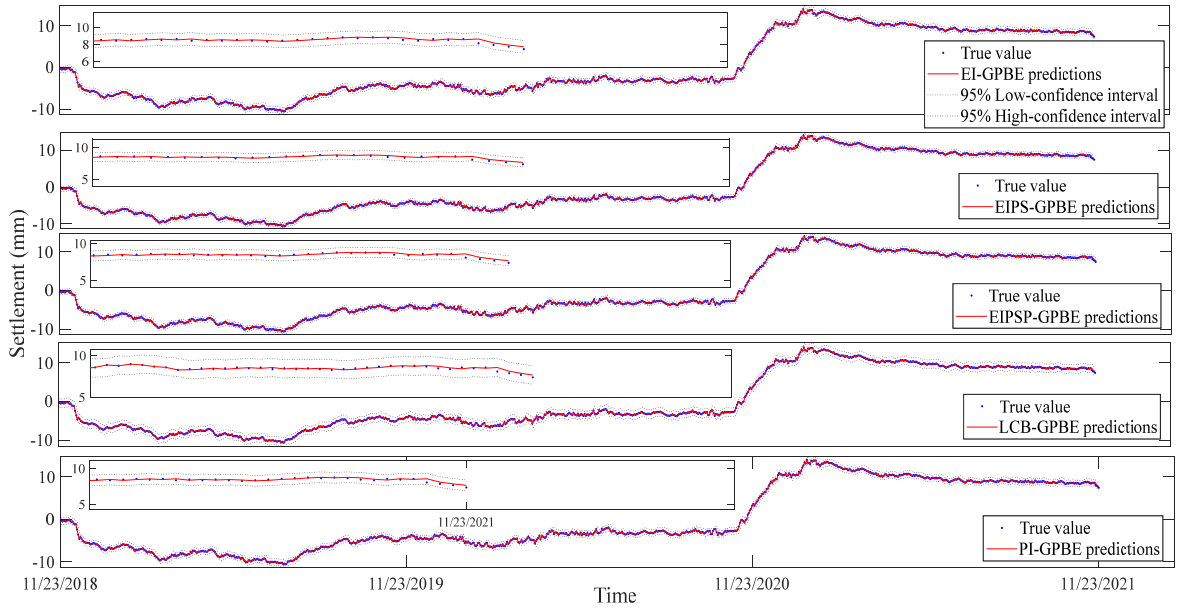


Fig. 6 Prediction results of point J16

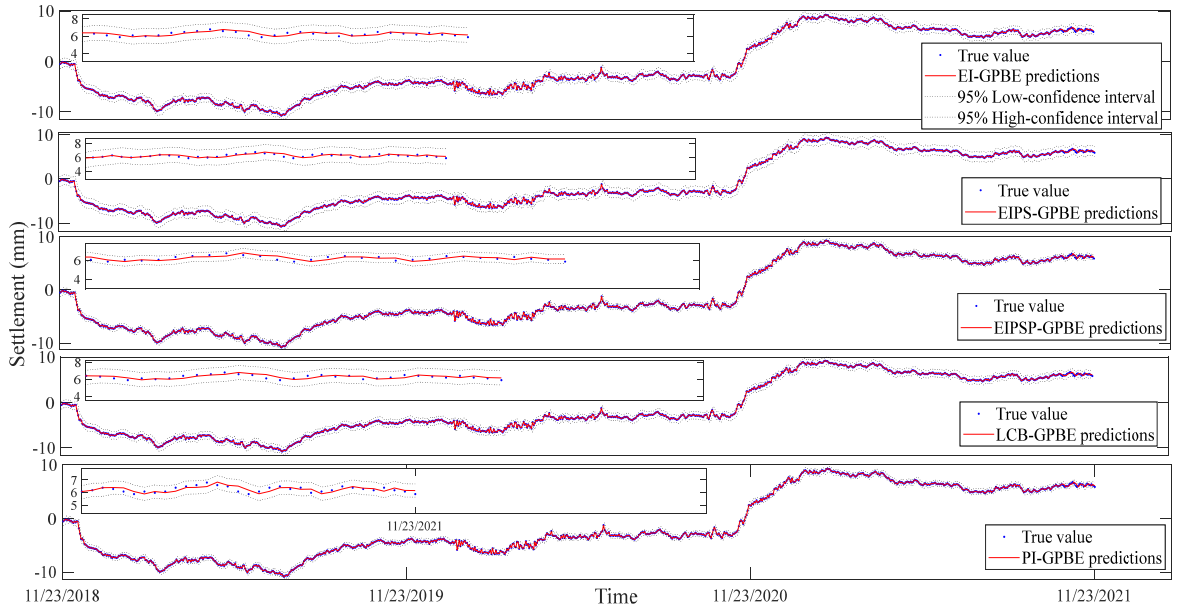


Fig. 7 Prediction results of point J17

that is

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_i - \mu_i)^2} \tag{15}$$

where y_i is the measured settlement and μ_i is the BO-GPBE prediction. Obviously, the smaller the RMSE value, the better the prediction performance.

The RMSE of GPBE prediction results based on different acquisition functions (BO) is shown in Fig. 8. As depicted in Fig. 8, the proposed BO-GPBE prediction model can well predict settlement, and its RMSE value of prediction at each monitoring point does not exceed 0.3. However, the prediction performance of the BO-GPBE model varies for different monitoring data. For instance, for the monitoring points J12-J19, the prediction performance of the established model is relatively poor, mainly because this part of the monitoring points is greatly affected by the construction of new subways, resulting in significant changes in settlement. Specifically, under the EI acquisition function, the RMSE value of J12 monitoring point is 0.212; The RMSE value of J13 monitoring point is 0.191; The RMSE value of J14 monitoring point is 0.136; The RMSE value of J15 monitoring point is 0.193; The RMSE value of J16 monitoring point is 0.162; The RMSE value of J17 monitoring point is 0.211; The RMSE value of J18 monitoring point is 0.165; The RMSE value of J19

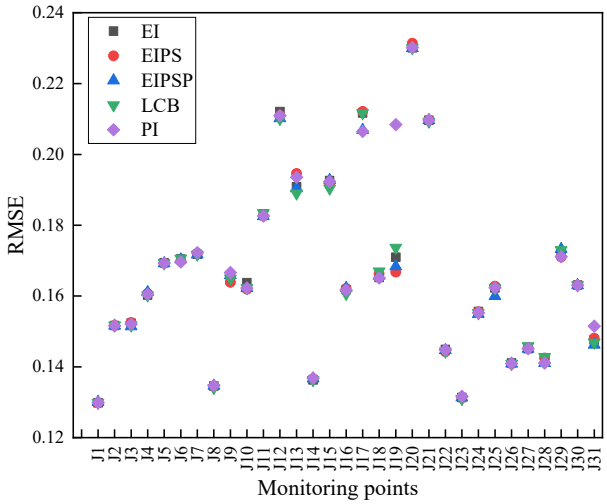


Fig. 8 RMSE of GPBE prediction results based on different acquisition functions

monitoring point is 0.171.

In the contrast, under the EI acquisition function, the RMSE value of J1 monitoring point is 0.130; The RMSE value of J2 monitoring point is 0.152; The RMSE value of J3 monitoring point is 0.152; The RMSE value of J4 monitoring point is 0.160; The RMSE value of J5 monitoring point is 0.170; The RMSE value of J6 monitoring point is 0.170; The RMSE value of J7 monitoring point is 0.172; The RMSE value of J8 monitoring point is 0.135; The RMSE value of J9 monitoring point is 0.164; The RMSE value of J10 monitoring point is 0.164; The RMSE value of J11 monitoring point is 0.180; The RMSE value of J20 monitoring point is 0.23; The RMSE value of J21 monitoring point is 0.210; The RMSE value of J22 monitoring point is 0.145; The RMSE value of J23 monitoring point is 0.131; The RMSE value of J24 monitoring point is 0.156; The RMSE value of J25 monitoring point is 0.162; The RMSE value of J26 monitoring point is 0.141; The RMSE value of J27 monitoring point is 0.145; The RMSE value of J28 monitoring point is 0.143; The RMSE value of J29 monitoring point is 0.172; The RMSE value of J30 monitoring point is 0.163; The RMSE value of J31 monitoring point is 0.148. Obviously, the change of RMSE value of settlement prediction of monitoring points J12-J19 is the largest, while the change of RMSE value of settlement prediction of monitoring points J1-J11 and J20-J31 is relatively small.

Moreover, the different acquisition functions (EI, EIPS, EIPSP, LCB, and PI) have no effect on settlement prediction, which fully demonstrates the robustness of the acquisition function. Specifically, the RMSE value of J12 monitoring point is 0.211, 0.210, 0.210, 0.211 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J13 monitoring point is 0.195, 0.190, 0.189, 0.194 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J14 monitoring point is 0.136, 0.137, 0.136, 0.137 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J15 monitoring point is 0.191, 0.193, 0.190, 0.192 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J16 monitoring point is 0.162, 0.162, 0.161, 0.162 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J17 monitoring point is 0.212, 0.207, 0.217, 0.206 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J18 monitoring point is 0.165, 0.166, 0.167, 0.165 with the EIPS, EIPSP, LCB, PI, separately. The RMSE value of J19 monitoring point is 0.171, 0.167, 0.174, 0.208 with the EIPS, EIPSP, LCB, PI, separately. It can be seen that the established BO-GPBE model is robust and universal, that is, the acquisition function has little effect on the settlement prediction value. Therefore, we can use this model to predict the deformation of other tunnel structures (settlement and float), which has universal applicability.

4. Conclusions

In this paper, this paper proposed a Gaussian Prior Bayesian Emulator (GPBE) probability prediction model based on Bayesian optimization (BO), namely BO-GPBE. We discuss the influence of five different acquisition functions on prediction performance, including expected improvement (EI), expected improvement per second (EIPS), expected improvement per second plus (EIPSP), lower confidence bound (LCB), and probability of improvement (PI). Additionally, we analyze the settlement change of the new metro on the existing metro based on the SHM data. Some conclusions are as follows: (1) From 2019 to 2020 year, the new Metro Line 7 has a significant impact on the settlement of the existing Metro Line 10, resulting in a significant increase in settlement (floating). However, with maintenance during the construction process, the settlement eventually stabilizes. (2) The proposed BO-GPBE model can effectively consider the uncertainty of settlement and predict settlement with good performance, which is verified by the SHM data. (3) Different acquisition functions have little influence on predicting settlement, indicating that the performance of the BO-GPBE model is similar and robustness of the acquisition function.

Acknowledgments

This paper was supported by the Ministry of education of Humanities and Social Science project (No. 23YJCZH037), the Foundation of the State Key Laboratory of Mountain Bridge and Tunnel Engineering (No. SKLBT-2210), China, the Educational Science Planning Project of Zhejiang Province (No. 2023SCG222), and the Scientific Research Project of Zhejiang Provincial Department of Education (No. Y202248682).

Data availability statement

All data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Brochu, E., Cora, V.M. and De Freitas, N. (2010), "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning", *Comput. Sci.*, 1012.2599.

Bull, A.D. (2011), "Convergence rates of efficient global optimization algorithms", *Mach. Learn.*, ***12***(10), arXiv: 1101.3501.

Bullock, Z., Karimi, Z., Dashti, S., Porter, K., Liel, A.B. and Franke, K.W. (2019), "A physics-informed semi-empirical probabilistic model for the settlement of shallow-founded structures on liquefiable ground", *Géotechnique*, **69**(5), 406-419. https://doi.org/10.1680/jgeot.17.P.174

Chen, Z. and Sun, H. (2021), "Sparse representation for damage identification of structural systems", *Struct. Health Monit.*, **20**(4), 1644-1656. https://doi.org/10.1177/1475921720926970

Deng, H.S., Fu, H.L., Yue, S., Huang, Z. and Zhao, Y.Y. (2022), "Ground loss model for analyzing shield tunneling-induced surface settlement along curve sections", *Tunn. Underg. Space Technol.*, **119**, 104250. https://doi.org/10.1016/j.tust.2021.104250

Ding, Y., Ye, X.W. and Guo, Y. (2023a), "Wind load assessment with the JPDF of wind speed and direction based on SHM data", *Structures*, **47**(1), 2074-2080. https://doi.org/10.1016/j.istruc.2022.12.028

Ding, Y., Ye, X.W. and Guo, Y. (2023b), "Data set from wind, temperature, humidity and cable acceleration monitoring of the Jiashao bridge", *J. Civil Struct. Health Monit.*, **13**(2-3), 579-589. https://doi.org/10.1007/s13349-022-00662-5

Ding, Y., Ye, X.W. and Guo, Y. (2023c), "Copula-based JPDF of wind speed, wind direction, wind angle, and temperature with SHM data", *Probab. Eng. Mech.*, **73**, 103483. https://doi.org/10.1016/j.probengmech.2023.103483

Ding, Y., Ye, X.W., Guo, Y., Zhang, R. and Ma, Z. (2023d), "Probabilistic method for wind speed prediction and statistics distribution inference based on SHM data-driven", *Probab. Eng. Mech.*, **73**, 103475. https://doi.org/10.1016/j.probengmech.2023.103475

Ding, Y., Ye, X.W. and Guo, Y. (2023e), "A multistep direct and indirect strategy for predicting wind direction based on the EMD-LSTM model", *Struct. Control Health Monit.*, 4950487. https://doi.org/10.1155/2023/4950487

Ding, Y., Ye, X.W., Su, Y.H. and Zheng, X.L. (2023f), "A framework of cable wire failure mode deduction based on Bayesian network", *Structures*, **57**, 104996. https://doi.org/10.1016/j.istruc.2023.104996

Ding, Y., Hang, D., Wei, Y.J., Zhang, X.L., Ma, S.Y., Liu, Z.X., Zhou, S.X. and Han, Z. (2023g), "Settlement prediction of existing metro induced by new metro construction with machine learning based on SHM data: a comparative study", *J. Civil Struct. Health Monit.*, **13**(6-7), 1447-1457. https://doi.org/10.1007/s13349-023-00714-4

Ding, Y., Ye, X.W., Ding, Z., Wei, G., Cui, Y.L., Han, Z. and Jin, T. (2023h), "Short-term tunnel-settlement prediction based on Bayesian wavelet: a probability analysis method", *J. Zhejiang Univ.-SCI A*, **24**(11), 960-977. https://doi.org/10.1631/jzus.A2200599

Fang, K., Yang, Z., Jiang, Y., Sun, Z. and Wang, Z. (2020), "Surface subsidence characteristics of fully overlapping tunnels constructed using tunnel boring machine in a clay stratum", *Comput. Geotech.*, **125**, 103679. https://doi.org/10.1016/j.compgeo.2020.103679

Gelbart, M.A., Snoek, J. and Adams, R.P. (2014), "Bayesian optimization with unknown constraints", *Mach. Learn.*, arXiv: 1403.5607.

Gong, W., Luo, Z., Juang, C.H., Huang, H., Zhang, J. and Wang, L. (2014), "Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays", *Comput. Geotech.*, **56**, 69-79. https://doi.org/10.1016/j.compgeo.2013.10.008

Han, Q., Ni, P.H., Du, X.L., Zhou, H. and Cheng, X. (2022), "Computationally efficient Bayesian inference for probabilistic model updating with polynomial chaos and Gibbs sampling", *Struct. Control Health Monit.*, **29**(6), e2936. https://doi.org/10.1002/stc.2936

Huang, H., Jia, R., Shi, X., Liang, J. and Dang, J. (2021), "Feature selection and hyper parameters optimization for short-term wind power forecast", *Appl. Intell.*, **51**(10), 6752-6770. https://doi.org/10.1007/s10489-021-02191-y

Jones, D.R., Schonlau, M. and Welch, W.J. (1998), "Efficient global optimization of expensive black-box functions", *J. Global Optim.*, **13**(4), 455-492. https://doi.org/10.1023/A:1008306431147

Li, B. and Wang, Z.Z. (2019), "Numerical study on the response of ground movements to construction activities of a metro station using the pile-beam-arch method", *Tunn. Underg. Space Technol.*, **88**, 209-220. https://doi.org/10.1016/j.tust.2019.03.014

Li, W., Feng, W. and Yuan, H.Z. (2020), "Multimode traffic travel behavior characteristics analysis and congestion governance research", *J. Adv. Transp.*, **5**, 1-8. https://doi.org/10.1155/2020/6678158

Liu, T., Wei, H., Liu, S. and Zhang, K. (2020a), "Industrial time series forecasting based on improved Gaussian process regression", *Soft Comput.*, **24**(20), 15853-15869. https://doi.org/10.1007/s00500-020-04916-6

Liu, B., Zhang, D.W., Yang, C. and Zhang, Q.B. (2020b), "Long-term performance of metro tunnels induced by adjacent large deep excavation and protective measures in Nanjing silty clay", *Tunn. Underg. Space Technol.*, **95**, 103147. https://doi.org/10.1016/j.tust.2019.103147

Luat, N.V., Nguyen, V.Q., Lee, S., Woo, S. and Lee, K. (2020), "An evolutionary hybrid optimization of MARS model in predicting settlement of shallow foundations on sandy soils", *Geomech. Eng., Int. J.*, **21**(6), 583-598. https://doi.org/10.12989/gae.2020.21.6.583

Madra, A., Causse, P., Trochu, F., Adrien, J., Maire, E. and Breitkopf, P. (2019), "Stochastic characterization of textile reinforcements in composites based on X-ray microtomographic scans", *Compos. Struct.*, **224**, 111031. https://doi.org/10.1016/j.compstruct.2019.111031

Miliziano, S. and de Lillis, A. (2019), "Predicted and observed settlements induced by the mechanized tunnel excavation of metro line C near S. Giovanni station in Rome", *Tunn. Underg. Space Technol.*, **86**, 236-246. https://doi.org/10.1016/j.tust.2019.01.022

Moaveni, B. and Najafi, S. (2017), "Metro traffic modeling and regulation in loop lines using a robust model predictive controller to improve passenger satisfaction", *IEEE Transact. Control Syst. Technol.*, **26**(5), 1541-1551. https://doi.org/10.1109/tcst.2017.2735945

Mu, B., Xie, X., Li, X., Li, J., Shao, C. and Zhao, J. (2021), "Monitoring, modelling and prediction of segmental lining deformation and ground settlement of an EPB tunnel in different soils", *Tunn. Underg. Space Technol.*, **113**, 103870. https://doi.org/10.1016/j.tust.2021.103870

Ni, Y.Q., Wang, Y.W. and Zhang, C. (2020), "A Bayesian approach for condition assessment and damage alarm of bridge expansion joints using long-term structural health monitoring data", *Eng. Struct.*, **212**, 110520. https://doi.org/10.1016/j.engstruct.2020.110520

Ni, P.H., Li, Q., Han, Q., Xu, K. and Du, X.L. (2023), "Substructure approach for Bayesian probabilistic model updating using response reconstruction technique", *Mech. Syst. Signal Proc.*, **183**, 109624. https://doi.org/10.1016/j.ymssp.2022.109624

Qiu, J., Qin, Y., Feng, Z., Wang, L. and Wang, K. (2020), "Safety risks and protection measures for city wall during construction and operation of Xi'an Metro", *J. Perform. Constr. Fac.*, **34**(2), 04020003. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001374

Qu, K., Xu, Y.Y. and Huang, J.X. (2023), "Numerical simulation of hydrodynamic characteristics of submerged floating tunnels under the action of focused waves", *J. Changsha Univ. Sci. Techn.* (*Natural Science*), (04):127-141. [In Chinese] https://doi.org/10.19951/j.cnki.1672-9331.20220425001

Shi, S., Zhao, R., Li, S., Xie, X., Li, L., Zhou, Z. and Liu, H. (2019), "Intelligent prediction of surrounding rock deformation of shallow buried highway tunnel and its engineering application", *Tunn. Underg. Space Technol.*, **90**, 1-11. https://doi.org/10.1016/j.tust.2019.04.013

Snoek, J., Larochelle, H. and Adams, R.P. (2012), "Practical bayesian optimization of machine learning algorithms", *Mach. Learn.*, arXiv:1206.2944.

Tan, L.S., Ong, V.M., Nott, D.J. and Jasra, A. (2016), "Variational inference for sparse spectrum Gaussian process regression", *Statis. Comput.*, **26**(6), 1243-1261. https://doi.org/10.1007/s11222-015-9600-7

Tu, H., Zhou, H., Qiao, C. and Gao, Y. (2020), "Excavation and kinematic analysis of a shallow large-span tunnel in an up-soft/low-hard rock stratum", *Tunn. Underg. Space Technol.*, **97**, 103245. https://doi.org/10.1016/j.tust.2019.10324

Vereecken, E., Botte, W., Lombaert, G. and Caspeele, R. (2020), "Bayesian decision analysis for the optimization of inspection and repair of spatially degrading concrete structures", *Eng. Struct.*, **220**, 111028. https://doi.org/10.1016/j.engstruct.2020.111028

Wan, H.P. and Ni, Y.Q. (2018), "Bayesian modeling approach for forecast of structural stress response using structural health monitoring data", *J. Struct. Eng.*, **144**(9), 04018130. https://doi.org/10.1061/(ASCE)ST.1943-541X.0002085

Wan, H.P. and Ni, Y.Q. (2019), "Bayesian multi-task learning methodology for reconstruction of structural health monitoring data", *Struct. Health Monit.*, **18**(4), 1282-1309. https://doi.org/10.1177/1475921718794953

Wang, F., Gou, B. and Qin, Y. (2013), "Modeling tunneling-induced ground surface settlement development using a wavelet smooth relevance vector machine", *Compu. Geotech.*, **54**, 125-132. https://doi.org/10.1016/j.compgeo.2013.07.004

Wang, Z.L., Ogawa, T. and Adachi, Y. (2019), "Influence of algorithm parameters of Bayesian optimization, genetic algorithm, and particle swarm optimization on their optimization performance", *Adv. Theory Simulat.*, **2**(10), 1900110. https://doi.org/10.1002/adts.201900110

Ye, X.W., Ding, Y. and Wan, H.P. (2019), "Machine learning approaches for wind speed forecasting using long-term monitoring data: a comparative study", *Smart Struct. Syst., Int. J.*, **24**(6), 733-744. https://doi.org/10.12989/sss.2019.24.6.733

Ye, X.W., Ding, Y. and Wan, H.P. (2020), "Statistical evaluation of wind properties based on long-term monitoring data", *J. Civ. Struct. Health Monit.*, **10**(5), 987-1000. https://doi.org/10.1007/s13349-020-00430-3

Ye, X.W., Ding, Y. and Wan, H.P. (2021), "Probabilistic forecast of wind speed based on Bayesian emulator using monitoring data", *Struct. Control Health Monit.*, **28**(1), e2650. https://doi.org/10.1002/stc.2650

Zhang, S., Shang, C., Wang, C., Song, R. and Wang, X. (2019), "Real-time safety risk identification model during metro construction adjacent to buildings", *J. Construt. Eng. Manag.*, **145**(6), 04019034. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001657

Zhang, K., Lyu, H.M., Shen, S.L., Zhou, A. and Yin, Z.Y. (2020a), "Evolutionary hybrid neural network approach to predict shield tunneling-induced ground settlements", *Tunn. Underg. Space Technol.*, **16**, 103594. https://doi.org/10.1016/j.tust.2020.103594

Zhang, P., Wu, H.N., Chen, R.P. and Chan, T.H. (2020b), "Hybrid meta-heuristic and machine learning algorithms for tunneling-induced settlement prediction: A comparative study", *Tunn. Underg. Space Technol.*, **99**, 103383. https://doi.org/10.1016/j.tust.2020.103383

Zhang, W., Zhao, M., Du, X., Gao, Z. and Ni, P. (2023), "Probabilistic machine learning approach for structural reliability analysis", *Probab. Eng. Mech.*, **74**, 103502. https://doi.org/10.1016/j.probengmech.2023.103502