Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering) 2024 25(9):732-748
www.jzus.zju.edu.cn; www.springer.com/journal/11582
E-mail: jzus_a@zju.edu.cn

Research Article

https://doi.org/10.1631/jzus.A2300441

Temperature field prediction of steel-concrete composite decks using TVFEMD-stacking ensemble algorithm

Benkun TAN[1,2]✉, Da WANG[1,3], Jialin SHI[3], Lianqi ZHANG[1]

[1]*School of Civil Engineering, Changsha University of Science & Technology, Changsha 410114, China*
[2]*School of Civil and Architecture Engineering, Hunan University of Arts and Science, Changde 415000, China*
[3]*School of Civil Engineering, Central South University of Forestry and Technology, Changsha 410004, China*

Abstract: This research aims to develop an advanced deep learning-based ensemble algorithm, utilizing environmental temperature and solar radiation as feature factors, to conduct hourly temperature field predictions for steel-concrete composite decks (SCCDs). The proposed model comprises feature parameter lag selection, two non-stationary time series decomposition methods (empirical mode decomposition (EMD) and time-varying filtering-based empirical mode decomposition (TVFEMD)), and a stacking ensemble prediction model. To validate the proposed model, five machine learning (ML) models (random forest (RF), support vector regression (SVR), multilayer perceptron (MLP), gradient boosting regression (GBR), and extreme gradient boosting (XGBoost)) were tested as base learners and evaluations were conducted within independent, mixed, and ensemble frameworks. Finally, predictions are made based on engineering cases. The results indicate that consideration of lag variables and modal decomposition can significantly improve the prediction performance of learners, and the stacking framework, which combines multiple learners, achieves superior prediction results. The proposed method demonstrates a high degree of predictive robustness and can be applied to statistical analysis of the temperature field in SCCDs. Incorporating time lag features helps account for the delayed heat dissipation phenomenon in concrete, while decomposition techniques assist in feature extraction.

Key words: Steel-concrete composite deck (SCCD); Temperature field; Time-varying filtering-based empirical mode decomposition (TVFEMD); Feature selection; Machine learning (ML)

1 Introduction

Steel-concrete composite beam bridges are widely used due to their competitive advantages in many aspects. They consist of steel girders and concrete flange slabs that are joined together using shear connectors (Shim et al., 2001). However, the non-uniform distribution of temperature across the cross-section caused by environmental temperature fluctuations can lead to thermal stress, which may affect the structural performance and service life of such bridges (Wang et al., 2021). It can cause cracking of concrete bridge decks and bending of steel girders (Branco and Mendes, 1993; Giussani, 2009), ultimately decreasing structural performance and posing a threat to structural safety and durability (Luo et al., 2023; Chen et al., 2024). Therefore, the study of temperature fields in steel-concrete composite beam bridges has received widespread attention.

The thermal effects of bridges are influenced by various meteorological factors, including atmospheric temperature and solar radiation (Zhang et al., 2022). Much research has been carried out on temperature distribution in bridge structures based on meteorological data, focusing on extreme temperature distribution in different seasonal patterns and the corresponding thermal behavior (Tong et al., 2002; Lee, 2012). Changes in the structural temperature field can lead to alterations in the structural dynamic characteristics which affect the displacement and strain responses in structural health monitoring (SHM) (Sohn et al., 1999; Innocenzi et al., 2022; Nicoletti et al., 2023). Once the

✉ Benkun TAN, tanbenkun@yeah.net
Benkun TAN, https://orcid.org/0000-0002-0949-9288
Jialin SHI, https://orcid.org/0009-0003-8616-8559

Received Aug. 25, 2023; Revision accepted Dec. 20, 2023;
Crosschecked Sept. 4, 2024

service scenario of a bridge structure is determined, a thermal load model can typically be established by collecting temperature data and conducting statistical analysis (Catbas et al., 2008). Based on this, further investigation into the extreme temperature distribution in different seasons and corresponding thermal response behavior can be carried out (Zhang et al., 2022). To obtain the required temperature data, on-site testing and monitoring are common methods. As SHM systems continue to evolve, researchers have proposed temperature gradient models for various regions and bridge types by collecting data through thermal sensors on bridges (Han et al., 2021). However, this method requires a sufficiently large sample size, which means that the required time cost is high (Zhang et al., 2023). Moreover, the maintenance of equipment and data loss are also issues that researchers face (Xin et al., 2023).

Predicting temperature fields of bridge structures under complex environmental conditions is a major focus of research. While standard values of temperature effects obtained through long-term measurements and statistical analysis can cover the most adverse temperature patterns during the design period, a single temperature gradient model may not adapt well to changing meteorological conditions. Some researchers have proposed empirical theoretical formulas based on fitting the relationship between solar radiation, wind speed, atmospheric temperature, and temperatures at different locations on the structure for temperature prediction (Qin and Hiller, 2011). However, the responses of concrete to atmospheric temperature and radiation are not synchronously produced at different depths and, due to temperature hysteresis, the temperature of concrete at other locations lags behind that of its surface. Moreover, the heat transfer coefficient of steel is higher than that of concrete, resulting in temperature gradients along the height direction of the cross-section. When meteorological conditions are complex, it is difficult to describe the heat transfer mode in a simple fitting formula. Finite element (FE) methods are often used for simulating temperature fields of bridge structures. The Fourier heat transfer equation is an effective tool for simulating heat conduction behavior and can be used to solve transient analysis of complex temperature fields with refined FE models and accurately calibrated thermal parameters (Zhang et al., 2020). However, this method requires significant computational resources and has low computational efficiency, making it challenging to apply it to the analysis of massive historical data for temperature load calculations in design (Fan et al., 2022). With the initial exploration of digital twin technology in the field of bridges and the development of SHM techniques (Opoku et al., 2021; Broo et al., 2022), maintenance of large-span bridges is gradually moving towards real-time monitoring (Zhao et al., 2023), requiring a real-time and accurate temperature field prediction method (Fu et al., 2022). Some researchers have suggested that future bridge design and maintenance decisions should take into account climate characteristics (Figueiredo et al., 2023). Machine learning (ML) methods have become an effective approach to solving this problem, but only a few studies have been devoted to their application in temperature prediction of steel-concrete composite decks (SCCDs) (Han et al., 2021). However, ML methods have achieved encouraging results in other engineering fields (Flah et al., 2021; Wedel and Marx, 2022).

In summary, accurate prediction of the temperature field in SCCDs under environmental influences is fundamental to analyzing their thermal response. However, existing methods based on health monitoring and numerical simulation have drawbacks such as high maintenance costs of equipment and low computational efficiency. Additionally, prediction methods based on single ML models tend to have weak parameter sensitivity or poor generalization capabilities. Recognizing these limitations, this paper aims to establish a new predictive model for improving the efficiency and accuracy of temperature field predictions for SCCDs. In this study, a stacking ensemble algorithm was developed for predicting the temperature field of SCCDs. The model's data samples were expanded using a numerical analysis model, which was validated with SHM data. Non-stationary time series were decomposed to capture input features, and the importance of these features was assessed, taking into account the heat transfer lag at different positions. Comparative analysis was conducted on the predictive performance of multiple single ML models versus the stacking model. Additionally, the statistical accuracy of temperature sample data predicted by the stacking model was evaluated.

2 Methodology

2.1 Numerical temperature field simulation

Fig. 1 shows that the temperature field of SCCDs can be simulated by inputting various thermal parameters and specifying the thermal boundary conditions described earlier (Zhang et al., 2020; Sheng et al., 2022). Transient thermal analysis requires time and space discretization, followed by the use of a numerical algorithm to obtain a solution. The implicit algorithm has a time step that is not limited by time stability, resulting in higher calculation accuracy and stability. The solution's efficiency is improved by using Newton's method to solve the nonlinear system of equations at each time step. The formulas for heat conduction theory involved in this study are shown in Section S1 of the electronic supplementary materials (ESM) (Branco and Mendes, 1993; Narasimhan, 1999; Shi et al., 2022). To solve the Fourier heat transfer equation, initial conditions must also be defined. The FE model is processed cyclically, with an additional 24 h cycle, until the temperature at each node at the end of the incremental step on day n matches the temperature at the end of the incremental step on day (n+1). The temperature field at the end of the incremental step on day (n+1) is then used as the initial temperature field for calculating the annual temperature field of the numerical model.

2.2 Feature importance scoring

Prior to using an ML algorithm, it is frequently necessary to perform feature selection. In this study, an algorithm based on the decision tree principle was used to determine the importance of input features (Sugumaran et al., 2007). To analyze the time lag effect of ambient temperature on structural temperature, the importance of different time lag quantities can be scored using an ML model with a decision tree function in which two methods can be used to calculate the importance of features: the information entropy approach and the Gini coefficient-based approach. The information entropy method is based on the concept of information gain, which is used to measure the contribution of features in the classification, while the Gini coefficient method is based on the concept of Gini impurity, which is used to measure the error rate of features in the classification (Liu et al., 2021).

2.3 Decomposition and reconstitution of non-stationary time series

2.3.1 Empirical mode decomposition (EMD)

The adaptive decomposition of non-linear and non-stationary time series into multiple intrinsic modal functions (IMFs) and a residual term is achieved using EMD, a commonly employed signal decomposition method (Boudraa and Cexus, 2007). The number of IMFs depends on the time series characteristics. For a temperature time series $T(t)$, this can be expressed as:

$$T(t)=\sum_{i=1}^{J}c^{i}(t), \tag{1}$$

where $c^i(t)$ denotes the ith IMF component and the last term is referred to as the residual term. J represents

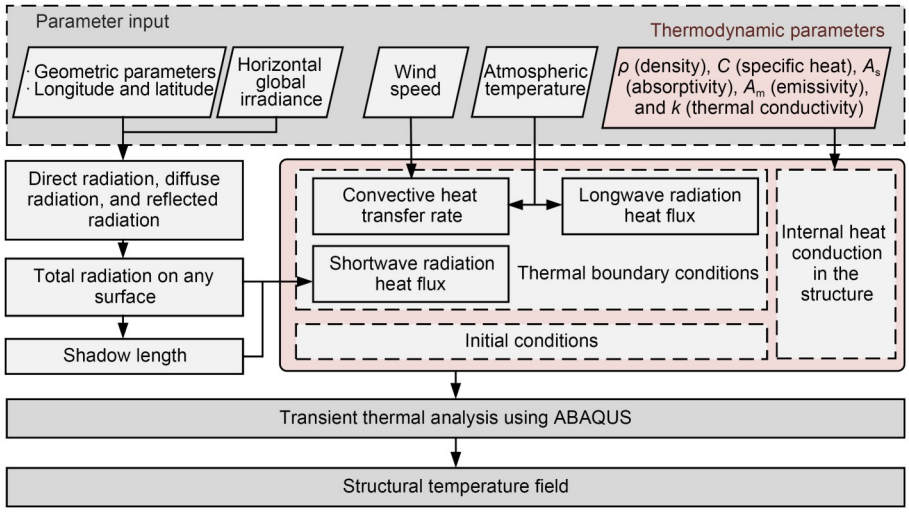

Fig. 1 Numerical simulation flow diagram of the temperature field

the total number of IMF components obtained from the decomposition.

2.3.2 Time-varying filtering-based empirical mode decomposition (TVFEMD)

TVFEMD is a variant of the EMD method that constructs a time-varying filter in each local region of the original signal to capture the characteristics of local frequency modulation (Jamei et al., 2023). The time-varying filter is constructed based on local extremum points, and B-sample interpolation is used to smooth the response curve of the filter. This allows the signal to be decomposed into multiple IMFs. The steps for implementing the TVFEMD method are shown in Section S2 of the ESM (Xin et al., 2023).

2.3.3 Sample entropy (SE)

The SE method is a technique proposed for detecting similarity between two different time series (Richman and Moorman, 2000). It quantifies the degree of similarity between the time series without making prior assumptions about the shape of the data distribution. The principle is based on comparing the distances between vectors of adjacent time points, and thus calculating the value of sample entropy, which can be used as a basis for recombining the time series components obtained from EMD and TVFEMD decomposition. Suppose there is a set of time series $X=[x(1), x(2), \cdots, x(N)]$ of length N. The steps for calculating the SE value are shown in Section S3 of the ESM.

2.4 ML methods

2.4.1 Random forest (RF)

The RF algorithm involves training several decision trees independently, each on a different set of randomly selected features and data samples, avoiding overfitting and learning different features and patterns. RF combines the predictions of all decision trees to obtain the final prediction. The classification and regression tree (CART) algorithm is usually used to train decision trees, where the Gini impurity or information gain is used as the classification criterion. RF improves performance and generalization capabilities by combining the prediction results of multiple decision trees (Liu et al., 2021).

2.4.2 Support vector regression (SVR)

SVR is based on the principle of support vector machines (SVMs). The core idea is to divide the data points into two classes by constructing an optimal hyperplane such that the distance between the two sides of the hyperplane is maximized, i.e., the interval is maximized. In SVR, we use the same idea but transform the hyperplane into a regression function that is used to fit the data points so that the prediction error is minimized (Liu et al., 2021).

$$\boldsymbol{y}=\boldsymbol{\beta}^{\mathrm{T}}\phi(\boldsymbol{x})+b, \tag{2}$$

where $\boldsymbol{x}$ represents the feature vector of the input data, and $\phi(\boldsymbol{x})$ is the transformation of the input data into a vector in a high-dimensional space using a mapping function. The normal vector of the hyperplane is represented by $\boldsymbol{\beta}$, and b is the intercept of the hyperplane. The corresponding prediction result is denoted by $\boldsymbol{y}$.

2.4.3 Multilayer perceptron (MLP)

MLP is a relatively common type of feed-forward neural network used to solve a variety of ML problems (Nguyen et al., 2021). MLP typically consists of multiple fully connected layers, where each neuron is connected to all neurons in the previous layer, with certain weights and bias values, and is used to compute a linear combination of input data. These linear combinations are transformed by the activation function and used as input to the next layer, which undergoes multiple layers of non-linear transformations to obtain the final output of the model. Suppose we have an input vector $\boldsymbol{x}\in\mathbb{R}^m$, an output vector $\boldsymbol{y}\in\mathbb{R}^m$, and a set of neurons with weights and bias values. For the jth neuron in layer i, its input $z_{i,j}$ and output $a_{i,j}$ can be expressed as follows, respectively:

$$z_{i,j}=\sum_{k=1}^{n_{i-1}} w_{i,j,k}a_{i-1,k}+b_{i,j}, \tag{3}$$

$$a_{i,j}=f(z_{i,j}), \tag{4}$$

where $w_{i,j,k}$ is the weight between the jth neuron in layer i and the kth neuron in layer $i-1$, $b_{i,j}$ is the bias of the jth neuron in layer i, n_{i-1} is the total number of neurons in layer $i-1$, and f is the activation function, which can be a Sigmoid function, rectified linear unit (ReLU) function, Tanh function, etc.

2.4.4 Gradient boosting regression (GBR)

GBR is an ML method that utilizes gradient boosting tree (GBT) to solve regression problems (Nguyen et al., 2021). Developed by Friedman, this ensemble learning model is designed to create a highly accurate strong learner by combining a sequence of weak or base learners, each of which performs better than random guessing. The model achieves this by adding new weak learners to minimize the total loss (Friedman, 2001). The gradient boosting algorithm works in iterations, where each iteration adds a new weak learner, typically in the form of decision trees, to the model. In the first iteration, the algorithm learns the first weak learner, which is the first tree, to reduce the overall training error. In the subsequent iterations, the algorithm learns additional trees to reduce the errors of the previously added trees. The algorithm repeats this process until a high-quality model is constructed, which typically involves minimizing the model's loss, also known as the overall error, to a desired level.

2.4.5 Extreme gradient boosting (XGBoost)

XGBoost is a highly scalable extension of the GBT algorithm, which has been used to handle regression tasks (Nguyen et al., 2021; Wang et al., 2022). For a data set $\boldsymbol{D}=[(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_n, y_n)]$, the model is initialized and can be set to the average of the training data, as shown in Eq. (5):

$$\hat{y}_i^{(0)}=\frac{1}{n}\sum_{i=1}^{n}y_i, \tag{5}$$

where $\boldsymbol{x}_i$ denotes the feature of the ith sample, y_i' is its prediction target value, and n is the number of samples in the training dataset, and $\hat{y}_i^{(0)}$ is the initial prediction of the ith sample. The specific derivation process is presented in Section S4 of the ESM.

2.4.6 Elastic net

The elastic net method is a linear regression method that combines L1 and L2 regularization terms (Zou and Hastie, 2005). The L1 regularization term generates sparsity in the regression coefficients, while the L2 regularization term generates smoothness in the regression coefficients. Typically, regularization methods take the following form:

$$\arg\min_{\boldsymbol{w}}\left\{L(\boldsymbol{w};\boldsymbol{X},\boldsymbol{y})+\alpha R(\boldsymbol{w})\right\}, \tag{6}$$

where $L(\boldsymbol{w};\boldsymbol{X},\boldsymbol{y})$ is the loss function, and $R(\boldsymbol{w})$ is the regularization term, which is used to control the complexity of the model and avoid overfitting. $\boldsymbol{w}$ denotes the parameter vector of the model, $\boldsymbol{X}$ denotes the feature matrix of the training data, and $\boldsymbol{y}$ denotes the prediction result of the training data. α is the regularization coefficient, which is used to control the importance of the regularization term.

Different regularization methods differ in the exact form of $R(\boldsymbol{w})$, but they all have the general form described above, and, by adjusting the magnitude of the regularization coefficient α, the fit and generalization ability of the model can be balanced. Elastic net is a regularization method based on square loss, and the optimal regression coefficient $\boldsymbol{w}^*$ in the method can be obtained by solving this optimization problem as follows:

$$\boldsymbol{w}^*=\arg\min_{\boldsymbol{w}}\left[\frac{1}{2}\sum_{i=1}^{n}(y_i-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i)^2+\lambda_1\|\boldsymbol{w}\|_1+\lambda_2\|\boldsymbol{w}\|_2^2\right], \tag{7}$$

where $\boldsymbol{x}_i$ denotes the eigenvector of the ith sample, λ_1 and λ_2 denote the L1 and L2 regularization coefficients, respectively, $\|\cdot\|_1$ denotes the L1 norm, and $\|\cdot\|_2$ denotes the L2 norm.

2.5 Intelligent expert framework establishment

2.5.1 Data feature processing

Prior to building an ML model, data preprocessing is necessary. In this study, the main steps of preprocessing include calculating time delays using the feature importance scoring function based on the XGBoost model, feature decomposition using EMD and TVFEMD methods, and recombination and dimensionality reduction of the decomposed features using SE.

SCCD's temperature field data, in addition to being influenced by its own structural parameters, is also affected by environmental factors such as atmospheric temperature and solar radiation. As the weather changes, the temperature gradient of the structure presents different patterns. For example, on sunny days with strong sunlight, the structure presents a positive temperature gradient pattern while, when the temperature drops significantly, the structure presents

a negative temperature gradient pattern because the surface in contact with the air and steel beams has a slower heat dissipation rate than other areas. The temperature gradient pattern is mainly influenced by the ambient temperature. The effect of ambient temperature on the structure has a certain time delay characteristic due to the different heat exchange and dissipation times at different positions. Therefore, it is necessary to analyze the influence of time lag (Liu et al., 2020). The ambient temperature has obvious periodic features that vary with day and night and season. However, because the ambient temperature is a non-stationary time series, simple ML models find it difficult to distinguish trends, periodicities, or seasonal changes in the sequence. Decomposition of the original sequence is thus necessary to address the issue of increased feature dimensions after decomposition; dimensionality reduction methods can be used to recombine some features. By combining these methods, the efficiency and accuracy of ML can be improved.

2.5.2 Stacking ensemble learning framework

The stacking ensemble learning framework is a method that integrates the prediction results of multiple base learners through a single meta-model. Its core principle is to improve prediction accuracy and stability by leveraging the complementarity between different base models. When using the stacking ensemble learning framework, it is necessary to select appropriate base models and meta-models, which represent a single model for training and prediction and a model that incorporates the prediction results of a single model, respectively. The selection of base models should be based on their predictive ability and stability. For the meta-model, its complexity and generalization ability should be considered. Here, RF, SVR, MLP, GBR, and XGBoost were selected as the base models. The elastic net model was chosen as the meta-model to learn respective features from different base models.

In addition, when using the stacking ensemble learning framework, cross-validation methods are used to ensure the performance of the models. Typically, the k-fold cross-validation method is used to divide the dataset into k mutually exclusive subsets. Then, each subset is used in turn as the test set and the remaining $k-1$ subsets are used as the training set. The resulting model is trained to make predictions, and error metrics are calculated. Finally, all the error metrics are averaged to obtain the performance evaluation metrics of the model. This prevents the model from overfitting to a particular training set, thus improving its generalization ability.

The summary of the model-building process is as follows: the dataset is divided into a training set and a test set. Multiple base models are trained using the training set, and multiple predictions are made on the test set. The predictions from the base models are then used to train a meta-model, which can be used to make predictions on the test data. Figs. 2 and 3 show the flow chart of the k-fold cross-validation of a single model and the framework of the stacking ensemble model, respectively. To begin, the data samples are prepared, and the sample set is divided into a training set and a test set. The features are normalized. For a single model, such as the base model Model 1, the training set is cross-validated in k folds. During the cross-validation process, the hyperparameters are optimized using the random grid search method. The prediction results of the training set, $P_{\text{train},i}$ (i=1, 2, ⋯, n), and the prediction results of the test set, $P_{\text{test},j}$ (j=1, 2, ⋯, n), are recorded in each fold, respectively. Then, the $P_{\text{train},i}$ obtained in each fold are combined to form a new training set (NewTrain_Model 1), and the $P_{\text{test},j}$ obtained in each fold are averaged to obtain a new test set (NewTest_Model 1). This operation is repeated m times, where m is the number of base models selected. As a result, the set of m NewTrain_Model k and m NewTest_Model k (k=1, 2, ⋯, m) can be obtained. The NewTrain_Model k is trained with the prediction labels (TrainSet_Label) in the training set using the meta-model as the input and target values to obtain the trained stacking ensemble model. Finally, the NewTest_Model k is input to the trained model to obtain the final prediction results of the test set.

2.5.3 Performance assessment indicators

This study employs three metrics to evaluate the predictive performance of ML models, including the mean absolute error (MAE), coefficient of determination (R^2), and root mean squared error (RMSE); higher R^2 values and smaller MAE and RMSE values indicate better model predictions. The above indicators are defined as follows:

$$R^2=1-\frac{\sum_{i=1}^{n}\left(y_i'-\hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i'-\bar{y}\right)^2}, \tag{8}$$

$$E_{\mathrm{MA}}=\frac{1}{n}\sum_{n=1}^{n}|\,y_i'-\hat{y}_i\,|, \tag{9}$$

$$E_{\mathrm{RMS}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i'-\hat{y}_i)^2}, \tag{10}$$

where E_{MA} is the MAE, E_{RMS} is the RMSE, n is the number of samples, y_i' is the measured value of the ith sample data, $\hat{y}_i$ represents the predicted value of the ith sample data, and $\bar{y}$ represents the average of all sample data.

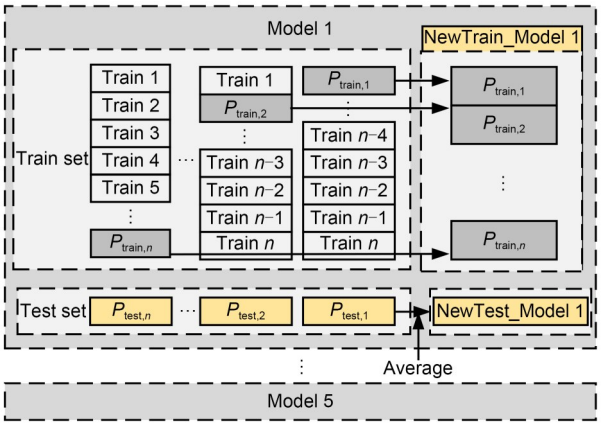

Fig. 2 Flow chart of the cross-validation of a single model

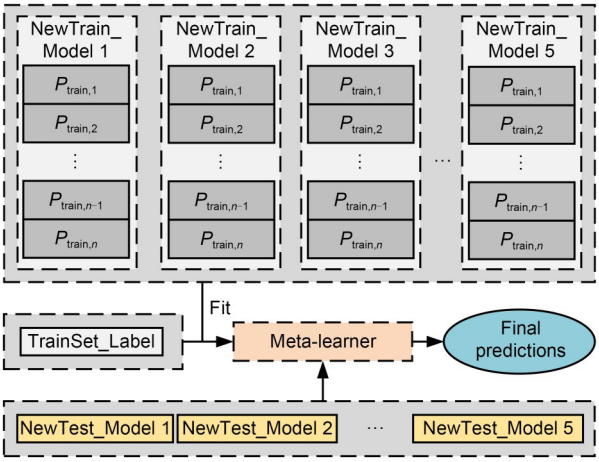

Fig. 3 Framework of the stacking ensemble model

3 Numerical analysis and validation of temperature field

3.1 Engineering background

A large-span steel truss suspension bridge utilizing SCCDs is considered with a length of 1176 m, a height of 7.5 m, a width of 27 m, and a truss node spacing of 7.25 m (Wang et al., 2018). The bridge deck includes left and right traffic lanes, arranged with spaced separation as shown in Fig. 4. The longitudinal spacing is unevenly arranged according to shear resistance requirements. Temperature monitoring sensors are installed at the mid-span, 1/4 span, and near the tower section of the bridge, with Sections A to E depicted in Fig. 4a. In this study, Section C was selected for detailed analysis. The diameter and height of the shear connectors are 22 and 190 mm, respectively. In the cross-sectional view in Fig. 4b, three shear connectors are welded onto the top surface of each longitudinal steel beam, with a transverse spacing of 120 mm.

A computational model was built using the general-purpose FE software ABAQUS. The radiation in shaded areas was implemented using the DFLUX interface in ABAQUS. This calculation is only necessary for the concrete slab and the steel girders on both sides. The process is divided into the following steps. First, based on the sun's elevation angle, it is determined whether it is day or night, as shadow effects on solar radiation are only calculated during the day. During daylight, the solar incident angle on each surface of the composite beam is used to combine the shadowed and sunlit areas. The top surface of the bridge deck is always exposed to solar radiation, while the sides are affected by their own shadow as the sun's position changes. The steel girder web has partially sunlit and shadowed areas, which are determined through calculations using Eq. (S7) of the ESM (Zhang et al., 2020; Fan et al., 2021). The bottom surface of the bridge deck and the bottom flanges of the steel girders do not receive direct solar radiation and are therefore not included in the shadow recognition calculation. Finally, the process ends once all the calculations are completed. A two-dimensional FE transient heat transfer analysis was conducted, ignoring temperature differences along the longitudinal direction of the girder segments. The dimensions of the computational model as well as the mesh division are shown in Fig. 4b. The composite beam consists of three parts: asphalt concrete, concrete, and steel beam, all simulated using four-node quadrilateral heat transfer elements (DC2D4) (Wang et al., 2021). The different materials are bonded to each other by tie constraints to ensure temperature and heat flow density transfer

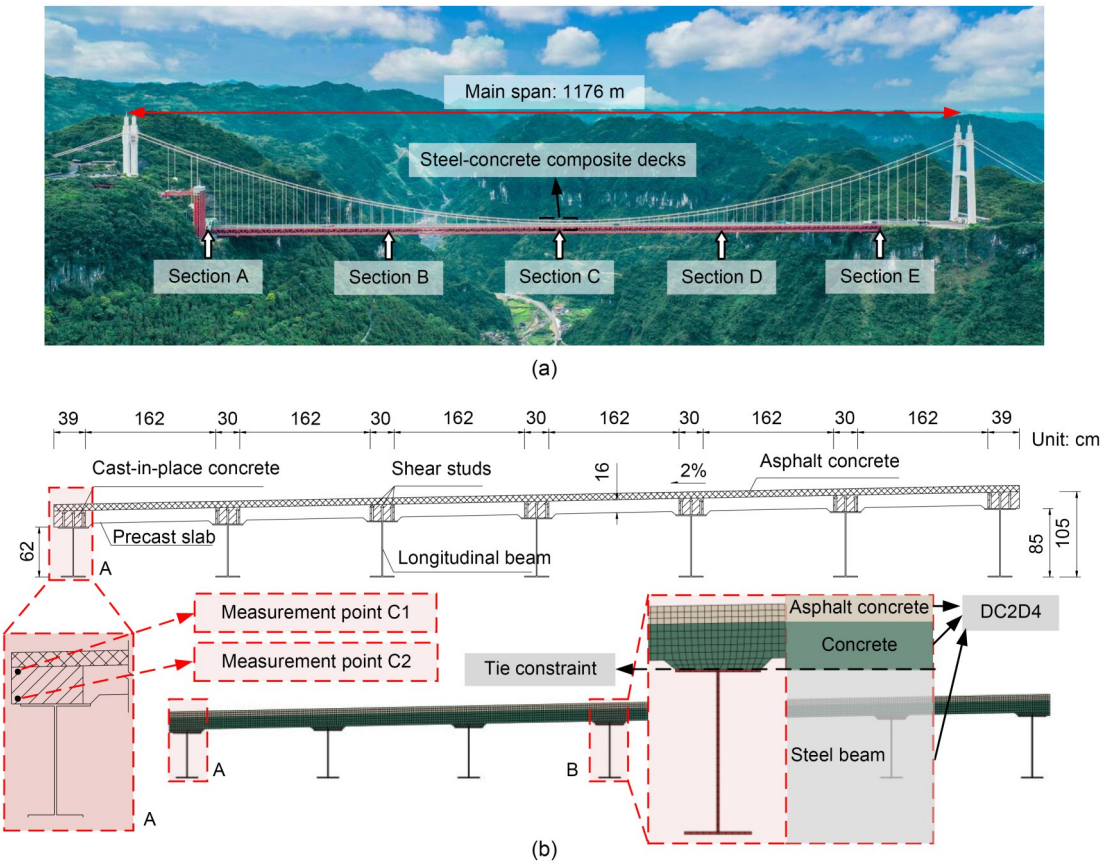

Fig. 4 Engineering background: (a) overall diagram of suspension bridge; (b) cross-sectional dimensions and FE model of the SCCDs

between the different material interfaces. In this study, the FE calculation verification was performed using the C1 and C2 measuring points in Fig. 4b as examples.

In this study, the temperature field of the bridge site was simulated for the entire year of 2013, and temperature data was sampled on an hourly basis. Hourly temperature and radiation data in the vicinity of the bridge site are given in Fig. 5 which could be determined by means of the atmospheric temperature data of local weather stations. The boundary conditions and calculation parameters required for the calculations are provided in Table S1 of the ESM. The longitude and latitude of the bridge site location are E109.610° and N28.333°, respectively, and the elevation is 571.1 m. The ground reflectance (r_e) used to calculate the reflected radiation was taken to be 0.2. Each surface has a different angle of incidence with respect to solar radiation and therefore receives a different type of radiation, including direct radiation (R_d), diffuse radiation (R_s), and reflected radiation (R_r). The parentheses in Table S1 indicate that the type of radiation received by the sides of the steel web and the top surface of the lower flange needs to be determined. This determination is based on combining the calculation results of the shaded lengths in Eqs. (S8) and (S9).

3.2 Validation of the temperature field calculation method

According to the calculation method described in Fig. 1, it is possible to simulate the temperature field of the SCCDs under sunlight radiation. To ensure the accuracy of the FE simulation results, actual measurement data from points C1 and C2 in Fig. 4b were selected, and the calculated hourly temperature results were verified. Fig. S2 of the ESM shows a comparison between the calculated values from the FE model and the actual measured values during some periods in winter and summer. It can be seen that the data obtained through the simulation are consistent with the actual measurement data, with a certain deviation at some moments but within the tolerance range.

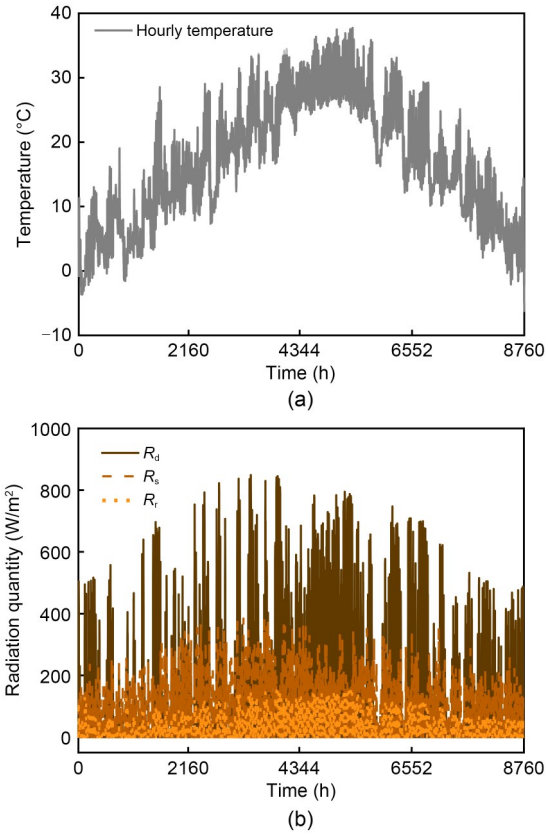

Fig. 5 Hourly temperature and radiation parameters: (a) temperature; (b) radiation

The above calculation method can better simulate the time-varying temperature field of the structure under sunlight. It proves that the temperature analysis model established above can be used to develop the following analyses.

4 Results assessment and discussion

4.1 Feature preprocessing analysis

The annual temperature data samples obtained from the numerical temperature field simulations were used to build a prediction model for the temperature field of the SCCDs based on the TVFEMD-stacking ensemble algorithm. As mentioned earlier, three additional stages of pre-processing of the time series were required before the model could be built. Firstly, the lagging terms of the data features were analyzed using XGBoost's feature-importance scoring. Then, each feature was decomposed using two effective data decomposition techniques, namely EMD and TVFEMD, to provide suitable data for ML to be built. Finally, the decomposed feature was restructured and dimensionality was reduced using the sample entropy method.

As the FE model contains many nodes, only five representative nodes are chosen to analyze the middle I-beam. These selected nodes are positioned at different locations, including the asphalt concrete layer, the center of the concrete slab, the steel-concrete joint, and the middle and bottom of the steel beam, respectively. For ease of reference, they are named M1 to M5 from top to bottom, as shown in Fig. 6. It should be noted that the temperature data of only these five measurement points are analyzed in this study to better illustrate the specific analysis steps. However, the research method used in this study is applicable to any location within the entire cross-section.

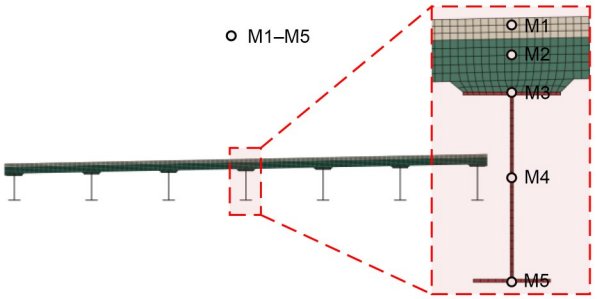

Fig. 6 Study subjects M1 to M5

The XGBoost method was used to analyze the influence of the atmospheric temperature lag term on each measurement point. Here, t represents the actual temperature sequence, and $t-n$ (n=1, 2, ⋯, 6) indicates a lag of n hours on the original basis. The results of each time lag characteristic score for measurement points M1 to M5 were calculated and are shown in Table 1. From the data in the table, it can be observed that the measurement points are slightly affected by the temperature lag due to their different spatial locations. M3, which is located at the junction of the section, is affected by the temperature at moment $t-1$,

Table 1 Feature importance scores based on the XGBoost model

Node	Feature importance score						
	t	$t-1$	$t-2$	$t-3$	$t-4$	$t-5$	$t-6$
M1	0.59	0.28	0.03	0.01	0.01	0.01	0.01
M2	0.00	0.00	0.04	0.74	0.22	0.00	0.01
M3	0.03	0.93	0.03	0.00	0.00	0.00	0.01
M4	0.67	0.33	0.00	0.00	0.00	0.00	0.00
M5	0.67	0.33	0.00	0.00	0.00	0.00	0.00

while M2, located inside the concrete, is affected by the temperature value at moment $t-3$. The other measurement points are closer to the contact surface with the air and are therefore directly affected by the ambient temperature at moment t. To take into account the significance of features across various locations, ambient temperature input variables including t, $t-1$, and $t-3$ were selected in addition to the input values including radiation data and time series. The raw data samples were divided in the ratio of 7:3 and set as the training and validation sets, respectively. Different features were transformed into the same scale through standardization to avoid the differences in values between different features.

The decomposition of features can better handle the mixed patterns of non-stationary time series and extract the more important variables. However, an increase in the number of components obtained after decomposition may increase the dimensionality of the features, resulting in an increase in the time required for the TVFEMD model prediction and in the computational cost. Therefore, the obtained IMF 1 through n are combined to obtain new components (Sub 1, Sub 2, ⋯, Sub m). As an example, the ambient temperature of the training set was decomposed using EMD and TVFEMD, and the results of recombination after decomposition are shown in Fig. 7.

The choice of hyperparameters has a significant impact on the prediction efficiency of the model. To set more optimal hyperparameters, commonly used methods include random search, Bayesian optimization, and random grid search. In this section, stochastic grid search is used to optimize the model hyperparameters in conjunction with the cross-validation shown in Fig. 2. The optimized hyperparameter settings are presented in Table S2 of the ESM. In the table, "original", "time lag-based", "EMD-based", and "TVFEMD-based" models represent models without time lag, models with a time lag, models with a time lag and EMD decomposition, and models with a time lag and TVFEMD decomposition, respectively.

4.2 Single model forecasting analysis

After processing the input features, the temperature values of the combined SCCDs can be predicted. As an example, using the XGBoost model (without hyperparameter tuning), the prediction results for the M2 and M3 measurement points with lag terms of 1 and 2, respectively, are shown in Table 2. The table presents results both with and without the time lag term being taken into account. Based on the data in the table, it can be observed that the errors in M2 and M3 are significantly reduced by the addition of the lag term. Therefore, it is necessary to consider the effect of the time lag term on the temperature field of the SCCDs at different locations.

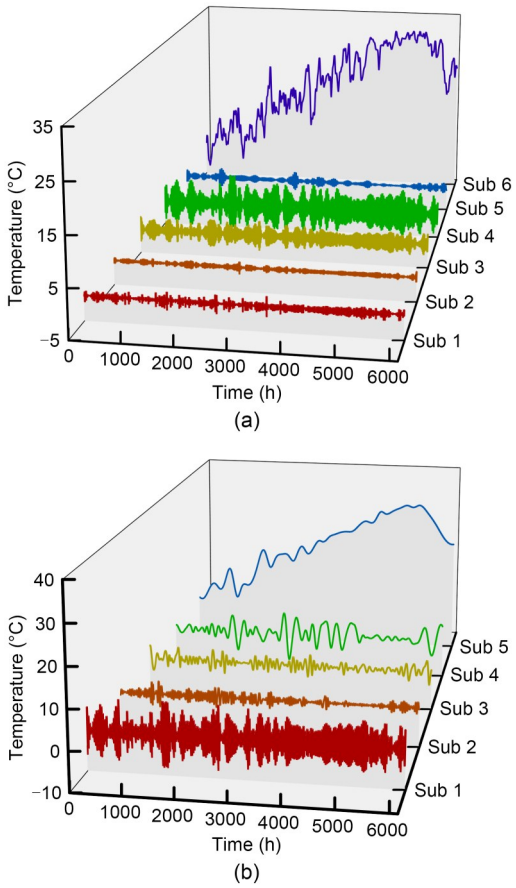

Fig. 7 Results of recombination after decomposition of ambient temperature sequences: (a) EMD; (b) TVFEMD

Table 2 Assessment of the impact of model time lags

Node	Original			Time lag-based		
	R^2	MAE	RMSE	R^2	MAE	RMSE
M1	0.90	2.24	2.94	0.95	1.54	2.05
M2	0.84	2.52	3.08	0.94	1.63	2.09
M3	0.96	1.18	1.16	0.99	0.56	0.71
M4	0.99	0.38	0.44	0.99	0.38	0.44
M5	0.99	0.50	0.41	0.99	0.50	0.41

Furthermore, it is important to note that the M2 measurement point has a relatively large time lag, and its temperature changes are more complex. Therefore,

its prediction accuracy has more room for improvement compared to the other measurement points. To improve the prediction accuracy of the M2 measurement point, the steady-state time series decomposition was used, taking into account the effect of temperature time lags. The prediction results for different single models are shown in Table S3 of the ESM, including the R^2, MAE, and RMSE results for the original model, the time lag-based model, the EMD-based model, and the TVFEMD-based model.

The results in Table S3 indicate that different models exhibit varying performance under different data processing methods. The TVFEMD decomposition-based data processing method performs better for most of the models, with the TVFEMD-XGBoost model and the TVFEMD-GBR model demonstrating higher prediction accuracy, achieving R^2 values of 0.97 for both models. Among these, the former has the lowest MAE (R^2=0.97, MAE=0.96, RMSE=1.24), while the latter has the lowest RMSE (R^2=0.97, MAE=0.98, RMSE=1.23). The performance of the SVR method was poor in the original and time lag-based data processing methods. However, by considering the time lag combined with the TVFEMD decomposition, the performance of each model was significantly improved. Even the SVR model, which had the worst prediction accuracy, could be improved from an R^2 of 0.86 to one of 0.95 through the TVFEMD decomposition. Moreover, the MAE and RMSE could be reduced from 2.27 and 2.86 to 1.45 and 1.80, respectively. These results demonstrate that the proposed method considering time lag combined with TVFEMD decomposition is an effective way of improving the temperature prediction effect using a temperature prediction model.

4.3 Stacking ensemble model prediction

The fundamental principle of ensemble learning is to combine multiple models or algorithms to improve overall predictive performance. To enhance the diversity and prediction accuracy in stacking ensemble learning, the RF, SVR, MLP, GBR, and XGBoost models were used as base learners, combined with the elastic net model as a meta-learner. The stacking integration prediction model was built using the framework shown in Fig. 3. The R^2, MAE, and RMSE results of the stacking ensemble models are presented in Table S4 of the ESM. The results show that the stacking, time lag-stacking, EMD-stacking, and TVFEMD-stacking models have R^2 values of 0.91, 0.96, 0.97, and 0.98, respectively. It is worth noting that the TVFEMD-stacking model outperformed the single model, with the highest prediction accuracy (R^2=0.98, MAE=0.79, RMSE=1.01).

Fig. 8 displays the scatter correlation plot between the predicted and target values, combining the

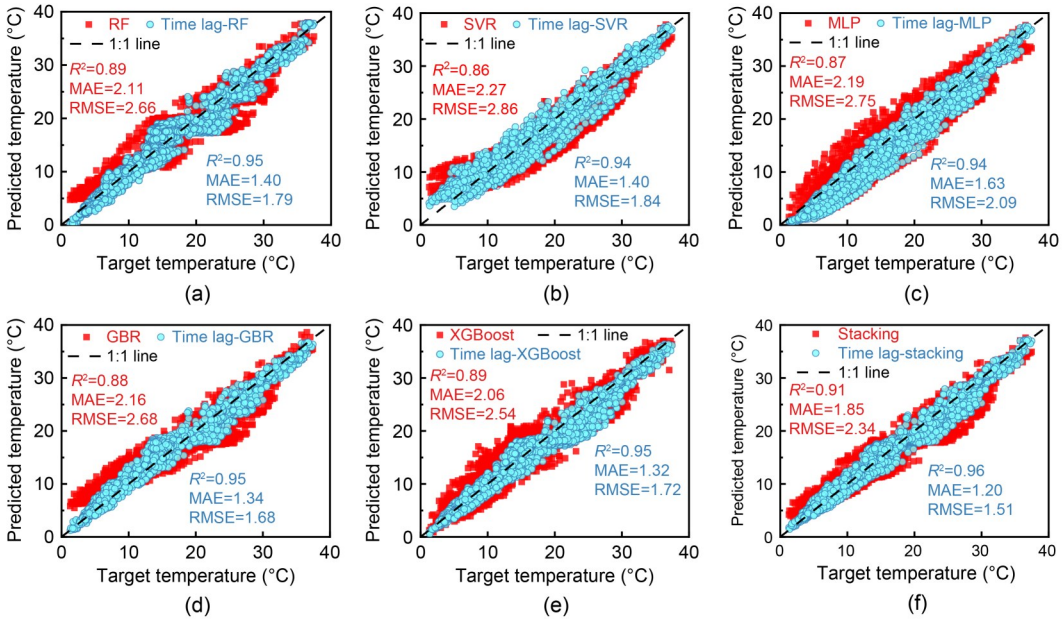

Fig. 8 Predicted temperature and target temperature scatter correlation plots (original and time lag-based models): (a) RF; (b) SVR; (c) MLP; (d) GBR; (e) XGBoost; (f) stacking

prediction results of the original model and the time lag-based model. It can be observed that the scatter distribution of the prediction results of the original model with respect to the target value is more dispersed, while the prediction results of the time lag-based model are closer to the scatter distribution of the target value and have a better correlation. This suggests that considering the time lag can improve the prediction accuracy of the model. Fig. 9 combines the prediction results of the EMD-based models and the TVFEMD-based models to give a scatter correlation plot between the predicted and target values. Both decomposition-based models have a more concentrated scatter distribution with a better correlation between the prediction results and the target values. This indicates that the decomposition-based models can capture the characteristics of temperature variation more accurately and thus improve the prediction accuracy.

The Taylor diagram was used to comprehensively compare the predictive performance of the models (Taylor, 2001). It is a graph that allows for the assessment of the similarities and differences between model outputs and observations, incorporating both the R^2 and RMSE metrics and adding the standard deviation (SD) between the output and target values for a more comprehensive evaluation. Based on Fig. 10a, it can be observed that the results of both the original and time lag-based models lie outside the arc of RMSE=2, indicating relatively large prediction errors. In contrast, the results of most of the models lie between the arcs of 2 and 4, suggesting relatively small prediction errors. Although the time lag-stacking model shows high accuracy for R^2 and RMSE, it is far from the target value in terms of standard deviation, indicating its predictions are more dispersed. On the other hand, according to Fig. 10b, the EMD-based and TVFEMD-based models show mostly clustered predictions on the arc of RMSE=2, indicating relatively small prediction errors. The results of the TVFEMD-stacking model are within the arc, with better standard deviation predictions, implying better performance in predicting temperature characteristics. Therefore, decomposition-based models exhibited better performance in predicting the temperature of the SCCDs.

To further compare the prediction errors of EMD and TVFEMD with the target values, the relative error distribution characteristics between the two models' predicted values and target values were compared. The upper quartile (Q75%), median, lower quartile (Q25%), mean, and interquartile range (IQR) of the relative error between the predicted values and target values of the EMD-based and TVFEMD-based models were calculated and are presented in Fig. 11 and Table S5. It was found that the stacking models based on EMD

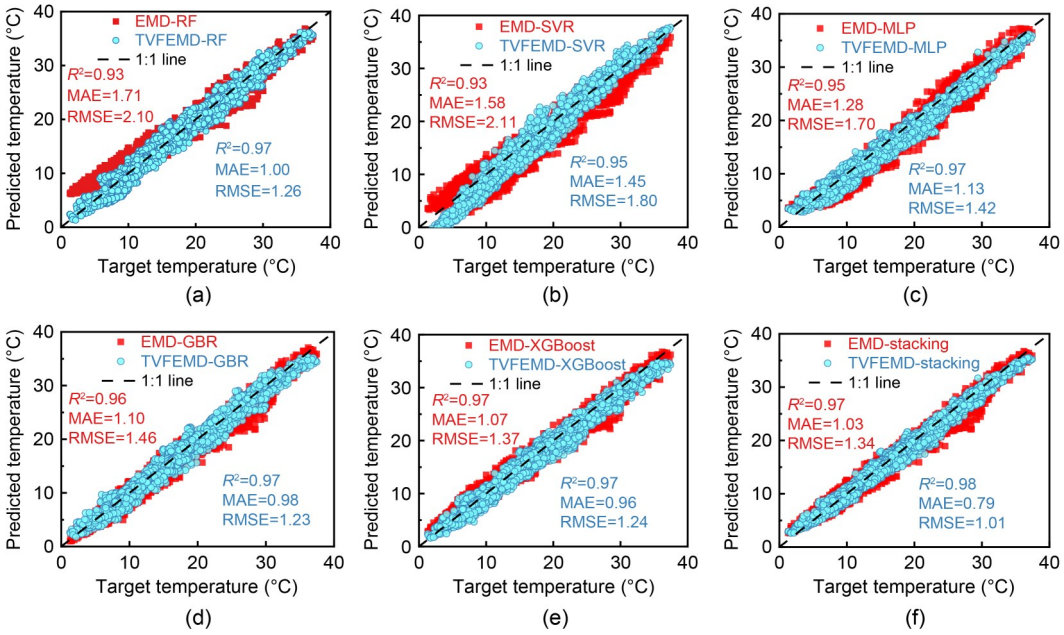

Fig. 9 Predicted temperature and target temperature scatter correlation plots (EMD-based and TVFEMD-based models): (a) RF; (b) SVR; (c) MLP; (d) GBR; (e) XGBoost; (f) stacking

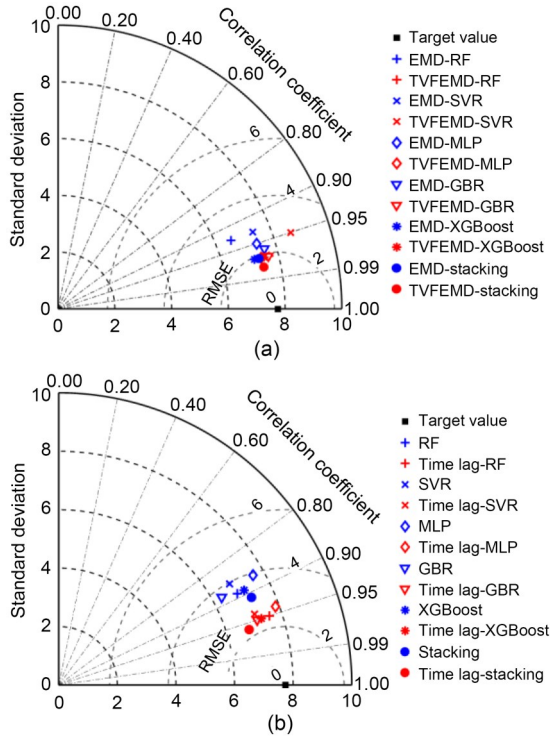

Fig. 10 Temperature prediction and target value Taylor diagram: (a) EMD-based and TVFEMD-based models; (b) original and time lag-based models. References to color refer to the online version of this figure

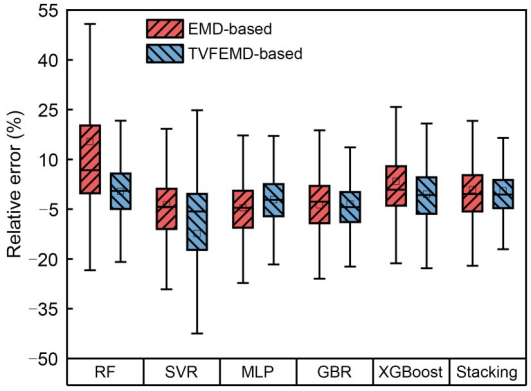

Fig. 11 Relative error box plot

and TVFEMD had better performance compared to other models. Specifically, the median relative error of the EMD-stacking model was −0.46, which was closer to 0 than the result of the TVFEMD-stacking model (−0.58). However, the TVFEMD-stacking model had a lower mean relative error, and a smaller IQR, indicating a more concentrated error distribution, and a closer proximity to 0. Therefore, overall, the TVFEMD-stacking model demonstrated better performance in temperature prediction accuracy.

4.4 Statistical analysis of temperature samples

To carry out studies such as reliability analysis of bridge structures, statistical analysis of their temperature characteristics is usually required. Therefore, the prediction results of the ML model should be able to describe the statistical characteristics of the actual temperature. Fig. 12 provides an empirical probability density plot of the prediction results for the EMD-based models and the TVFEMD-based models. It can be observed that the distribution characteristics of the prediction results differ significantly from the target values before the input characteristics of the model are processed. However, by considering the amount of time lag and performing the TVFEMD decomposition, the distribution of the empirical probability density function (PDF) is very close to the target value. Furthermore, Fig. 13 presents the empirical cumulative probability density distribution of the TVFEMD-based models, which indicates that the TVFEMD-stacking model has a better prediction and the smallest error from the target value.

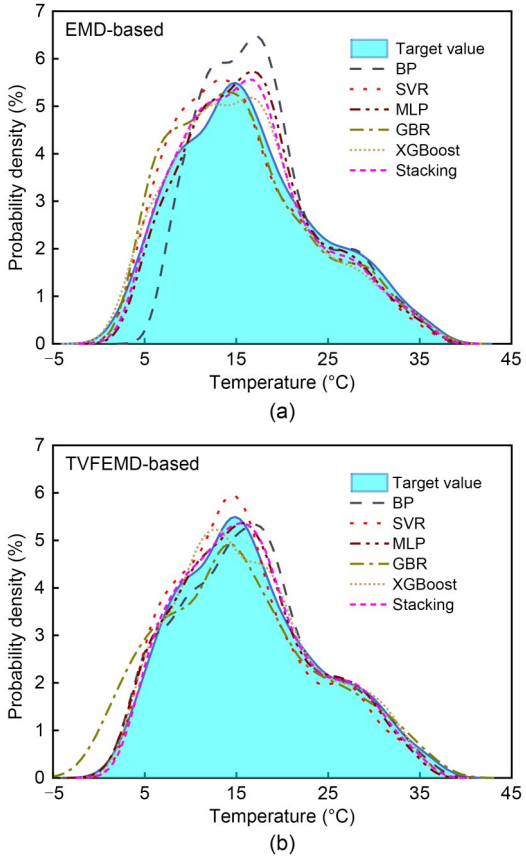

Fig. 12 Comparison of empirical probability density plots: (a) EMD-based models; (b) TVFEMD-based models

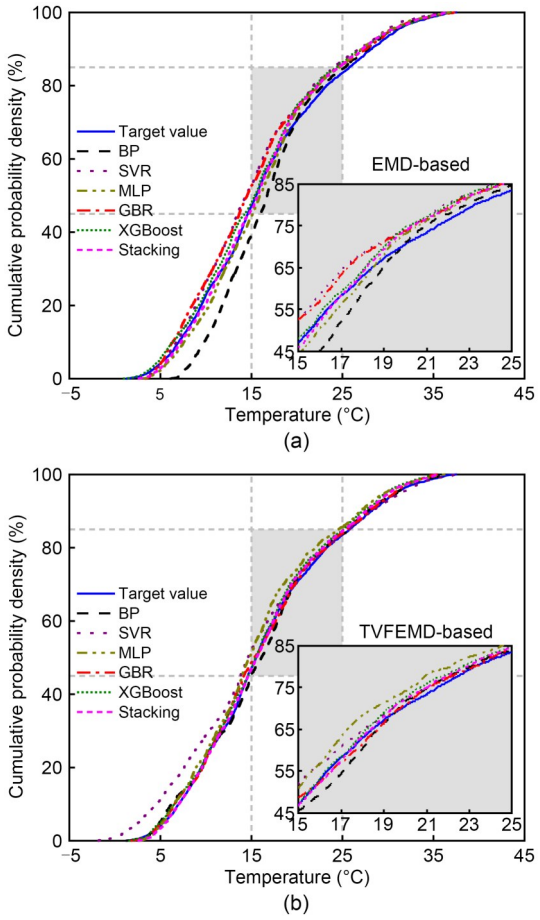

Fig. 13 Comparison of cumulative probability density plots of experience: (a) EMD-based models; (b) TVFEMD-based models

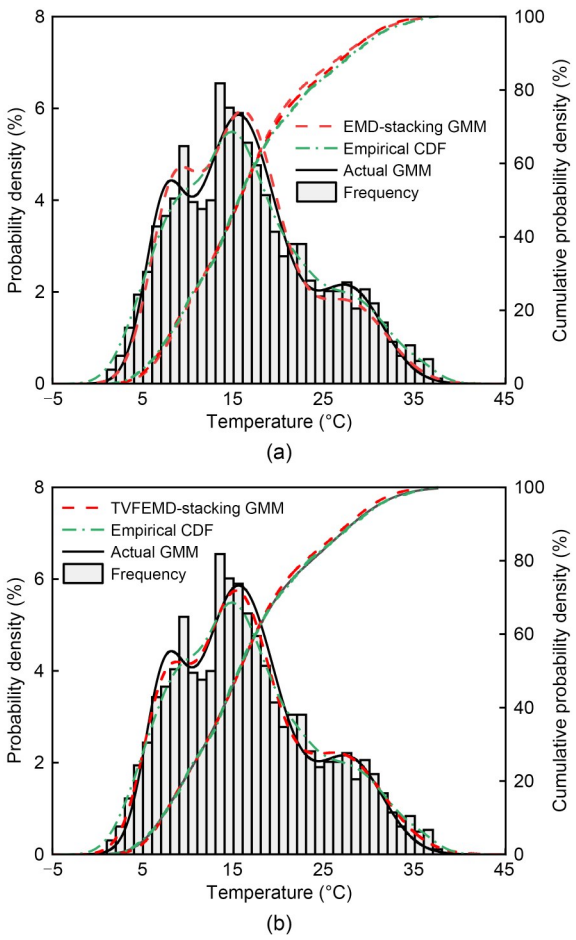

Fig. 14 Prediction and target value distribution fitting results: (a) EMD-based models; (b) TVFEMD-based models

The temperature distribution characteristics of the actual values and the predictions of the EMD-stacking and TVFEMD-stacking models can be modeled using a Gaussian mixed model (GMM) represented by Eq. (11), and the fitting results are presented in Fig. 14.

$$f(x)=\sum_{i=1}^{n} w_i \frac{1}{\sqrt{2\pi}\,\sigma_i}\exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right), \quad (11)$$

where n is the number of Gaussian components, w_i is the weight of the ith component, and μ_i and σ_i are the mean and standard deviation of the ith component, respectively. The optimal number of Gaussian components n for the fit is obtained by calculating the lowest value of the Akaika information criterion (AIC) and the Bayesian information criterion (BIC).

The goodness of fit was assessed using the Kolmogorov-Smirnov (KS) test by comparing the empirical cumulative distribution function (CDF) of two samples to determine whether they come from the same overall distribution (Wang et al., 2020). The null hypothesis of the KS test is that the samples come from the same overall distribution, while the alternative hypothesis is that they come from different overall distributions. The fitted probabilistic model parameters and the results of the goodness-of-fit tests are presented in Table S6, which also includes the RMSE values of the CDF and PDF between the fitted probabilistic model and the sample values. The RMSE values of the PDF and CDF for the actual values are 0.0024 and 0.0043, respectively; for the EMD-stacking model, they are 0.0048 and 0.0150, respectively; for the TVFEMD-stacking model, they are 0.0038 and 0.0086, respectively.

After comparison, the optimal number of components is 3 for all cases. Based on the analysis of the fitted Gaussian mixed distribution probability models

and the results of the goodness-of-fit tests, it can be concluded that both the EMD-stacking model and the TVFEMD-stacking model provide good predictions of the temperature distribution, as their fitting functions overlap with the actual values and the KS tests are accepted. However, the TVFEMD-stacking model performs slightly better in terms of describing the probability distribution of the actual temperature, as indicated by the lower RMSE of PDF and CDF compared to the EMD-stacking model. Therefore, it can be concluded that the TVFEMD-stacking model is the better choice for predicting the temperature distribution of the SCCDs.

5 Conclusions

The TVFEMD-stacking method proposed in this paper combines actual measurements and FE analysis with ML techniques, demonstrating good predictive performance by considering time lag and non-stationary time series features to improve prediction accuracy. Based on experimental verification and FE analysis, a TVFEMD-based stacking ensemble prediction model was developed by optimizing hyperparameters for BP, SVR, MLP, GBR, and XGBoost models. The main conclusions are as follows:

1. The FE model was validated by on-site monitoring, which accurately calculated the temperature field of SCCDs under solar radiation. A single ML model prediction method can be developed based on FE data. The prediction accuracy was relatively low, with the SVR model performing the worst (R^2=0.86, MAE=2.27, RMSE=2.86).

2. The predictive performance of a total of 20 models, which include four scenarios, i.e., considering no time delay, considering time delay, considering time delay and EMD, and considering time delay and TVFEMD, each combined with five individual ML models, was evaluated. The results show that considering time delay and input features from EMD can effectively improve the predictive accuracy of individual models.

3. The models that combined time lag and feature decomposition had higher prediction accuracy. Among them, the TVFEMD-GBR model and TVFEMD-XGBoost model had relatively higher prediction accuracy, with evaluation indicators of (R^2=0.97, MAE=0.98, RMSE=1.23) and (R^2=0.97, MAE=0.96, RMSE=1.24), respectively.

4. The stacking ensemble models based on EMD and TVFEMD can both further improve the prediction performance of the single ML model. The predicted distribution of the TVFEMD-stacking model is the closest to the fitted distribution of the actual values, with RMSE differences in PDF and CDF of 0.0014 and 0.0043, respectively.

5. The model proposed here is highly accurate and can be used for rapid prediction of the temperature distribution of SCCDs under environmental influences. The content of this study can provide a reference for further analysis of temperature gradients over long service lifecycles, as well as real-time analysis of thermal deformation behavior and thermally induced strains.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 52278235), the Science and Technology Program of Hunan Provincial Department of Transportation (No. 202309), China.

Author contributions

Benkun TAN designed the research. Da WANG processed the corresponding data. Benkun TAN and Lianqi ZHANG wrote the first draft of the manuscript. Jialin SHI helped to organize the manuscript. Benkun TAN and Da WANG revised and edited the final version.

Conflict of interest

Benkun TAN, Da WANG, Jialin SHI, and Lianqi ZHANG declare that they have no conflict of interest.

References

Boudraa AO, Cexus JC, 2007. EMD-based signal filtering. *IEEE Transactions on Instrumentation and Measurement*, 56(6):2196-2202. https://doi.org/10.1109/TIM.2007.907967

Branco FA, Mendes PA, 1993. Thermal actions for concrete bridge design. *Journal of Structural Engineering*, 119(8): 2313-2331. https://doi.org/10.1061/(ASCE)0733-9445(1993)119:8(2313)

Broo DG, Bravo-Haro M, Schooling J, 2022. Design and implementation of a smart infrastructure digital twin. *Automation in Construction*, 136:104171. https://doi.org/10.1016/j.autcon.2022.104171

Catbas FN, Susoy M, Frangopol DM, 2008. Structural health monitoring and reliability estimation: long span truss bridge application with environmental monitoring data. *Engineering Structures*, 30(9):2347-2359.

https://doi.org/10.1016/j.engstruct.2008.01.013

Chen FH, Zhang HP, Li ZC, et al., 2024. Residual stresses effects on fatigue crack growth behavior of rib-to-deck double-sided welded joints in orthotropic steel decks. *Advances in Structural Engineering*, 27(1):35-50. https://doi.org/10.1177/13694332231213462

Fan JS, Liu YF, Liu C, 2021. Experiment study and refined modeling of temperature field of steel-concrete composite beam bridges. *Engineering Structures*, 240:112350. https://doi.org/10.1016/j.engstruct.2021.112350

Fan JS, Li BL, Liu C, et al., 2022. An efficient model for simulation of temperature field of steel-concrete composite beam bridges. *Structures*, 43:1868-1880. https://doi.org/10.1016/j.istruc.2022.05.079

Figueiredo E, Santos LO, Moldovan I, et al., 2023. A roadmap for an integrated assessment approach to the adaptation of concrete bridges to climate change. *Journal of Bridge Engineering*, 28(6):03123002. https://doi.org/10.1061/JBENF2.BEENG-5735

Flah M, Nunez I, Chaabene WB, et al., 2021. Machine learning algorithms in civil structural health monitoring: a systematic review. *Archives of Computational Methods in Engineering*, 28(4):2621-2643. https://doi.org/10.1007/s11831-020-09471-9

Friedman JH, 2001. Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29(5):1189-1232. https://doi.org/10.1214/aos/1013203451

Fu WW, Sun BC, Wan HP, et al., 2022. A Gaussian processes-based approach for damage detection of concrete structure using temperature-induced strain. *Engineering Structures*, 268:114740. https://doi.org/10.1016/j.engstruct.2022.114740

Giussani F, 2009. The effects of temperature variations on the long-term behaviour of composite steel–concrete beams. *Engineering Structures*, 31(10):2392-2406. https://doi.org/10.1016/j.engstruct.2009.05.014

Han QH, Ma Q, Xu J, et al., 2021. Structural health monitoring research under varying temperature condition: a review. *Journal of Civil Structural Health Monitoring*, 11(1):149-173. https://doi.org/10.1007/s13349-020-00444-x

Innocenzi RD, Nicoletti V, Arezzo D, et al., 2022. A good practice for the proof testing of cable-stayed bridges. *Applied Sciences*, 12(7):3547. https://doi.org/10.3390/app12073547

Jamei M, Karbasi M, Ali M, et al., 2023. A novel global solar exposure forecasting model based on air temperature: designing a new multi-processing ensemble deep learning paradigm. *Expert Systems with Applications*, 222:119811. https://doi.org/10.1016/j.eswa.2023.119811

Lee JH, 2012. Investigation of extreme environmental conditions and design thermal gradients during construction for prestressed concrete bridge girders. *Journal of Bridge Engineering*, 17(3):547-556. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000277

Liu HJ, Chen C, Guo ZQ, et al., 2021. Overall grouting compactness detection of bridge prestressed bellows based on RF feature selection and the GA-SVM model. *Construction and Building Materials*, 301:124323. https://doi.org/10.1016/j.conbuildmat.2021.124323

Liu J, Liu YJ, Zhang CY, et al., 2020. Temperature action and effect of concrete-filled steel tubular bridges: a review. *Journal of Traffic and Transportation Engineering*, 7(2):174-191. https://doi.org/10.1016/j.jtte.2020.03.001

Luo Y, Liu XF, Chen FH, et al., 2023. Numerical simulation on crack–inclusion interaction for rib-to-deck welded joints in orthotropic steel deck. *Metals*, 13(8):1402. https://doi.org/10.3390/met13081402

Narasimhan TN, 1999. Fourier's heat conduction equation: history, influence, and connections. *Reviews of Geophysics*, 37(1):151-172. https://doi.org/10.1029/1998RG900006

Nguyen H, Vu T, Vo TP, et al., 2021. Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials*, 266:120950. https://doi.org/10.1016/j.conbuildmat.2020.120950

Nicoletti V, Quarchioni S, Tentella L, et al., 2023. Experimental tests and numerical analyses for the dynamic characterization of a steel and wooden cable-stayed footbridge. *Infrastructures*, 8(6):100. https://doi.org/10.3390/infrastructures8060100

Opoku DGJ, Perera S, Osei-Kyei R, et al., 2021. Digital twin application in the construction industry: a literature review. *Journal of Building Engineering*, 40:102726. https://doi.org/10.1016/j.jobe.2021.102726

Qin YH, Hiller JE, 2011. Modeling temperature distribution in rigid pavement slabs: impact of air temperature. *Construction and Building Materials*, 25(9):3753-3761. https://doi.org/10.1016/j.conbuildmat.2011.04.015

Richman JS, Moorman JR, 2000. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6):H2039-H2049. https://doi.org/10.1152/ajpheart.2000.278.6.H2039

Sheng XW, Zhou TM, Huang SJ, et al., 2022. Prediction of vertical temperature gradient on concrete box-girder considering different locations in China. *Case Studies in Construction Materials*, 16:e01026. https://doi.org/10.1016/j.cscm.2022.e01026

Shi T, Lou P, Zheng WQ, et al., 2022. A hybrid approach to predict vertical temperature gradient of ballastless track caused by solar radiation. *Construction and Building Materials*, 352:129063. https://doi.org/10.1016/j.conbuildmat.2022.129063

Shim CS, Lee PG, Chang SP, 2001. Design of shear connection in composite steel and concrete bridges with precast decks. *Journal of Constructional Steel Research*, 57(3):203-219. https://doi.org/10.1016/S0143-974X(00)00018-3

Sohn H, Dzwonczyk M, Straser EG, et al., 1999. An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge. *Earthquake Engineering & Structural Dynamics*, 28(8):879-897. https://doi.org/10.1002/(SICI)1096-9845(199908)28:8<

879::AID-EQE845>3.0.CO;2-V

Sugumaran V, Muralidharan V, Ramachandran KI, 2007. Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing. *Mechanical Systems and Signal Processing*, 21(2):930-942. https://doi.org/10.1016/j.ymssp.2006.05.004

Taylor KE, 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7):7183-7192. https://doi.org/10.1029/2000JD900719

Tong M, Tham LG, Au FTK, 2002. Extreme thermal loading on steel bridges in tropical region. *Journal of Bridge Engineering*, 7(6):357-366. https://doi.org/10.1061/(ASCE)1084-0702(2002)7:6(357)

Wang D, Liu YM, Liu Y, 2018. 3D temperature gradient effect on a steel-concrete composite deck in a suspension bridge with field monitoring data. *Structural Control and Health Monitoring*, 25(7):e2179. https://doi.org/10.1002/stc.2179

Wang D, Tan BK, Wang X, et al., 2021. Experimental study and numerical simulation of temperature gradient effect for steel-concrete composite bridge deck. *Measurement and Control*, 54(5-6):681-691. https://doi.org/10.1177/00202940211007166

Wang J, Du XY, Qi X, 2022. Strain prediction for historical timber buildings with a hybrid Prophet-XGBoost model. *Mechanical Systems and Signal Processing*, 179:109316. https://doi.org/10.1016/j.ymssp.2022.109316

Wang ZW, Zhang WM, Tian GM, et al., 2020. Joint values determination of wind and temperature actions on long-span bridges: copula-based analysis using long-term meteorological data. *Engineering Structures*, 219:110866. https://doi.org/10.1016/j.engstruct.2020.110866

Wedel F, Marx S, 2022. Application of machine learning methods on real bridge monitoring data. *Engineering Structures*, 250:113365. https://doi.org/10.1016/j.engstruct.2021.113365

Xin JZ, Zhou CY, Jiang Y, et al., 2023. A signal recovery method for bridge monitoring system using TVFEMD and encoder-decoder aided LSTM. *Measurement*, 214:112797. https://doi.org/10.1016/j.measurement.2023.112797

Zhang CY, Liu YJ, Liu J, et al., 2020. Validation of long-term temperature simulations in a steel-concrete composite girder. *Structures*, 27:1962-1976. https://doi.org/10.1016/j.istruc.2020.07.070

Zhang PJ, Wang CS, Wu GS, et al., 2022. Temperature gradient models of steel-concrete composite girder based on long-term monitoring data. *Journal of Constructional Steel Research*, 194:107309. https://doi.org/10.1016/j.jcsr.2022.107309

Zhang ZJ, Liu YJ, Liu J, et al., 2023. Experimental study and analysis for the long-term behavior of the steel–concrete composite girder bridge. *Structures*, 51:1305-1327. https://doi.org/10.1016/j.istruc.2023.03.052

Zhao HW, Ding YL, Li AQ, et al., 2023. Digital modeling approach of distributional mapping from structural temperature field to temperature-induced strain field for bridges. *Journal of Civil Structural Health Monitoring*, 13(1):251-267. https://doi.org/10.1007/s13349-022-00635-8

Zou H, Hastie T, 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Electronic supplementary materials

Sections S1–S4, Eqs. (S1)–(S36), Tables S1–S6, Figs. S1 and S2