RESEARCH

Machine Learning Approach to Rapidly Evaluate Curling of Concrete Pavement

Sangyoung Han[1], Taemin Heo[2], Chul Min Yeum[3], Kukjoo Kim[4*], Jongkwon Choi[5] and Mang Tia[6]

Abstract

This paper focuses on the methodology for evaluating the degree of total curling in concrete pavement using machine learning. Deflection induced by falling weight deflectometer (FWD) testing is known as a direct correlation to total curling including built-in and daily curling. However, deflection measurement in the in-service road is also affected by others, such as environmental conditions, pavement geometry, subgrade stiffness, and mixture design. Thus, it is challenging to determine the level of curling from FWD data due to the complexity of influencing parameters. To navigate this complexity, prominent machine learning models are exploited to identify a non-linear relationship between curling and FWD deflections. A finite-element simulation of FWD is conducted to generate a vast data set, and a robust regression model is trained to estimate the total effective temperature difference (TETD) to quantify the effects of curling. Since input parameters for testing pavements can be measurable in the field, curling from TETD can be readily obtained using the proposed methodology. Comparative simulations highlight that the proposed models, with an MAE less than 0.5 °C significantly outperform the multiple regression performance, which registers an MAE of 3.45 °C in TETD, particularly in offering cost-effective and noise-tolerant prediction.

Keywords Concrete pavement, Machine learning, Finite element, TETD, FWD

*Correspondence:
Kukjoo Kim
kukjoo.kim@mnd.go.kr
[1] Department of ICT Integrated Ocean Smart City Engineering, Dong-A University, Busan 49315, South Korea
[2] The Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712, USA
[3] Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[4] Defense Installations Agency, Ministry of Defense, 22 Itaewon-Ro, Hangangno-Dong, Yongsan-Gu, Seoul, South Korea
[5] Department of Civil and Environmental Engineering, Hongik University, Seoul 04066, South Korea
[6] Engineering School of Sustainable Infrastructure and Environment, University of Florida, Gainesville, FL 32611, USA

1 Introduction

The Florida Department of Transportation has operated the Florida Concrete Test Road, a 4-km long test road (Greene, 2016). One of the major challenges in collecting data from this test road is how to determine the magnitude of total curling, which is defined as the degree of combined built-in and reversible daily curvature of a concrete slab. It is well-known that a high degree of concrete pavement curling deformation can result in undesirable issues that affect rideability (Johnson et al., 2010), cracking (Beckemeyer et al., 2002), and durability (Lange & Shin, 2001). Therefore, it is imperative to recognize the curling response of concrete pavement in the design process. Moreover, in the proper evaluation of concrete pavement response to in-service conditions, information on curling deformation should be considered. If the effect of curling is not considered in existing analysis tools, such as finite-element model (FEM) and the mechanistic-empirical pavement design

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 2 of 15

guide (MEPDG [ARA, Inc., 2007]), the concrete pavement stress state will be misjudged.

Built-in curling is known as a permanent form of curling deformation in hardened concrete pavement (Franklin, 1969; Poblete et al., 1990). The early age curing behavior of hardening concrete is closely associated with irreversibility mainly due to the temperature gradients at the concrete setting time, known as the built-in temperature gradient. In the long term, this irreversibility is further exacerbated by nonlinear factors such as excessive drying shrinkage in the internal pore region as well as creep close to the surface (Jeong & Zollinger, 2003; Wei & Hansen, 2011; Yeon et al., 2013). On the other hand, the temperature and moisture gradient through the concrete slab depth by diurnal events can lead to cyclic differential deformation between the top and bottom of the concrete slab, referred to as daily curling. The coexistence of built-in and daily curling on concrete pavement makes the total curling measurement difficult without long-term deformation data from the initial hardening phase.

The process of measuring the built-in curling of concrete pavement involves decomposing the irreversible built-in curling from the total curling. Numerous researchers have measured and determined the built-in curling of early age concrete slabs in controlled experiments (Alland et al., 2017; Asbahan & Vandenbossche, 2011; Hansen et al., 2006; Hiller & Roesler, 2010; Rao & Roesler, 2005). However, such approaches are not applicable to determine the degree of built-in curling of in-service concrete pavement at a hardened stage, where daily curling is also present. In regard to implementing existing tools or methods, most pavement surface profilers require time-consuming and costly instrumentation plans for long-term monitoring and data collection (Asbahan & Vandenbossche, 2011; Ceylan et al., 2016; Hansen et al., 2006; Rao & Roesler, 2005; Sridhar et al., 2022). Moreover, due to the high level of dependence on historic records, sparse data collection could result in an unreliable prediction of the performance of concrete pavement (Asbahan & Vandenbossche, 2011; Brody et al., 2023). This led the significant need for a robust and cost-effective methodology to determine the in situ curling of concrete pavement.

This study focuses on the development of a methodology for quantifying the total curling of concrete pavement using well-known machine learning algorithms. The deflection behavior of concrete pavement by downward loading increases with the loss of subgrade support by the deformation of slab curling (Huang, 2003). Therefore, it is hypothesized that the surface deflection induced by loading can be correlated to the magnitude of total curling. Falling Weight of Deflectometer (FWD) testing is a type of nondestructive testing method to capture surface deflection characteristics under various loading conditions. However, total curling is not the only factor that determines the amount of concrete pavement deflection, and many other geometric and environmental parameters are also affected. The core idea of the proposed methodology is to build a machine learning model to identify the complex relationship between total curling and the relevant influence parameters. Here, these parameters can be easily gathered from construction records (e.g., concrete mix design, pavement geometry) as well as measured in the field (e.g., FWD testing); thus, the proposed method permits the rapid evaluation of total curling without expensive instrumentation and tremendous monitoring effort.

To pursue this goal, the FE model simulates FWD testing to obtain the data set needed for training and testing the machine learning model for total curling estimation. Note that the FE model is designed specifically for one of the major types of concrete pavement, i.e., jointed plain concrete pavement (JPCP). Consequently, the findings and results are primarily applicable to JPCP. However, the presented methodology can be extended to other types of pavement, such as continuously reinforced concrete pavement (CRCP) or precast prestressed concrete pavement (PPCP), where curling exists. The FE simulation of FWD testing is conducted on various field-representative concrete models and testing environments. The model parameters are selected by considering the concrete mix design, structural geometric profile, boundary condition, and total curling. After aggregating the database, machine learning algorithms are used to develop the candidate models, and the performance of the total curling estimation for these models is thoroughly tested.

While the effects and measurements of curling have been extensively studied, there is still a gap in effective and efficient methodologies to determine the degree of total curling (Asbahan & Vandenbossche, 2011; Ceylan et al., 2016; Hansen et al., 2006; Rao & Roesler, 2005; Sridhar et al., 2022). The novel methodology presented in this paper significantly contributes to the fact that the proposed method plays a pivotal role in efficiently and rapidly calculating the degree of total curling from the vast secondary data of FWD and providing the design input of total curling needed for existing design and analytic tools. To give you an idea, the developed regression model in this study will promptly output the intensity of total curling from the data collected by users. The trained regression model will be shared with the public, so that they can easily surrogate complex and computationally demanding FE model. By inputting

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 3 of 15

total curling information directly to existing design tools, researchers and/or engineers can efficiently analyze the structural performance and obtain accurate curling behavior information.

2 FE Model for FWD Simulation

2.1 Concept of TETD for Curling in FE

The total curling behavior of concrete pavement is influenced by several nonlinear factors. Reversible curling behavior arises due to temperature ($\triangle T_{tg}$) as well as the moisture gradient through the concrete slab ($\triangle T_{mg}$). On the other hand, irreversible curling aspects are attributed to the built-in temperature gradient at the concrete setting time ($\triangle T_{bi}$), differential drying shrinkage gradient through the concrete slab ($\triangle T_{shr}$), and creep ($\triangle T_{crp}$). The accumulated degree of curling for a hardened concrete slab primarily results from the aforementioned nonlinear factors. This accumulation can be represented in terms of the temperature degree, referred to as the total effective temperature difference (TETD) $\triangle T_{tot}$, as calculated below:

$$\Delta T_{tot} = \Delta T_{tg} + \Delta T_{mg} + \Delta T_{bi} + \Delta T_{shr} + \Delta T_{crp} \tag{1}$$

Given the fact that modeling the curling geometry is exceedingly difficult in a FE model (Sridhar et al., 2022) and MEPDG (Vandenbossche et al., 2011) directly as a function of curvature degree, the concept of the total effective temperature difference (TETD) as a function of the temperature difference is adopted from Rao and Roesler (2005). The TETD allows us to simulate the same effect of built-in and daily curling in FE and MEPDG simulations using the temperature gradient through the depth of the concrete slab model. In brief, the TETD deforms the slab into a concave shape in a way that is similar to the deformation by total curling. Fig. 1 exhibits an example of how to apply daily and built-in curling to a FE model. Fig. 1a addresses a general method to simulate daily curling behavior by making the temperature difference at the time of FWD testing through the slab depth during the daytime (e.g., a temperature difference of +5.0
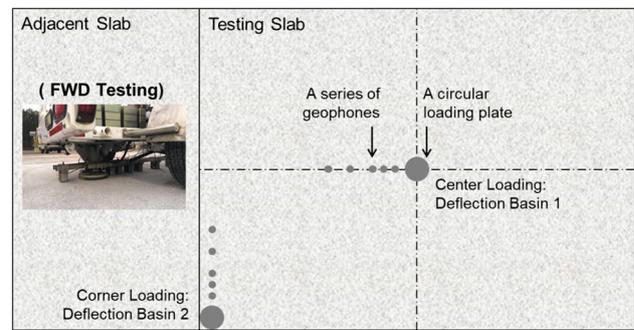

Fig. 2 FWD testing configurations: Two impact locations are selected, one at the corner and another in the center of the slab.

°C, reflecting a top slab surface temperature of 25.0 °C and an interlayer temperature of 20.0 °C). Fig. 1b provides the effective temperature difference (ETD), which represents the degree of built-in curling, i.e., the effect of deforming a concrete slab by applying the top surface temperature to obtain a concave shape (e.g., − 10.0 °C). The superimposed temperature, defined as the TETD (e.g., − 5.0 °C), reflecting the effects of daily and built-in curling, is used for this study, as shown in Fig. 1c. In this study, the TETD is the target output of the proposed regression model, allowing us to quantify the amount of slab curling.

2.2 FE Design for Achieving the FWD Data Set

FWD testing is defined as "a device designed to simulate deflection of a pavement surface caused by a fast-moving truck" according to FHWA (2006). This nondestructive testing is devised to measure the surface response behavior by dropping a circular plate from an array of installed geophones. To make the FE simulation of FWD testing, six geophones placement named as deflection basin is modeled according to the FHWA guidelines (2006). Moreover, a major pavement system of a jointed plain concrete pavement (JPCP) design is simulated. In this study, the configurations of FWD testing are design
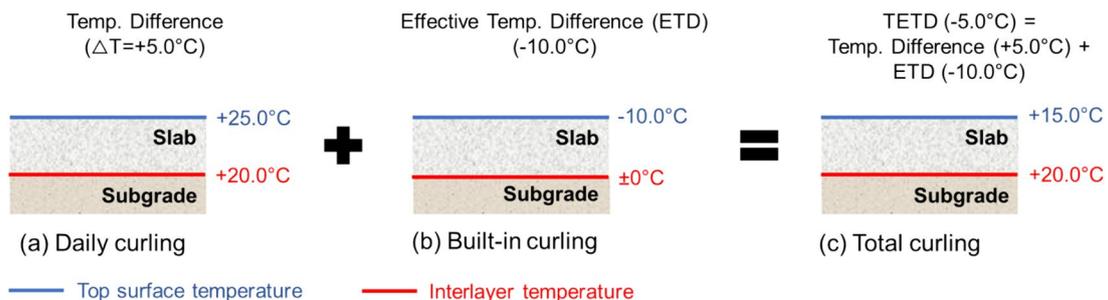

Fig. 1 Daily and built-in curling modeling through the total effective temperature difference (TETD).

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 4 of 15

to obtain the optimal deflection response from FWD testing.

Some modifications to the configurations of FWD testing are made for the purpose of predicting the TETD of JPCP. Two configurations of the FWD testing are located at the corner of the slab expected to experience the maximum deflection and another is located in the center of the slab, which experiences minimum deflection in the daytime. Fig. 2 illustrates the loading locations and their corresponding geophone locations (two directions of deflection basins). The two locations of FWD loadings and collected defection basins capture sufficient information about the characteristics of deflection responses as the input data for the regression model.

Fig. 3a illustrates the 3D FE model used for generating FWD responses from the input parameters, which include the concrete mix design, geometry profile, boundary conditions, and TETD. The inputted values considering the combined properties are used to design

the FE model. A full-size pavement system is designed, and the subgrade extends to each side to eliminate the effect of boundary conditions. Because of the JPCP design, a half-size adjacent slab—a symmetry condition that does not affect the FE results—is modeled, and each adjacent slab impacts the corner deflection due to the effect of load transfer on the joint. The FWD plate loading is treated as a static loading condition because the primary concern is to determine the maximum deflection corresponding to a specific magnitude of loading to satisfy the purpose of this study. This approach contingently saves tremendous computational time. To further enhance the FE model, the incorporation of dynamic loading conditions is recommended for future studies to account for the effect of the strain rate. In the FE model, rather than replicating the actual geometric conditions of dowel bars, the effect of load transfer is simulated by considering the behavior of dowel bars at the joint location of JPCP. To compute the actual magnitude of load transfer
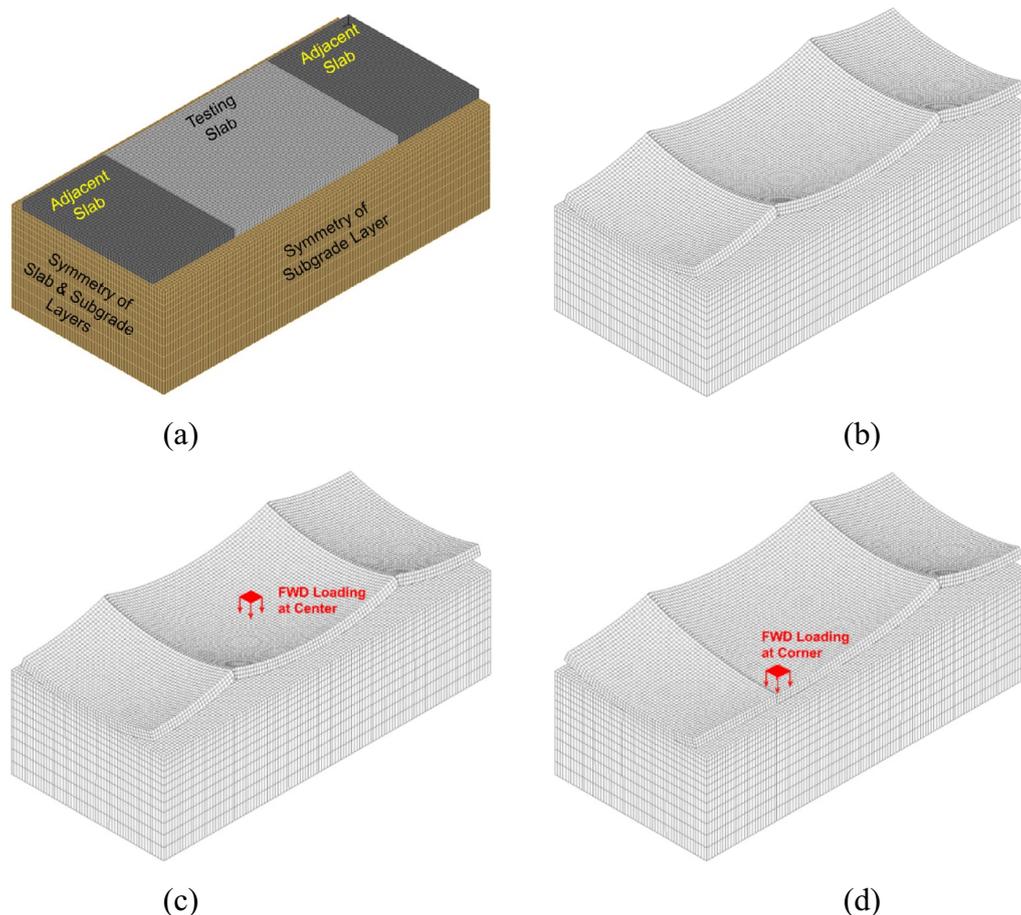

(a)

(b)

(c)

(d)

Fig. 3 3D FE model for generating FWD response data: **a** developed FE model, **b** results of the FE model after the TETD is applied (concave-up shape) for providing the initial deformation data, **c** results of the FE model after TETD and FWD loading at the center of a concrete slab, and **d** results of the FE model after TETD and FWD loading at the corner of the concrete slab.

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 5 of 15

in this FE model, the effective load transfer is calibrated using the FWD field test data on JPCP along with actual dowel bars at the joint location.

To address FWD protocols for generating a machine learning data set, two FWD deflection basins, such as those in the center and corner joints, should be computed. To calculate the deflection by FWD loading, the designed FE model should account for the initial deformation information that considers the TETD without FWD loading, as shown in Fig. 3b. Specifically, the concrete slab is already deformed by the effect of the TETD before FWD loading. Thus, the initial deformed positions need to be determined from the FE model with the TETD, referred to as the initial deformation. Note that this structural analysis can consider the predeformation in the concrete slab before applying FWD loading. From the FE results of initial deformation, the final deflection, such as center and corner joint deflection, can be computed by subtracting the center FWD loading or the corner FWD loading, as shown in Fig. 3c, d, respectively. Therefore, 1179 samples incorporating simulated center and corner joint deflection are generated in this study as the supervised training data set for machine learning regression, which requires a model trained on a data set that is triple the size, i.e., 3537 samples.

Table 1 presents the modeling parameters for the FE model categorized by the geometry condition, concrete mix design, boundary condition, and TETD. The ranges of the parametric values are obtained from a normal concrete pavement design based on AASHTO PP84 (2017). The number of feasible combinations with these factors is enormous; thus, 1179 randomly generated combinations are used for the input data of the FE model. Three different FE models for initial deformation, center, and corner joint loadings are analyzed using 1179 input values to produce two FWD responses, one at the center and the other at a corner joint. The FWD response results are the independent variables used to predict the TETD in this study. It is rational to anticipate that the effect of built-in curling reduces the subgrade support at the corner of a concrete slab due to its deformation (i.e., concave-up shape). As a result, the surface deflection measured by the FWD increases as the degree of built-in curling increases. Therefore, quantifying the total curling can be achieved by evaluating the concrete slab deflection response due to the corresponding FWD loading and its location. In this study, extensive FWD data reinforce establishing a meaningful pattern to achieve the relationship between the various individual parameters. Note that the data set maintains a consistent number of input parameters in this study. For future studies, it is strongly advised to segment data sets based on varying counts of input parameters.

2.3 FE Calibration from the Field Data

To improve the reliability of the FE results, the FE model was calibrated using collected field data (Tia et al., 2020). The properties of the concrete material were determined from the results of laboratory tests on the sampled concrete. Surface deflections in the concrete pavement caused by FWD loading were used to estimate the values of the elastic modulus of the subgrade in the FE model. Based on the backcalculation methodology guided by FHWA (2006), the curling behavior of a concrete slab influenced by temperature and moisture gradients is theoretically negated when FWD testing is conducted at the different times corresponding to specific locations on JPCP. For instance, the center FWD loading is applied at night or in early morning—times when the concrete slab curls upward—to backcalculate the effective modulus of the subgrade (k value). In contrast, corner FWD loading is carried out during the daytime, when the concrete slab curls downward, facilitating the determination of the

Table 1 Lower and upper bounds of the finite-element analysis inputs.

Input values		Lower bound	Upper bound
Geometry condition	Wide (m [ft])	3.7 (12)	3.7 (12)
	Depth (cm [in.])	15.2 (6)	30.4 (12)
	Length (m [ft])	4.3 (14)	4.9 (16)
Concrete mix design	MOE (MPa [ksi])	24,131 (3500)	37,921 (5500)
	Unit Weight (kg/m³ [pcf])	1922 (120)	2563 (160)
	CTE (/°C [/°F])	10.8 (6.0)	18.0 (10.0)
	Poisson Ratio	0.23	0.25
Boundary condition	Subgrade Modulus (MPa [ksi])	344 (50)	2757 (400)
	Load Transfer (kg/m [lb/in.])	5760 (500,000)	34,563 (3,000,000)
TETD	TETD (°C [°F])	− 33 (− 60)	+11 (+20)

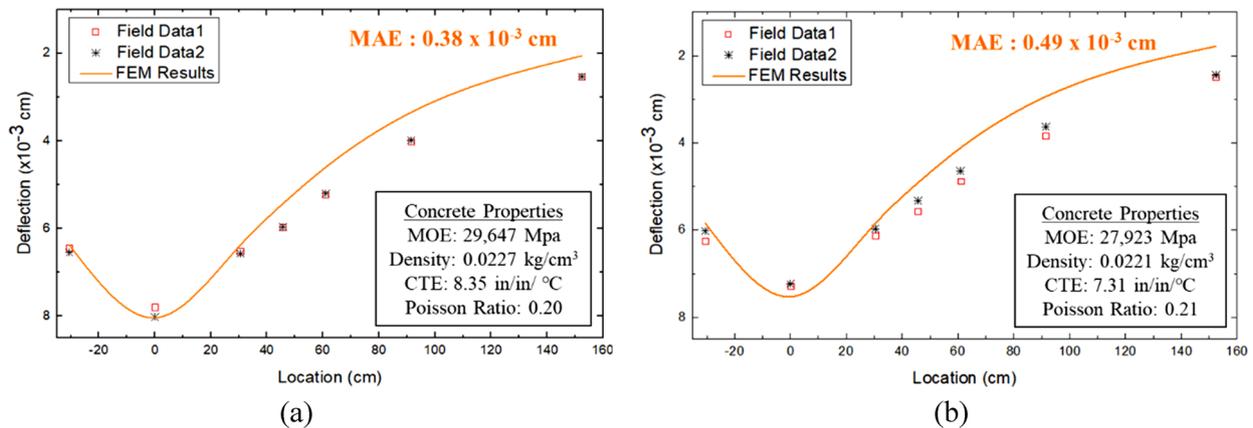Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 6 of 15

Fig. 4 Determination of the effective subgrade modulus using an FWD basin caused by a 5.4-ton (12-kip) FWD load at the center for the (**a**) testing slab and (**b**) adjacent slab.

load transfer. In these scenarios, the FWD loading location on the concrete slab achieves full contact conditions with the subgrade, effectively eliminating the influence of temperature and moisture gradients, irrespective of their intensity. Fig. 4 shows the measured and computed deflection basin caused by a 5.4-ton (12-kip) FWD load. The analytical deflection basin was calculated using an elastic modulus of the subgrade of 482 MPa for the testing slab and 551 MPa for the adjacent slab, respectively. Using the previously estimated parameters and material properties, fairly well-matched results between the measured and the calculated deflection basins based on the calculated mean absolute error (MAE) were obtained.

3 Development of Regression Models

This section addresses the development of regression models and the evaluation of their performance to predict the TETD that is used in existing designs and analytic tools and for estimating the degree of curling. In this study, prominent machine learning algorithms are used to create the best-performing regression model. Three regression models are built and tested for the performance of TETD estimation. Noise sensitivity analysis is also conducted to validate the model developed under a field setting involving FWD testing and input parameter measurements.

Fig. 5 illustrates the steps for training the regression model. First, the data set is prepared by conducting a simulation of FWD testing using random combinations of properties. The range of each concrete mix property is in accordance with AASHTO PP84 (2017), as summarized in Table 2. Commercial finite-element software, ADINA, is used to iteratively simulate FWD testing under different parametric setups. A total of

1179 samples are created from the FE simulation. The TensorFlow, Keras, and Sklearn libraries in Python are implemented to develop the regression models using the simulation data set. Here, surface deflection is the output of the simulation and field-measurable through FWD testing. Therefore, the regression model uses the surface deflection as an input as well, but the TETD becomes the output of the model. Note that due to the lack of extensive field FWD data, FE simulations are utilized to create a vast amount of defection basin data, which genuinely serve as the dependent variable in this process. However, the FE-driven deflection basin data are then used to train the machine learning models, positioning these data as an independent variable for predicting the TETD output. This scheme is commonly termed '*surrogate modeling*'. Here, a low-fidelity machine learning model is trained to act as a stand-in for a more computationally intensive high-fidelity numerical simulator. This surrogate modeling approach might prompt concerns regarding the reliability of predictions, including potential overfitting issues, from the trained models. To address these concerns and validate the models, k-fold cross-validation is conducted, and the model is tested with noise sensitivity analysis. By randomly subdividing the data into k segments and using k-1 folds to train and the remaining fold to test, we can ensure the independence between the training and testing data. Furthermore, we generate adversarial samples, defined as noise-added deflection basins, which add 5%, 10%, and 15% Gaussian random noise into the FE-derived deflection basin data. The robustness of trained models using noise-added deflection basins to random noise shows that the model is well trained on the general data not included in the training
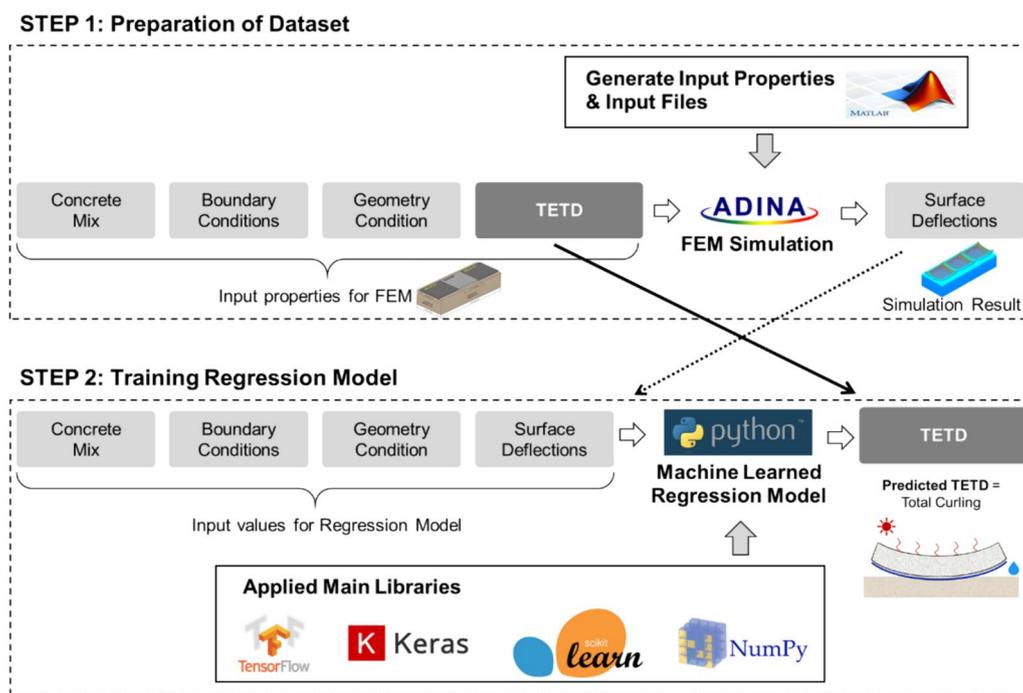
Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 7 of 15

Fig. 5 Regression model development environment and the input(s) and output(s) for each step.

and validation data. The detailed process and subsequent results are addressed in the following sections.

3.1 Preparation of the Data Sets

After analyzing the FE model, a series of surface responses regarding deflection basins 1 and 2 can be achieved. Since the objective of this study is to estimate the TETD as one of the input properties of the FE model, a method of supervised learning—one of the big data learning categories among supervised, unsupervised, and reinforcement learnings—is selected. Supervised learning makes use of both the independent variable(s) and the dependent variable of the TETD as one of the needed input data for model training. The data set combining input(s) and output data from the FE simulation is then used to train and test the regression models based on the technique of supervised learning.

To improve the performance of the regression model, the normalization of the data samples is recommended to calculate the numerical matrix by removing the magnitude difference between independent values (Chou & Pham, 2013). Min–max normalization is a well-known and effective method for avoiding the magnitude difference between inputs that convert all values to the range between 0 and 1. Fig. 6 clearly illustrates input data that are uniformly distributed and free of outliers, although the target data for the TETD appear biased due to the randomly combined matrix of input parameters. In this figure, the data distribution is presented in its nonnormalized form, confirming the absence of outliers. This suggests that the application of min–max normalization is suitable, as this technique performs well in scenarios without outliers.

Table 2 Results of the performance indicators for SVM, DT, and DNN.

Performance indicator		SVM (Quadratic)	DT (LGBM)	DNN (Deep Learning)
MAE (Degree, °C)	Training	0.35	0.02	0.32
	Test	0.71	0.99	0.78
adjusted_R^2	Training	0.998	0.999	0.998
	Test	0.991	0.984	0.987
MSE	Training	0.0001540	0.0000003	0.0000860
	Test	0.0005102	0.0009628	0.0007355

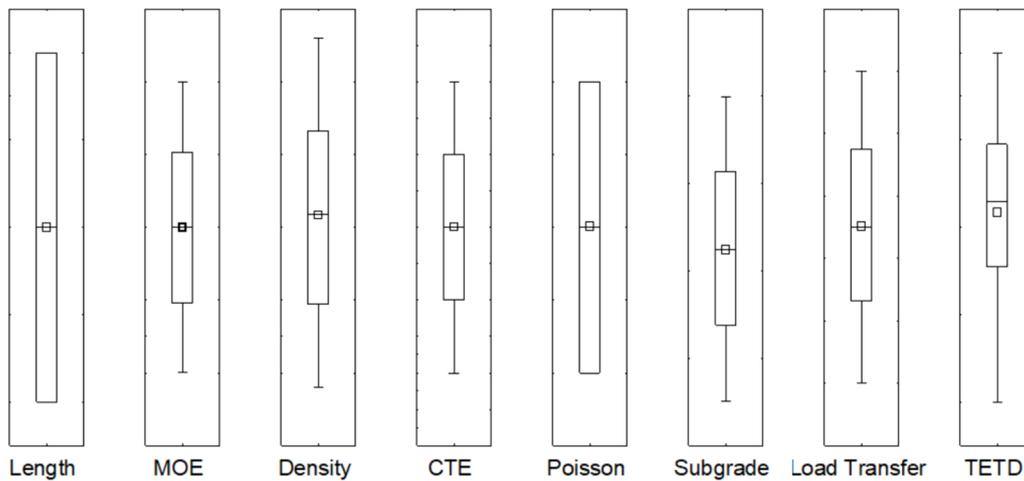Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 8 of 15

Fig. 6 Non-normalized data set distribution for training the machine learning regression model.
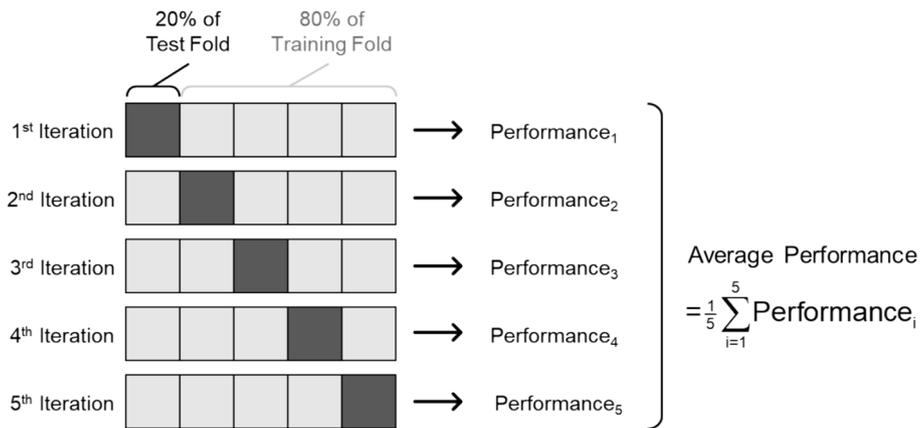

Fig. 7 Designed k-fold cross-validation (i = 5 iterations) to prepare the data set in this study.

3.2 Training the Regression Model (Supervised Learning)

Three machine learning algorithms are utilized: the support vector machine (SVM)-based quadratic model, decision tree (DT)-based light gradient boosting machine, and deep neural network (DNN) model. The SVM builds the feature of a kernel to create a hyperplane and then computes the optimized margin between the different objectives. The DT involves splitting the variables at discrete cutting points from the specified criteria. The DNN is a perceptron-based method that updates the weights until the minimum of the cost function is found. These algorithms are compared to each other to select the best-performing algorithm.

Although all machine learning models can achieve high accuracy, they can overfit the data, limiting successful prediction results to those within the training data. To avoid this issue, first, the data set is divided as follows: 80% for the training data set and 20% for the test data set. The divided samples for the training and test data sets are then applied to the technique of k-fold cross-validation, as illustrated in Fig. 7. This technique is constantly applied to all regression models to avoid overfitting issues. Moreover, the grid search technique using the Sklearn application determines the best combination of hyperparameters for each model to optimize each regression model. Note that every model has different types of fitting parameters to achieve various levels of predictive performance.

3.2.1 Multiple Regression

Multiple regression predicts the TETD with the same data set to provide benchmarks for assessing the performance of the machine learning models. Among the multiple regression methods, principal component regression (PCR) is used for a large number

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 9 of 15

of independent variables. This method generates new predictor variables referred to as the component, and tuning the number of components improves the accuracy of the regression model. By applying 21 components, the regression model achieves the best results in terms of predicting the TETD. This method builds new predictor variables as linear combinations from the original predictor variables. Despite this widely used solution for regression, the predictive performance could be limited in regard to complex nonlinear cases. In this paper, PCR serves as the benchmark for assessing the predictive performance of the machine learning models. Future studies should consider comparisons with other conventional nonlinear regression approaches.

3.2.2 Support Vector Machine (SVM)

The support vector machine (SVM) is a popular machine learning algorithm that is used to solve regression and classification cases and was introduced by Cortes and Vapnik (1995). The algorithm minimizes the loss that is measured based on the distance between the observed value and the boundary. The SVM can be divided into four different methods in terms of the type of kernel function, such as the linear kernel, radial basis function kernel, sigmoid kernel, and quadratic (polynomial) kernel. In this study, due to the characteristics of the nonlinear relationship between dependent and independent variables, the quadratic SVM aids in predicting the values from the prepared data set. This study makes use of the grid search technique, thereby building a robust regression model based on the optimized combination of the fitting parameters. This technique determines the hyperparameters in terms of degree (degree of the polynomial kernel function), C (regularization parameter), epsilon (epsilon in the epsilon–SVR model), and gamma (kernel coefficient) for the SVM-based regression model. Regarding the tuning parameters for this model, more detailed information can be found in the Scikit-learn documentation (Pedregosa et al., 2011).

3.2.3 Decision Tree (DT)

The light gradient boosting machine (LightGBM) is a decision tree (DT)-based model composed of a weighted combination of multiple regression trees introduced by Microsoft Research (Ke et al., 2017). By combining multiple regression trees, the accuracy is generally improved. In other words, the LightGBM model aggregates all results of each single regression tree and then uses the average of the results for regression. Depending on the method used to select the random subset of the original data, such as boosted tree or bagged tree, the ensemble algorithms generate different

predictions. In this study, the boosted tree algorithm predicts the dependent value. Recalling the grid search used for each model, the hyperparameters in terms of the learning rate (boosting learning rate), number of leaves (maximum tree leaves for base learners), and n_estimators (number of boosted trees to fit) are determined.

3.2.4 Deep Neural Network (DNN)

Deep learning uses neural networks to learn useful representations of features and to find a specific pattern directly from inputted data. This deep neural network algorithm combines multiple nonlinear hidden processing layers. In the parametric study to tune the number of hidden layers by manual calibration, the use of nine hidden layers generates the best performance for predicting the dependent value. Moreover, the aforementioned grid search technique employed in this developed model determines the hyperparameter tuning regarding optimizers (method of updating the deep neural network based on the loss function), epoch (number of data sets passed forward and backward through the deep neural network; 500 in this study), batch (size of a data sample; 5 in this study), and learning rate (deep neural network learning rate; 0.001 in this study).

4 Evaluation of the Regression Models

4.1 Assessing the Goodness-of-Fit of the Regression Models

Three different regression models, including SVM, DT, and DNN, have been tested on the database from the FE simulation in the previous section, and their performance is compared with that of the control model by a simple statistical approach. The accuracy of each selected regression model is evaluated by comparing the values of the true and the predicted TETD from the same data set. For the quantitative evaluation, the mean absolute error (*MAE*) is primarily used as the performance indicator, which intuitively reports the actual difference of the TETD in the degree unit. The mean absolute error (*MAE*) determines the level of error between the true and predicted values. The *MAE* explains the absolute difference between the predicted and actual values. The results of the *MAE* scores are from zero to any positive value. The *MAE* score indicates how much error the model produces. In addition, both the adjusted *R*-squared (*adjusted_R^2*) and mean square error (*MSE*) are provided as supplemental information. Recall that the developed models uniformly apply k-fold cross-validation, and then the performance indicators, including the *MAE, adjusted_R^2*, and *MSE,* are computed from the average results of K validations.

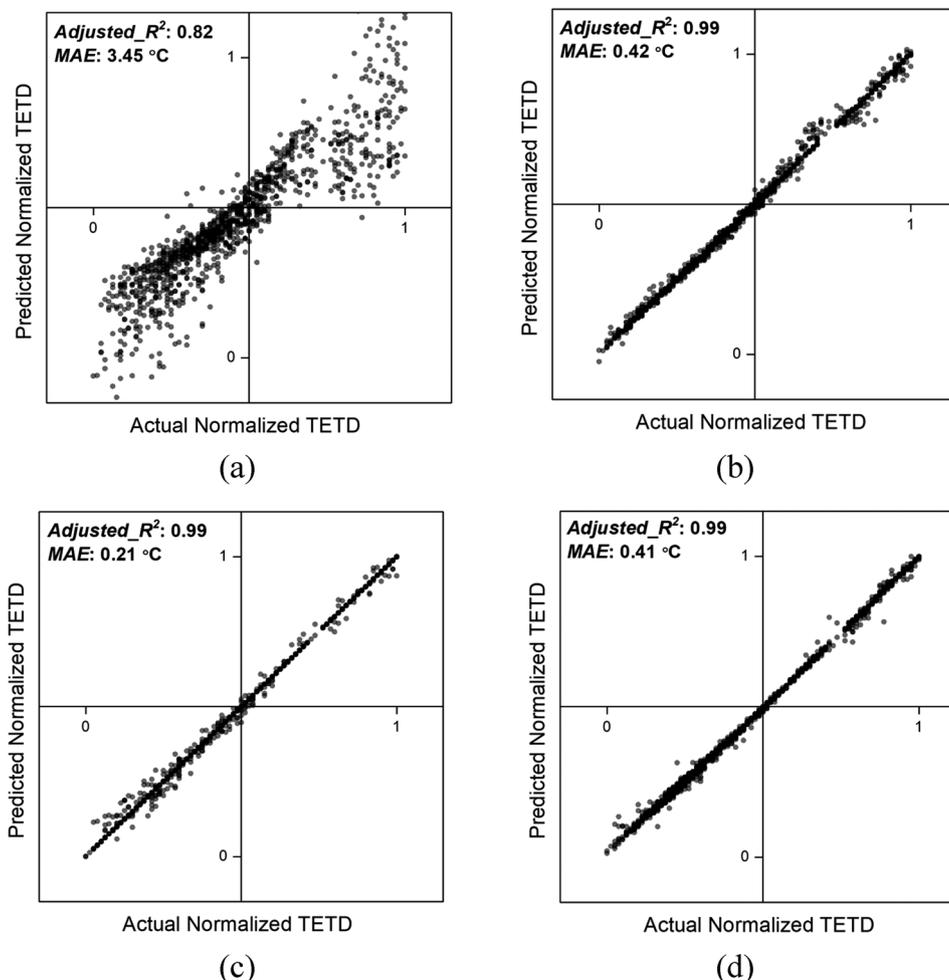Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 10 of 15

Fig. 8 Comparison between the true and predicted TETD calculated by **a** multiple regression, **b** SVM, **c** DT, and **d** DNN.

Fig. 8 exhibits the relationship between the true and predicted TETD. Here, the true values of the TETD are normalized ranging from 0 to 1, and the corresponding predicted TETD values are plotted. For the analysis of the quantitative measurements, from the overall regions identified through machine learning, an *adjusted_R²* of 0.99 was obtained, closely aligning with the ideal linear correlation coefficient of 1.0. On the other hand, the static method spreads the predicted values. Overall, every machine learning model exhibits a small-scale difference between the true and predicted values and an unbiased fit to the ideally upward trend.

Table 2 presents the performance of each regression model using performance indicators such as the *MAE*, *adjusted_R²*, and *MSE*. The computed *adjusted_R²* provides sufficient evidence of the outstanding performance of the machine learning models compared to that of multiple regression. Even though the *adjusted_R²* is an effective evaluation indicator based

on variability, overfitting the model or multiplicated independent variables could artificially inflate the result of the *adjusted_R²*. To overcome this limitation, another performance indicator, the *MAE*, evaluates the prediction errors with regard to the training and test data sets separately. More specifically, the TETD leads to residual stresses caused by the curling deformation subjected to vehicle loading. As a point of reference, a TETD value of 0.7 °C induces residual stresses between approximately 0.09 MPa and 0.14 MPa with standard concrete slab depths of 15 cm to 30 cm for JPCP. Additional discussion and insights on this matter can be found in Sect. 4.3.

The DT exhibits the lowest *MAE* on the training set but the highest on the test set. This proves that the model is overfitted to the training data, so it is not suitable for use on the actual testing data set. The SVM achieves the highest *MAE* on the training set; however, this model exhibits the smallest *MAE* difference between the training and testing sets. Thus, it could be

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 11 of 15

possible to expect the highest robustness from the SVM. Researchers and engineers will input the collected field data to the model suggested in this study, and they will acquire the predicted TETD in return. Therefore, the SVM is recommended because of its low *MAE* on the test data set represented by the field data. However, there are still too minor differences in the prediction errors to choose the best model distinctively. If candidate models show similar prediction errors, then the most robust model is the best for practical use, and this will be a great benefit to prospective users. Therefore, more refined studies, such as noise sensitivity analysis, are additionally introduced to quantify the robustness of each model. Moreover, the *MAE* calculates the actual difference value (degree, °C) between the true and predicted values along with all ranges of the TETD, which is meaningful in this study to intuitively provide the actual TETD differences.

It is mathematically proven that the machine learning model's approximation error to the true system is bounded. This means that the prediction performance is near-optimum when a rich enough data set is used for training. Therefore, the indistinguishable performance difference between the machine learning algorithms in Table 2 is natural. It also implies that all models converge to some level, so our data set is rich enough for analysis. However, the actual level of convergence to the true system can differ between models. In other words, the required richness of the training data set for each machine learning model is different. Therefore, the model that requires the smallest richness is the best for practical use since the cost is directly proportional to the number of data points.

4.2 Noise Sensitivity Analysis using Gaussian Noise-Added Adversarial Samples

This study pursues the noise-sensitivity analysis to evaluate the machine learning models based on this point of view. Since all models are equally trained on the same data set, it is possible to directly compare each model in terms of its robustness on the adversarial samples. The most converged model produces the smallest error on the adversarial samples. In brief, the most robust model can be considered the most efficient as well as the most practical model for future use.

In this analysis, Gaussian noise is added to generate adversarial samples. Perturbations are applied on only twelve vertical surface responses from deflection basins 1 and 2. The reason that surface responses are selected for this analysis involving added noise is the susceptibility of the nondestructive FWD test to signal noise and/or user error more than other features, which is assumed to correspond to aleatoric measurement errors, and thus, the twelve noises are considered

statistically independent under this assumption. This study defines the $q\%$ level noise for the ith vertical displacement $\varepsilon_i \sim N(0, \sigma_i), i = 1, \ldots, 12$, where $\sigma_i = 0.01 \cdot q \cdot \mu_i / \Phi^{-1}(0.975)$; μ_i is the mean value of the ith vertical displacement, and $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function of the standard normal. The standard deviation of the ith noise, σ_i, is designed to ensure that the ith noise, ε_i, is less than $q\%$ of the ith mean value ($= 0.01 \cdot q \cdot \mu_i$) with a probability of 95%. A total of 1179 random samples from Gaussian noise are generated and added to the original data set to produce a set of adversarial samples. This procedure, which creates a random data set, repeats until 100 sets of adversarial samples are generated. A total of 1179 multiplied by 100 adversarial samples (a total of 117,900 samples) are input into the trained models, and the same number of predicted TETDs is then produced.

Fig. 9 shows a plot of the 5–95%-tile prediction band—the interval from the 5%-tile to the 95%-tile predicted TETD out of 100 sets of adversarial samples—of three models. A noise-sensitivity analysis with different noise levels, i.e., 5%, 10%, and 15%, is performed. It is indicated that the DNN is the most noise-insensitive model and, therefore, it is the most robust model for prospective users. In addition, an average *MAE* from the 100 sets of adversarial samples is measured according to three models with three noise levels of 5%, 10%, and 15%. Overall, the *MAE* increase as the level of noise increases for all models. As proven in Fig. 9, the DNN is the most efficient model at all ranges of noise levels, and the *MAE* is measured at a degree of less than 1 °C. Conversely, the SVM has a high *MAE*, which is approximately 5 °C. To analyze the effect of the calculated *MAE* (e.g., 0.6 °C for the DNN and 4.5 °C for the SVM) on a concrete slab, the supplemental FE study addresses the effect of various TETDs on the residual stress of concrete pavement in the next section.

4.3 Effect of TETD Errors on Concrete Pavement

The presence of the TETD as the degree of total curling leads to residual stress on the concrete pavement before being subjected to the vehicle load. Thus, the accurate measurement of the TETD should be pursued to properly diagnose the structural condition of the concrete pavement. Through the noise-sensitivity analysis, the regression models TETD estimation errors, ranging from 0.6 °C for the DNN to 4.5 °C for the SVM. Therefore, it is prudent to determine the acceptable limits of miscalculated TETD to greatly benefit field design engineers and researchers. To accomplish this, the residual stress of the FE model is computed concerning the TETD. Since the structural design maintains the concrete stress to ensure a value below its strength
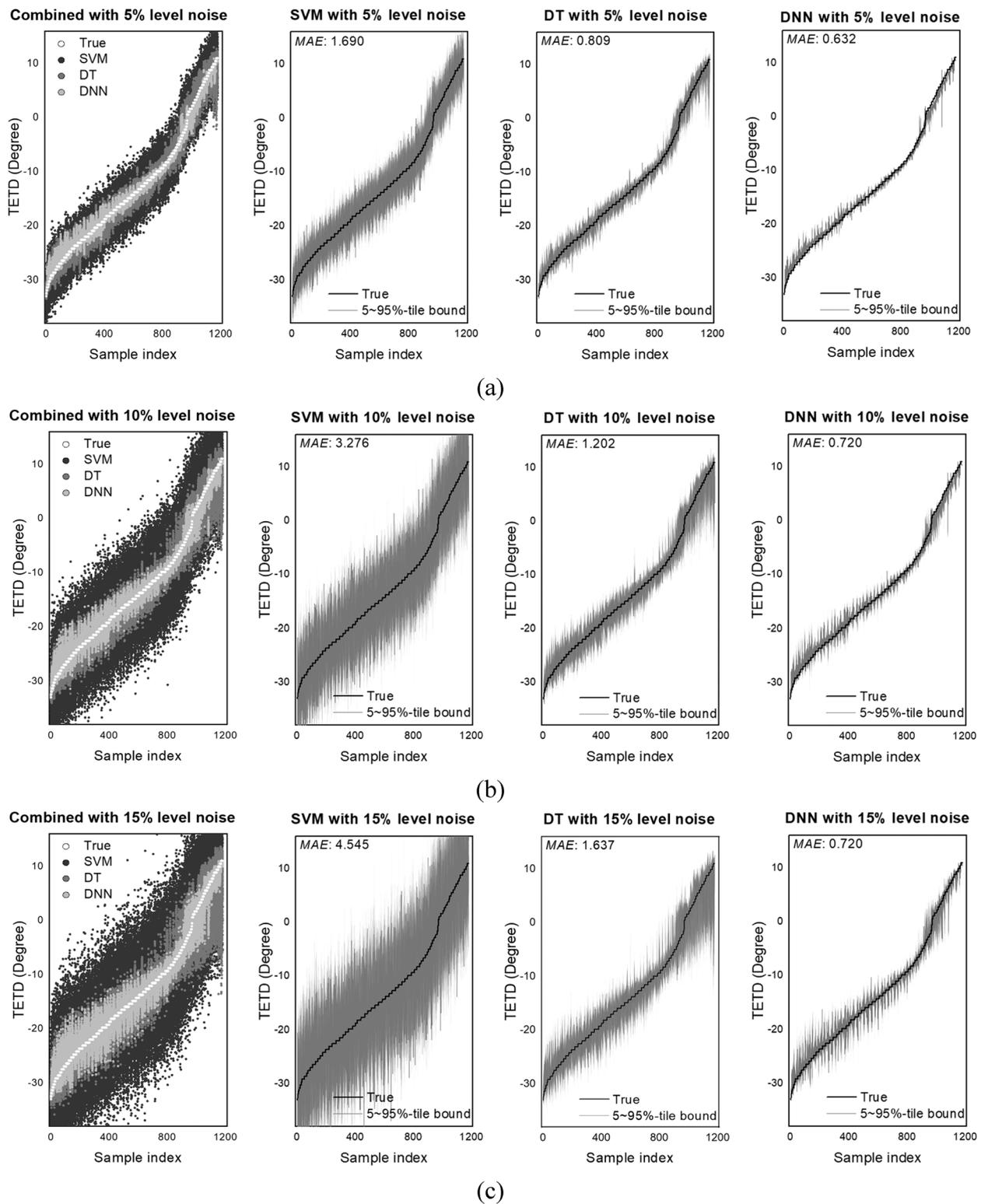
Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 12 of 15

Fig. 9 Plots of 5–95%-tile prediction band—the interval from the 5%-tile to the 95%-tile predicted TETD out of 100 sets of adversarial samples generated using a **a** Gaussian noise of 5% level, **b** Gaussian noise of 10% level, and **c** Gaussian noise of 15% level.
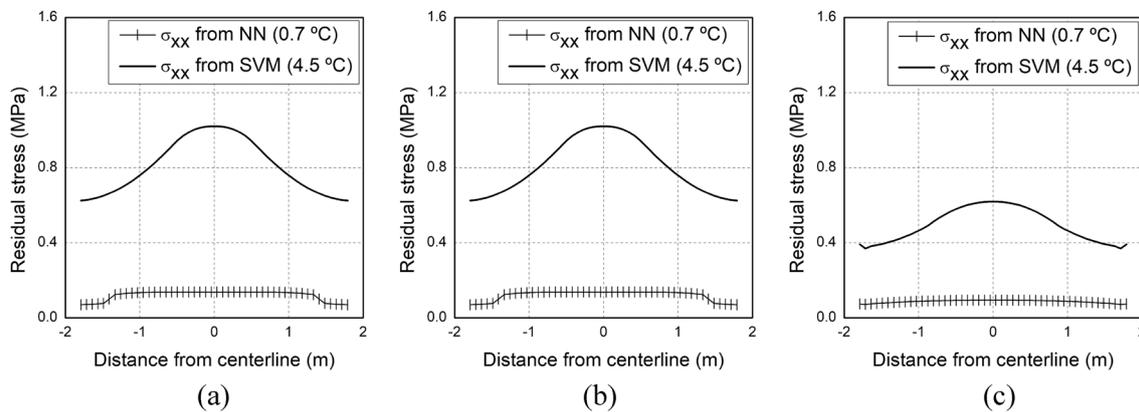
Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 13 of 15

Fig. 10 Comparison of residual stress from the computed TETD with the designed FE models: **a** a slab depth of 15 cm, **b** 23 cm, and **c** 30 cm.

multiplied by the designed safety factors, high-level residual stress could lead to the failure of the structural system or misjudgment during the structural analysis.

Fig. 10 shows the calculated residual stresses in the mid-transverse direction of the concrete slab from the FE models after applying two TETDs, 0.7 °C from the DNN with a noise level of 15% and 4.5 °C from the SVM with a noise level of 15%. These TETDs are chosen to consider conservative cases. Moreover, the analyzed FE models have three different slab depths, 15 cm, 23 cm, and 30 cm, because the depth of the slab plays a critical role in impacting the curling behavior. Therefore, more levels of residual stresses can be expected at each depth of the concrete slab. The degree of the TETD from the DNN yield unelevated maximum residual stresses ranging from 0.09 MPa to 0.14 MPa; however, the SVM not only produces maximum residual stresses between 0.62 MPa and 1.02 MPa but also exhibits an enormous gap of developed maximum stresses among the three different depths of FE models. This indicates that the error of the DNN has an insignificant influence on calculating stresses and is largely reliable for applying the various conditions of concrete pavement. However, the above is not true for the SVM. Although more parameters, such as the concrete mix and boundary conditions, are not investigated in this study, the result is still sufficient to provide insight into the outstanding performance of DNNs. Due to the limited scope of this study, only the slab depth is studied. However, further meta-studies should be performed to determine the refined level of error acceptance.

5 Conclusion and Recommendations

In this paper, we recommend the deep neural network as the most robust machine learning model to predict the TETD. Note that this result cannot be generalized to other civil engineering applications. To use the DNN, we need enough data to train, validate, and test. For some civil applications, collecting such data sets may be impossible due to the expensive and time-consuming experimental procedures involved. The required size of the data depends on the complexity of the applications, a larger database is needed to learn the complex relationships between the input variables and the target variable. Without knowing the complexity of the problem, training the deep neural network involves understanding the black box nature of the model, and there is a risk of overfitting, rending it impractical for real-world use. This limitation can only be overcome by combining our traditional understanding of civil applications into machine learning models, as practiced in this paper (Lunderg & Lee, 2017). As long as the DNN is trained, validated, and tested as demonstrated in this paper, the model will be fast and reliable in the field. For other civil applications, the model learning should be rigorously cross-validated based on a thorough fundamental understanding of the application, as introduced in this paper.

The main findings are summarized as follows:

- The created data set is used to successfully train prominent machine learning models, such as the SVM, DT, and DNN, with satisfactory levels of the *MAE.*
- Through the Gaussian noise-sensitivity analysis, the DNN is found to be the most robust model to predict the TETD of concrete pavement based its cost-effective and error-tolerant performance.
- The DNN with a noise-added level of 15% leads to negligible residual stress in all depths of concrete slabs. Based on all of the analyses, such as the Gaussian noise sensitivity and residual stress analy-

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 14 of 15

ses, the DNN is the most robust model to be recommended.

- The proposed methodology will allow users to determine the total curling of concrete slabs from conveniently collected field data that can be directly applied to the in-service road.
- Given the nature of civil engineering, structural design and testing data inherently come with a certain degree of uncertainty. This can affect the reliability of this methodology, which depends on the data sets for machine learning training. To address this, we have introduced noise sensitivity across varying noise levels, ranging from 5 to 15%, to address such uncertainty.

For future improvements of this methodology, the following aspects should be considered:

- In this paper, we introduce a pioneering methodology for determining the total curling intensity of concrete slabs of JPCP. Since curling is present the CRCP and PPCP, future studies should extend the scope of the concrete pavement system to check the applicability of this methodology. In this paper, the FE model is crucial for simulating the nondestructive FWD test. To better represent real-life situations, the FE model can be improved by applying the dynamic loading and actual size of the dowel bar at the joint.
- In this paper, we maintain a consistent number of input parameters for the data set. For future studies, it is strongly advised to segment data sets based on varying counts of input parameters.

Declarations

Competing interests
The authors declare that they have no competing interests.

References
AASHTO PP84. (2017). Standard practice for developing performance engineered concrete pavement mixtures (PP84). American Association of State Highway and Transportation Officials, Washington, D.C.

Alland, K. D., Vandenbossche, J. M., & Melo de Soursa, A. (2017). Daily cycles of temperature-independent curvature in jointed plain concrete pavements. *Journal of Transportation Engineering, Part a: System, 143*(8), 04017034.

ARA, Inc. (2007). Interim mechanistic-empirical pavement design guide manual of practice. National Cooperative Highway Research Program Project 1-37A, Campaign, IL.

Asbahan, R. E., & Vandenbossche, J. M. (2011). Effects of temperature and moisture gradients on slab deformation for joint plain concrete pavements. *Journal of Transportation Engineering, ASCE, 137*(8), 563–570.

Beckemeyer, C. A., Khazanovich, L., & Thomas, Y. H. (2002). Determining amount of built-in curling in jointed plain concrete pavement: Case study of Pennsylvania 1–80. *Transportation Research Record, 1809*, 85–92.

Brody, Z. A., Donnelly, C. A., & Vandenbossche, J. M. (2023). Temporal and spatial variation in drying shrinkage and thermal response of a jointed plain concrete pavement. *Transportation Research Record, 2677*(9), 81–93.

Ceylan, H., Yang, S., Gopalakrishnan, K., Kim, S., Taylor, P., & Alhasan, A. (2016). Impact of curling and warping on concrete pavement. Final Report, Iowa Department of Transportation, Ames, IA.

Chou, J., & Pham, A. (2013). Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength. *Construction and Building Materials, 49*, 554–563.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297.

FHWA. (2006). *Long term pavement performance program manual for falling weight deflectometer measurements*. Federal Highway Administration.

Franklin, R. E. (1969). *The effect of weather conditions on early strain in concrete slabs*. Transportation Research Board.

Greene, J. (2016). FDOT's concrete test road, Florida Department of Transportation, Gainesville, FL.

Hansen, W., Wei, Y., Smiley, D. L., Peng, Y., & Jensen, E. A. (2006). Effects of paving conditions on built-in curling and pavement performance. *International Journal of Pavement Engineering, 7*(4), 291–296.

Hiller, J. E., & Roesler, J. R. (2010). Simplified nonlinear temperature curling analysis for jointed concrete pavements. *Journal of Transportation Engineering, 136*(7), 654–663.

Huang, Y. H. (2003). *Pavement analysis and design* (2nd ed.). Pearson Prentice Hall.

Jeong, J., & Zollinger, D. G. (2003). Development of test methodology and model for evaluation of curing effectiveness in concrete pavement construction. *Transportation Research Record, 1861*, 17–25.

Johnson, A. M., Smith, B. C., & Gibson, L. W. (2010). Evaluating the effect of slab curling on IRI. South Carolina Concrete Pavements and the Federal highway Administration, Columbia, SC.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems, 30*, 3149–3157.

Lange, D., & Shin, H. C. (2001). Early age stresses and debonding in bonded concrete overlays. *Transportation Research Record, 1778*, 174–181.

Lundeberg, S. M. & Lee, S. I. (2017). A unified approach to interpreting model predictions. Neural Information Processing System 30 (NIPS), Long Beach, CA.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, T., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research, 12*, 2825–2830.

Poblete, M., Garcia, A., David, J., Ceza, P., & Espinosa, R. (1990). "Moisture effects on the behavior of PCC pavements." Proceedings of the 2nd International Workshop on the Theoretical Design of Concrete Pavements, Siquenza, Spain.

Han *et al. Int J Concr Struct Mater* (2024) 18:71

Page 15 of 15

Rao, S., & Roesler, J. R. (2005). Characterizing effective built-in curling from concrete pavement field measurements. *Journal of Transportation Engineering Part B, ASCE, 131*(4), 320–327.

Sridhar, R. K., Sandesh, P., & Amaranatha, R. M. (2022). Field experiments and numerical analysis of curling behavior of cast-in-situ short paneled concrete pavement on lean concrete base. *International Journal of Pavement Engineering, 23*(11), 3743–3756.

Tia, M., Subgranon, T., Chung, H. W., & Han, S. (2020). Mitigation of cracking in Florida structural concrete, Final Report, Florida Department of Transportation, Gainesville, FL.

Vandenbossche, J. M., Mu, F., Gutierrez, J. J., & Sherwood, J. (2011). An evaluation of the built-in temperature difference input parameter in the jointed plain concrete pavement cracking model of the Mechanistic-Empirical Pavement Design Guide. *International Journal of Pavement Engineering, 12*(3), 215–228.

Wei, Y., & Hansen, W. (2011). Characterization of moisture transport and its effect on deformations in jointed plain concrete pavement. *Transportation Research Record, 2240*, 9–15.

Yeon, J. H., Choi, S., Ha, S., & Won, M. C. (2013). Effects of creep and built-In curling on stress development of Portland cement concrete pavement under environmental loadings. *Journal of Transportation Engineering Part B, ASCE, 139*(2), 147–155.

Publisher's Note

Sangyoung Han is an Assistant Professor in the Department of ICT Integrated Ocean Smart City Engineering at Dong-A University, Busan, South Korea. He received his PhD from the University of Texas at Austin, Austin, TX, in 2022. His research interests include the concrete mixture characteristic and analysis of reinforced and prestressed concrete structures.

Taemin Heo is a Postdoctoral Fellow at the Oden Institute for Computational Engineering and Sciences at the University of Texas at Austin, US. His research area is to create data scientific frameworks that help us to make the best decisions for dealing with climate change challenges. He developed novel methods for time-dependent reliability and risk-based design framework accounting for nonstationarity and serial correlation embedded in the stochastic climate-ocean processes.

Chul Min Yeum is an Associate Professor in the Department of Civil and Environmental Engineering at University of Waterloo, Waterloo, Canada. His research interests include the smart structure, computer vision, augmented reality, machine learning, robotics, nondestructive testing, and sensing technologies.

Kukjoo Kim is a Professional Engineer in Defense Installations Agency, Ministry of Defense, South Korea. He received his PhD from the University of Florida in 2017. His research interests include the Ultra High Performance Concrete and analysis of protective structure against blast loadings.

Jongkwon Choi is an Assistant Professor in the Department of Civil and Environmental Engineering at Hongik University, Seoul, South Korea. He received his BS and MS from Seoul National University, Seoul, Korea, and PhD from the University of Texas at Austin, Austin, TX. His research interests include but are not limited to mechanical behavior, experiment, and analysis of reinforced concrete and prestressed concrete structures, and structural assessment of aging concrete structures.

Mang Tia is a Professor of civil engineering at the University of Florida, where he has been a faculty member since 1982. His research interests include concrete pavements and materials, asphalt pavements and materials, accelerated pavement testing, and instrumentation for pavements and materials research.