Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering) in press
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.A2400568

Machine learning for soil parameter inversion enhanced with Bayesian optimization

Anfeng HU[1], Chi WANG[1,2], Senlin XIE[1], Zhirong XIAO[3]✉, Tang LI[1], Ang XU[1]

[1]*Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang University, Hangzhou 310058, China.*
[2]*MOE Key Laboratory of Soft Soils and Geoenvironmental Engineering, Zhejiang University, Hangzhou 310058, China*
[3]*School of Civil Engineering and Architecture, Zhejiang University of Science and Technology, Hangzhou 310023, China*

Abstract: Machine learning has strong potential for soil settlement prediction, but determining hyperparameters for machine learning models is often intricate and laborious. Therefore, we apply Bayesian optimization to determine optimal hyperparameter combinations, enhancing the effectiveness of machine learning models for soil parameter inversion. The machine learning models are trained using numerical simulation data generated with the Modified Cam-Clay model in ABAQUS, and their performance is evaluated using ground settlement monitoring data from an airport runway. Five optimized machine learning models—decision tree, random forest, support vector regression, deep neural network, and one-dimensional convolutional neural network—are compared in terms of their accuracy for soil parameter inversion and settlement prediction. The results indicate that Bayesian optimization efficiently utilizes prior knowledge to identify optimal hyperparameters, significantly improving model performance. Among the evaluated models, the one-dimensional convolutional neural network achieves the highest accuracy in soil parameter inversion, generating settlement predictions that closely match real monitoring data. These findings demonstrate the effectiveness of the proposed approach for soil parameter inversion and settlement prediction, and how Bayesian optimization can refine the model selection process.

Key words: ABAQUS; Bayesian optimization; Machine learning algorithms; Parameter inversion; Settlement prediction

Uncorrected

1 Introduction

Airports are integral parts of public transport infrastructure, and airport runways are particularly sensitive to uneven surface settlement (Wu et al., 2020). To ensure safe operation, it is crucial to carry out deformation monitoring and settlement predictions on airport runways (Yao et al., 2020). Traditional settlement prediction methods can be divided into three types: curve-fitting, theoretical calculation, and numerical simulation. The curve-fitting method involves calculating the relationship between settlement and time through field-measured data (Asaoka, 1978; Sridharan et al., 1987). However, the chosen shape of the settlement-fitting curve depends on experience-based human judgement. The theoretical calculation method employs the layer-wise summation method along with various theoretical models to estimate the foundation's final settlement and consolidation degree, as well as the settlement's development pattern (Zhou et al., 2020). Nevertheless, because of the difficulty of obtaining undisturbed soil parameters and the complexity of construction sites, such calculated results often differ significantly from reality. The numerical simulation method combines various soil constitutive models and consolidation theory to calculate settlement (Lester et al., 2019). This method is widely utilized due to its broad applicability, high accuracy, and flexibility. However, there is often a significant challenge in calibrating the parameters that define the constitutive model (Hu et al., 2023).

Traditionally, soil parameters are determined using two methods: laboratory tests and field tests. Due to the disparity between the soil used in laboratory tests and soil in the real environment, measured parameterS deviate from the true values. The parameters obtained

✉ Zhirong XIAO, 100106@zust.edu.cn
Anfeng HU, https://orcid.org/0000-0002-3278-0238
Zhirong XIAO, https://orcid.org/0000-0003-3723-3978

Received Dec. 10, 2024; Revision accepted Mar. 26, 2025; Crosschecked

© Zhejiang University Press 2025

by field in-situ tests usually display localized characteristics and are affected by the test methods, equipment accuracy, and soil heterogeneity (Hu et al., 2021). In recent years, with the rapid increase in data volume and the continuous improvement of computer hardware, machine learning (ML) has been widely applied in the field of geotechnical engineering (Benzaamia et al., 2024a; Xie et al., 2024). It has been applied for tasks such as maximum surface settlement prediction (Chen et al., 2019; Mahmoodzadeh et al., 2020), slope stability evaluation (Qi and Tang, 2018; Wang et al., 2024), soil mechanical parameter estimation (Meng et al., 2021; Zhang et al., 2021a), tunnel excavation displacement (Huang et al., 2022; Zhang et al., 2022), and deformation control and shield tunneling guidance (Pan et al., 2023; Glab et al., 2024). Based on field monitoring data, various ML algorithms have been used to predict settlement and control development, such as the Back-Propagation Neural Network (BPNN) (Ye et al., 2022; Bardhan et al., 2024), Random Forest (RF) (Tang and Na, 2021; Zhang et al., 2021b), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) (Zhang et al., 2017). However, research on inversion for soil mechanical parameters based on airport runway settlement monitoring data remains scarce. The powerful nonlinear fitting ability of ML algorithms will help to accurately obtain soil layer parameters directly from settlement data, and enable robust predictions of airport runway settlement development. This can in turn aid the formulation of safety and control measures.

Hyperparameters are parameters set before training an ML model, which are usually set manually and cannot be learned directly from training data (Victoria and Maragatham, 2021). The hyperparameter space is typically vast and interdependent. To find the optimal combination of hyperparameters efficiently, ML practitioners use a variety of strategies, such as Grid Search (GS) (Fayed and Atiya, 2019), Random Search (RS) (Bergstra and Bengio, 2012), and Evolutionary Algorithms (EAs). The commonly used EAs include the Genetic Algorithm (GA) (Wijesinghe et al., 2022), Particle Swarm Optimization (PSO) (Hasanipanah et al., 2016), Ant Colony Optimization (ACO) (Wang and Vanapalli, 2024), and Simulated Annealing (SA) (Ma et al., 2023). However, the GS algorithm is susceptible to gradient explosion, while the RS algorithm struggles to balance efficiency and precision. EAs demand a substantial quantity of initial points, and yet its optimization efficiency remains suboptimal. Currently, most hyperparameter optimization methods fail to effectively leverage optimization experiences from other datasets. Consequently, within a constrained optimization timeframe, they are unable to fully explore the performance potential of different models, thus restricting their applicability in geotechnical engineering.

Bayesian optimization is a global optimization method based on probabilistic models, which excels at leveraging known experimental data points. Using limited observational data, it selectively identifies the most promising parameters for subsequent iterations. Ultimately, Bayesian optimization efficiently finds the global optimum within a specified number of iterations (Akiba et al., 2019). This characteristic has led to its widespread application in the field of geotechnical engineering, particularly for inversion of soil parameters (Yin et al., 2018; Gu et al., 2023; Zhang et al., 2024b). For instance, a Bayesian method based on Markov Chain Monte Carlo (MCMC) simulation was used to identify the optimal constitutive model for a sandy soil (Jin et al., 2019b), and a Bayesian method based on Differential Evolution Transitional MCMC could accurately determine the parameters of advanced soil models (Jin et al., 2019a). Moreover, the Bayesian Optimization – Genetic Algorithm (BO-GA) model has demonstrated superior performance in soil parameter calibration compared to the traditional GA (Liu et al., 2024). Additionally, an XGBoost (Extreme Gradient Boosting) model enhanced by Bayesian optimization was utilized to accurately predict the undrained shear strength (USS) of soft clays, avoiding inefficient brute-force searches (Zhang et al., 2021c).

In this study we propose an enhanced ML method for soil parameter inversion based on Bayesian optimization. First, a large dataset for training the ML models is generated through numerical modeling using the nonlinear finite element software ABAQUS, adopting the Modified Cam-Clay (MCC) constitutive model widely utilized in geotechnical engineering. Second, five ML models are established and their optimal hyperparameter combinations are determined using Bayesian optimization, followed by comparisons of their performance on the testing dataset. Third,

Unedited

the soil parameters required to predict settlement are inverted for using the optimized ML models, based on monitoring settlement data from the Xiamen Xiang'an Airport project in China. Finally, the inverted parameters from the five ML models are input to the numerical model, and respectively compared with the settlement monitoring data. Fig. 1 delineates the overall workflow of our method.

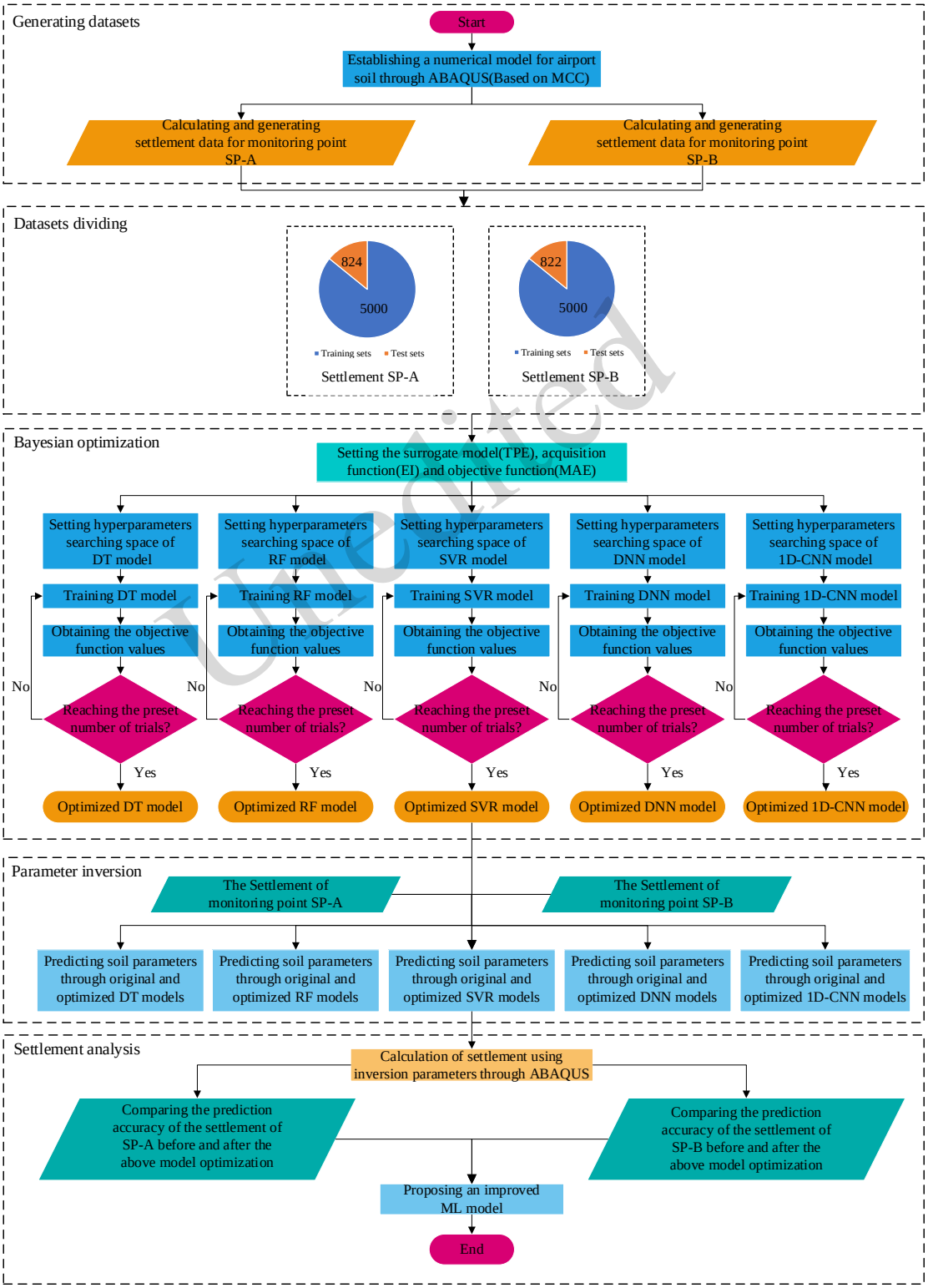

Fig. 1 The workflow of the proposed framework for obtaining the optimal soil parameter inversion mode

2 Bayesian optimization

Bayesian Optimization (BO) employs a surrogate model to approximate an objective function, and utilizes an acquisition function to balance exploration and exploitation so as to achieve efficient global optimization. In this study, the Tree-structured Parzen Estimator (TPE) serves as the surrogate model, which divides observed data points into two groups (optimal and suboptimal) and constructs Truncated Gaussian Mixture Models (TGMM) for each group. This approach enables flexible modeling of complex data distributions, while reducing computational complexity and enhancing robustness to noise. Eq. (1) defines the mathematical framework of the modeling process.

$$p\left(x|\ y\right)=\begin{cases} l\left(x\right) & \text{if } y<y^{*} \\ g\left(x\right) & \text{if } y\geq y^{*} \end{cases} \tag{1}$$

where $l(x)$ is the density function formed by the observed value $\{x^{(i)}\}$, with its corresponding loss $f(x^{(i)})$ being less than y^{*}, and $g(x)$ is the density function formed from the remaining observations. y^{*} corresponds to a percentile γ of the observed values of y, and its range is (0,1).

The mathematical expression for a truncated Gaussian mixture model consisting of K Gaussian distribution components can be formulated as follows:

$$p(x)=\sum_{k=1}^{K}\frac{\pi_k}{(2\pi)^{d/2}\left|\Sigma_k\right|^{1/2}}\exp\left(-\frac{1}{2}\left(x-\mu_k\right)^{T}\Sigma_k^{-1}\left(x-\mu_k\right)\right)I(x), \tag{2}$$

$$I\left(x\right)=\begin{cases} 1 & \text{if True} \\ 0 & \text{if False} \end{cases} \tag{3}$$

where x represents the input vector, and d signifies the vector's dimension. π_k stands for the mixing coefficient of the kth Gaussian distribution, and $\left|\Sigma_k\right|$ denotes the determinant of the covariance matrix $\left|\Sigma_k\right|$. μ_k represents the mean of the kth Gaussian distribution, and $I(x)$ denotes the indicator function utilized for model truncation.

The TPE further adopts the Expected Improvement (EI) acquisition function to guide the search process toward promising regions. In this way, we can effectively combine probabilistic modeling with adaptive sampling strategies to optimize objective functions in bounded domains (Jones, 2001).

$$\mathrm{EI}_{y^{*}}\left(x\right)=\int_{-\infty}^{y^{*}}\left(y^{*}-y\right)p\left(y\mid x\right)\mathrm{d}y=\int_{-\infty}^{y^{*}}\left(y^{*}-y\right)\frac{p\left(x\mid y\right)p\left(y\right)}{p\left(x\right)}\mathrm{d}y. \tag{4}$$

3 Datasets and ML models

3.1 Data source

ML models with robust performance and generalization capabilities typically require a substantial amount of training data. Given the limited availability of on-site settlement monitoring data, a viable alternative approach involves generating a synthetic dataset using numerical methods, and using this synthetic data for model training. This approach can be complemented by evaluating the model's accuracy with real settlement monitoring data, thereby enhancing the precision and robustness of the settlement prediction (Xie et al., 2025). Consequently, the data in this study encompasses two primary components: numerical simulation data generated by the finite element software ABAQUS, and settlement monitoring data from the Xiang'an Airport in Xiamen, China.

3.1.1 Monitoring dataset

In order to wholistically comprehend the distribution and developmental trends of runway settlement at the airport, monitoring points were strategically positioned at critical locations along the runway, as depicted in Fig. 2. The settlement development records from two distantly positioned monitoring points on the airport runway, SP-A and SP-B, were chosen to validate the ML-based parameter inversion model. The inverted "T" markers in Fig. 2 symbolize settlement monitoring plates located at the base of the load-bearing sand layer.

Unedited

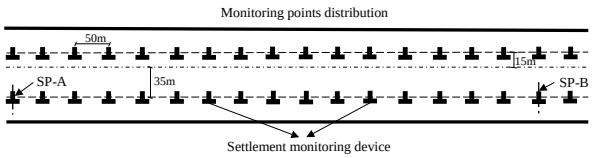

Fig. 2 The distribution of monitoring points along the runway

3.1.2 Numerical simulation dataset

Because of substantial ground settlement in the studied region, we incorporate the effects of significant soil deformation into the finite element modeling. The numerical modeling is based on rational assumptions, including subsoil stratification, soil homogenization, construction conditions, incompressibility of pore fluids, and the incompressibility of fully weathered granite layers, ensuring that the model accurately portrays real-world conditions. Fig. 3 illustrates a geological cross-section diagram of the monitoring points at the Xiang'an Airport.

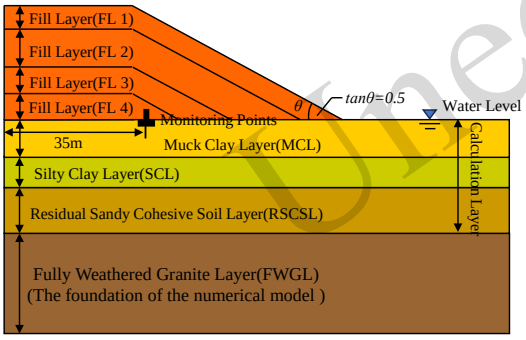

Fig. 3 A geological cross-sectional diagram with the monitoring points

The consolidation settlement of the airport runway under the influence of superimposed preloading is regarded as a plane strain problem. Various combinations of soil parameters are considered, and a semi-structured model is established using symmetry. For practical engineering considerations, the MCC model is employed to describe the relationship between soil stress and deformation. The foundation soil is saturated, and the elements are defined as 4-node plane strain pore pressure elements. To ensure computational efficiency and accuracy, the mesh size is set to 0.02D (where D represents the maximum length in the horizontal direction). The finite element models for cross-sections SP-A and SP-B are comprised of 339 and 386 nodes, respectively. Regarding the model's boundary conditions, the bottom is constrained in both the horizontal and vertical directions, while the left and right sides are constrained in the horizontal direction. Fig. 4 provides a visualization of the boundary conditions and numerical mesh. The load consists of four layers of fill (FL), and the subsoil compression layer comprises three parts: the muck clay layer (MCL), the silty clay layer (SCL), and the residual sandy cohesive soil layer (RSCSL). Table S1 presents the geometric and physical properties of these different soil layers.

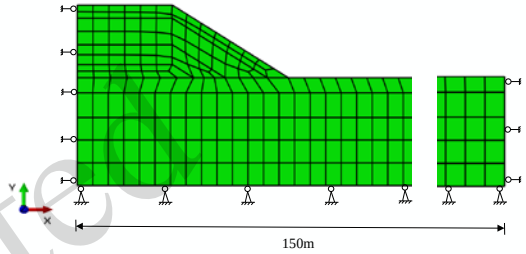

Fig. 4 Numerical model and constraint condition of the airport soil layers

The numerical simulation dataset is automatically generated in ABAQUS 2019 by importing prepared FEM model scripts. Different settlement curves are created by defining various ranges of soil constitutive parameters within the script. Table S2 presents the typical ranges of soil constitutive parameters as outlined in the geological survey report.

The numerical simulation dataset consists of 5,822 and 5,824 data points for points SP-A and SP-B, respectively. The random sampling method is used to divide the dataset into training and testing sets. Fig. 5 illustrates the dataset partitioning.

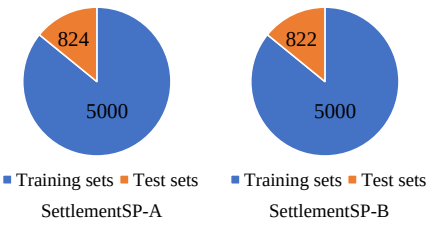

Fig. 5 The dataset partitioning

3.2 Data pre-processing

The generated numerical simulation dataset was normalized to mitigate potential biases in model pre-

Unedited

dictions due to variance among different features. All datasets were normalized within the range of (0, 1) according to the following scaling rule:

$$y_{\text{zoom}}^{i} = y^{i} \times K, i = 1, 2, \cdots, k \tag{5}$$

where y_{zoom}^{i}, y^{i}, K, and k represent the scaled label value, the original label value, the scaling constant, and the label dimension of the data, respectively.

Given the significant impact of the MCL on the overall settlement development, in this study we consider four key soil parameters within the MCL that affect settlement development. Table 1 shows the scaling constants corresponding to the four labels of the SP-A and SP-B datasets.

Table 1 The predicted labels along with their respective scaling constants

Label	λ_{MCL} (i = 1)	κ_{MCL} (i = 2)	M_{MCL} (i = 3)	k_{MCL} (i = 4)
C	4	40	1	10000

3.3 ML algorithms

In order to obtain the model with the best inversion performance on the datasets, we select five representative ML models. Here, the underlying principles of the models are first outlined.

3.3.1 Decision Tree

Utilizing the Classification and Regression Tree (CART) algorithm, a regression tree is constructed where the internal node features take on values of "yes" and "no" (Lewis, 2000), forming a binary tree structure. Fig. S1 illustrates the principles underlying this algorithm.

The regression tree algorithm operates by recursively splitting data through binary segmentation, starting from the root node and iterating until a termination condition is met. To mitigate overfitting, the CART algorithm prunes the tree model based on whether the loss function of the decision tree (DT) decreases. The expression for the loss function in the CART algorithm is given in Eq. (S1) of the Supplementary Material.

3.3.2 Random Forest

The Random Forest (RF) algorithm is an ensemble learning method involving the construction of multiple decision trees, which are then combined for predictive analysis. Each tree is trained independently, typically utilizing bootstrap sampling to establish distinct training datasets (Breiman, 2001). During the prediction phase, the Random Forest algorithm aggregates the predictions from all trees, commonly through averaging or voting, to produce an ultimate prediction. This approach frequently yields high predictive accuracy and robust performance, particularly when handling multidimensional datasets. Fig. S2 shows a visualization of the random forest workflow as applied to regression problems.

3.3.3 Support Vector Regression

Support Vector Regression (SVR), a variant of Support Vector Machine (SVM) tailored for regression problems, is depicted in Figure S3 along with its distinctions from SVM. In contrast to SVM, SVR aims to obtain a regression model where $f(x)$ and y are as close as possible, with ω and b representing parameters to be determined (Zhang et al., 2024a).

SVR employs a kernel function (commonly used types are detailed in the Supplementary Material) to map samples from the original space into a higher-dimensional feature space, thereby enabling linear separability in the transformed space (Boser et al., 1992).

3.3.4 Deep Neural Network

The fully connected deep neural network (DNN) is a classic neural network model belonging to the category of feedforward neural networks. It mainly consists of an input layer, hidden layers, and an output layer, with multiple neurons potentially included in each hidden layer. The input layer only receives external input without applying any function, while the neurons in the hidden and output layers process the data, with the final result being output by the neurons in the output layer (Cheng et al., 2023). Various neural networks can be formed based on the number of layers, the number of neurons in each layer, and the method of information propagation between layers. Fig. S4 presents a visualization of the DNN model adopted in this study.

3.3.5 1D Convolutional Neural Network

Convolutional Neural Networks (CNN) are a cat-

Unedited

egory of feedforward neural networks, which primarily consist of convolutional layers, pooling layers, and fully connected layers (Abbas et al., 2023). Within these convolutional layers, a computation known as the dot product is performed between specific regions of the input data and a weight matrix referred to as a filter. The function of the pooling layers is to reduce the spatial dimensions of the data, thereby decreasing its complexity without diminishing the network's depth. After the data undergoes a series of convolutional and pooling operations, it is processed by the fully connected layers. Here, the multitude of signals from the previous stages are integrated into a unified set of signals through connection operations, culminating in the network's output.

Unlike traditional CNNs used for image processing, the one-dimensional Convolutional Neural Network (1D-CNN) executes convolution operations along a single dimension (Chaaban et al., 2024). It can effectively capture local patterns and features within the input data, demonstrating exceptional performance in numerous sequential and time series data tasks. Fig. S5 illustrates the architecture of the 1D-CNN used in this study.

Unedited

4 Hyperparameter optimization

4.1 Selection of hyperparameters

The optimization of hyperparameters is a critical step in both the construction and refinement of ML models. However, this process does not require optimizing all possible hyperparameters. Instead, optimizing only the hyperparameters that most strongly impact model performance is often more effective (Snoek et al., 2012). In this study we use the Python library scikit-learn 1.0.2 to construct DT, RF, and SVR models, and Tensorflow-gpu 2.2.0 and Keras 2.3.1 to build DNN and 1D-CNN models.

Table S3 delineates the specific hyperparameters that require optimization for each tested ML model. For DT and RF models, “Max depth” specifies the maximum allowable depth of the tree, while “Min samples split” and “Min samples leaf” set the minimum sample thresholds for node splitting and terminal leaves, respectively. “Estimators” in the context of RF denotes the number of decision trees within the ensemble. In the SVR model, *C* represents the regularization strength, “Kernel” determines the feature space mapping strategy, γ_G controls the influence radius of samples in Gaussian kernels, “Degree” defines the polynomial order for polynomial kernels, and “Epsilon” sets the error tolerance margin. For the DNN model, “N_neurons” indicates the number of neurons in the fully connected layers, “Activation” selects nonlinear functions, “Learning rate” adjusts the gradient descent step size, “Dropout” applies regularization via stochastic neuron deactivation, “Epochs” sets the total number of training iterations, and “Batch size” determines the number of samples processed in each weight update. Unlike the DNN, the 1D-CNN model introduces the parameter “N_kernel” to specify the number of convolutional filters per layer, and “N_Dense” to configure neurons in the final dense layer. Architecturally, the DNN employs three fully connected hidden layers, while the 1D-CNN incorporates three convolutional layers with a fixed kernel size of 9 and a stride of 1 to enhance localized feature extraction. Detailed parameter selection criteria and rationales for the chosen architecture are provided in the Supplementary Material.

4.2 Determination of search scope

After identifying the hyperparameters that require optimization, it is essential to conduct a sensitivity analysis. This analysis provides insights into the influence of each hyperparameter on the model's performance, thereby facilitating determination of an appropriate search space for hyperparameter tuning, and ultimately enhancing the model's performance.

Tables S4 and S5 present the ranges of hyperparameters used for our sensitivity analysis of the RF and 1D-CNN models, respectively. Furthermore, Tables S6, S7, and S8 delineate the search ranges for hyperparameters specific to each type of model. Notably, in scikit-learn, the SVR model can only estimate one parameter at a time. Therefore, we constructed four SVR models for the SP-A and SP-B datasets, respectively.

Fig. S6 presents the results of the hyperparameter sensitivity analysis of the RF and 1D-CNN models, using the SP-A and SP-B datasets. It is important to note that the outcomes of a hyperparameter sensitivity analysis cannot be directly regarded as final optimal model results, because hyperparameters often

exhibit interdependencies and mutual influences. Therefore, by employing Bayesian optimization to simultaneously adjust multiple hyperparameters and evaluate the performance of each combination, the optimal set of hyperparameters for a given model can be identified.

4.3 Objective function

The objective function is the criterion for evaluating the performance of different hyperparameter configurations during optimization. Utilizing Mean Absolute Error (MAE) as the loss function, the optimal hyperparameter configuration is determined by minimizing the validation loss on the test dataset. Eq. (S3) gives the expression for MAE.

4.4 Optimization process

Hyperparameter optimization exhibits diminishing returns, with the most significant improvements often occurring in early optimization stages. Considering computational costs and memory consumption, and aiming to balance computational efficiency with thorough searching, the number of trials for the Bayesian hyperparameter optimization is set to 100. The MAE of the five ML models on the test dataset is recorded after each optimization. The optimization processes for each model are illustrated in Fig. 6. Tables 2-5 present the hyperparameter configurations of the optimized models on the SP-A and SP-B datasets.

Table 2 The optimized hyperparameter configurations for SVR on the SP-A dataset

SVR	SP-A dataset			degree	epsilon
	C	kernel	γ_G		
Model1	1.0	RBF	0.003	/	/
Model2	10	Poly	/	4	0.01
Model3	15	Poly	/	2	0.04
Model4	1.0	Poly	/	2	0.05

Table 3 The optimized hyperparameter configurations for SVR on the SP-B dataset

SVR	SP-B dataset			degree	epsilon
	C	kernel	γ_G		
Model1	10	Poly	/	4	0.01
Model2	8	Poly	/	3	0.03
Model3	10	Poly	/	2	0.01
Model4	10	Poly	/	3	0.04

Table 4 The optimized hyperparameter configurations for DT, RF, DNN, and 1D-CNN on the SP-A dataset

Model	Optimized hyperparameters on the SP-A dataset
DT	Max depth=8, Min samples leaf=2, Min samples split=7
RF	Max depth=10, Min samples leaf=1, Min samples split=3, Estimators=41
DNN	N_neuronsi(i=1,2,3)=28/20/32, Activation=ReLU, Learning rate=0.0008, Batch size=24, Dropout = 0.2, Epochs=1000
1D-CNN	N_neuronsi(i=1,2,3)=16/20/28, N_dense=36, Activation=ELU, Learning rate=0.0008, Batch size=40, Dropout=0.4, Epochs=1200

Table 5 The optimized hyperparameter configurations for DT, RF, DNN, and 1D-CNN on the SP-B dataset

Model	Optimized hyperparameters on the SP-B dataset
DT	Max depth=11, Min samples leaf=1, Min samples split=4
RF	Max depth=13, Min samples leaf=1, Min samples split=4, Estimators=71
DNN	N_neuronsi(i=1,2,3)=20/20/32, Activation=ReLU, Learning rate=0.00002 Batch size=20, Dropout = 0.4, Epochs=1200
1D-CNN	N_neuronsi(i=1,2,3)=32/28/28, N_dense=48, Activation=ELU, Learning rate=0.0002, Batch size=24, Dropout=0.2, Epochs=1200

Fig. 6 illustrates the variation of the MAE with the number of trials during the Bayesian optimization process for each ML model. The MAE after each optimization for the models is represented by solid black circles, while the minimum value of the objective function during the optimization process is indicated by solid red circles, corresponding to the hyperparameter combination that is considered optimal for the given ML model. The sharp initial decrease in the objective function demonstrates the high efficiency of the Bayesian optimization in identifying suitable hyperparameter combinations within a limited search space. The non-monotonic decrease of the objective function during optimization is due to the Bayesian optimization's strategy of exploring new potential hyperparameter combinations while seeking to minimize the objective function. This approach helps to avoid getting trapped in local optimal solutions, and increases the likelihood of finding the global optimal solution.

Unedited

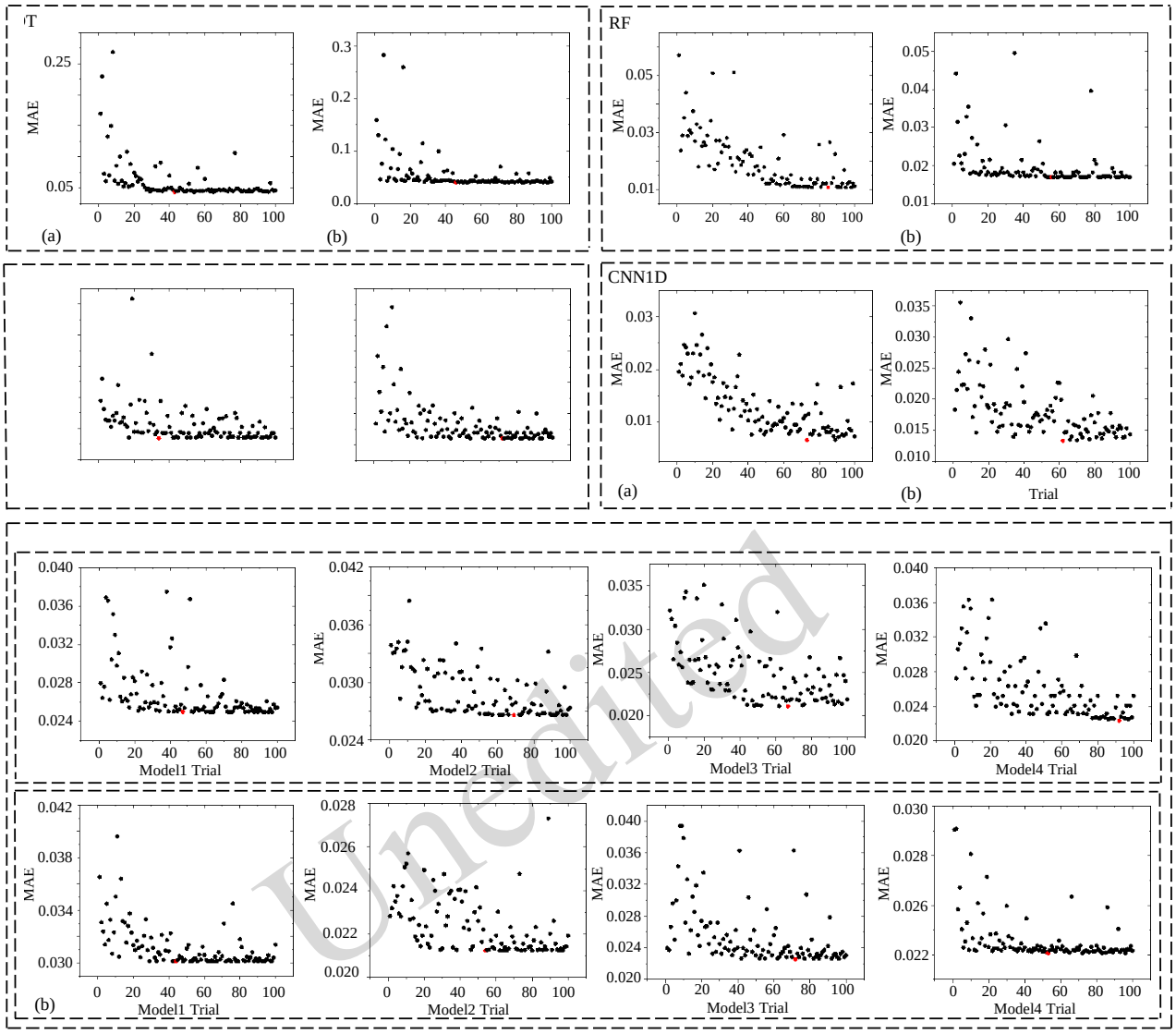

Fig. 6 The variation of MAE during the Bayesian optimization process for the five ML models. (a) The optimization process on SP-A; (b) The optimization process on SP-B

5 Application of optimized models

5.1 Performance of the ML models

Following Bayesian optimization, five optimal ML models are obtained according to minimization of the objective function. To evaluate the performance of these models more reliably across the datasets, metrics such as Mean Absolute Percentage Error (MAPE), R^2, Root Mean Squared Error (RMSE), and the a20-index (Asteris et al., 2024) are applied, with their mathematical definitions provided in Eqs. (S4-S6). Tables S9 and S10 present the evaluation results of the optimized models on the SP-A and SP-B datasets, respectively.

In Tables S9 and S10, the R^2 error ratios among the five ML models are all within 10%. Additionally, other metrics such as MAPE and RMSE show minimal differences between the training and testing datasets. These results suggest that the models exhibit small prediction errors and demonstrate consistent performance on both the training and test data, indicating no significant overfitting (Asteris et al., 2021; Benzaamia et al., 2024b). The a20-index values indicate that the prediction accuracy of all models is notably high following Bayesian optimization. All prediction results from the optimized RF, DNN, and 1D-CNN models on the testing datasets exhibit errors within 20%, demonstrating their consistent and reliable performance.

Figs. 7 and 8 depict the performance of the five models on the test datasets. Among the five optimized models, the 1D-CNN demonstrates the highest prediction accuracy, with its results on the test dataset closely aligning with the true values. In contrast, the DNN exhibits more scattered predictions. The DT model

performs the worst in terms of both accuracy and robustness. Notably, the ML models show significant variability in their predictions of the parameter M_{MCL}, which is a critical factor influencing the development of airport soil settlement. Given this observation, the 1D-CNN stands out for its ability to accurately identify and predict the key parameters affecting soil settlement, highlighting its superior performance.

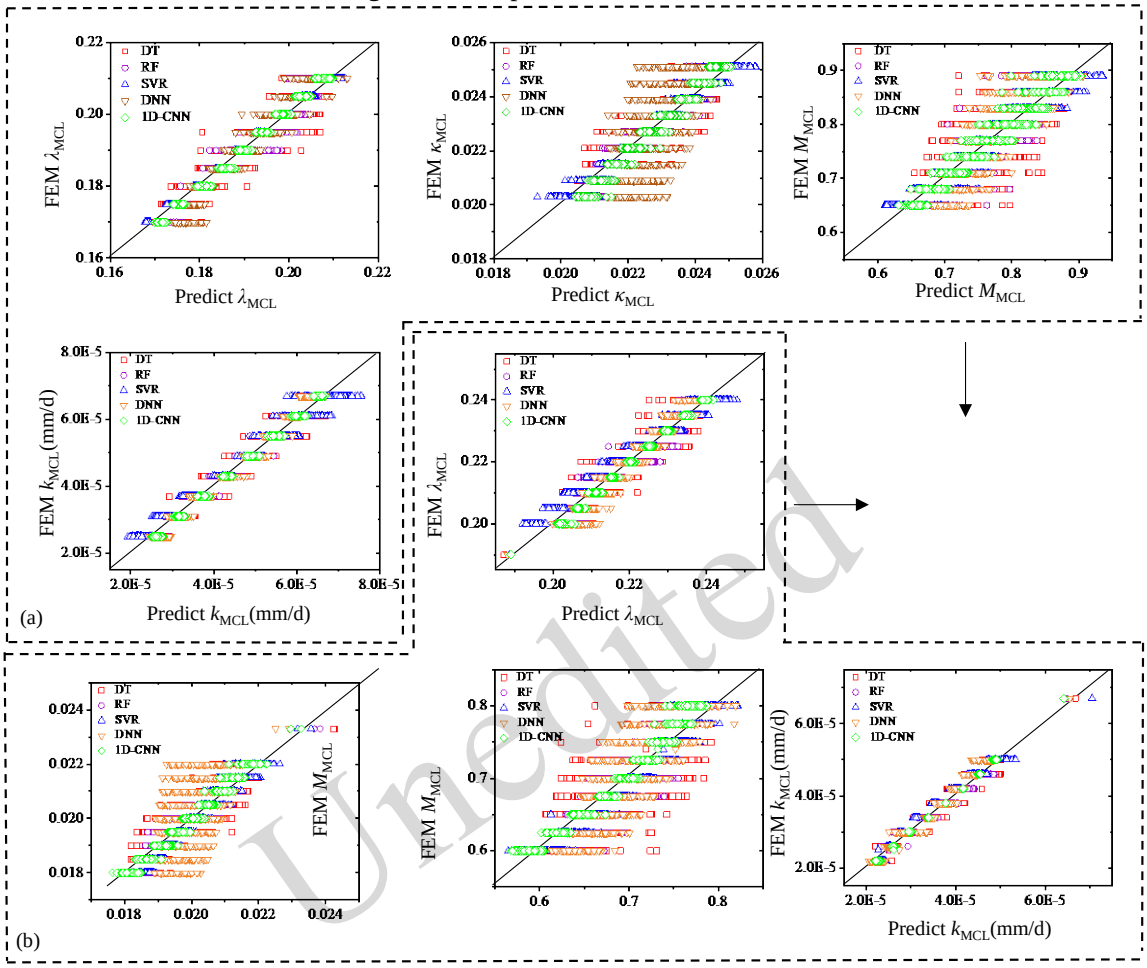

Fig. 7 Comparison of the true and predicted results of each ML model on the test datasets. (a) SP-A; (b) SP-B

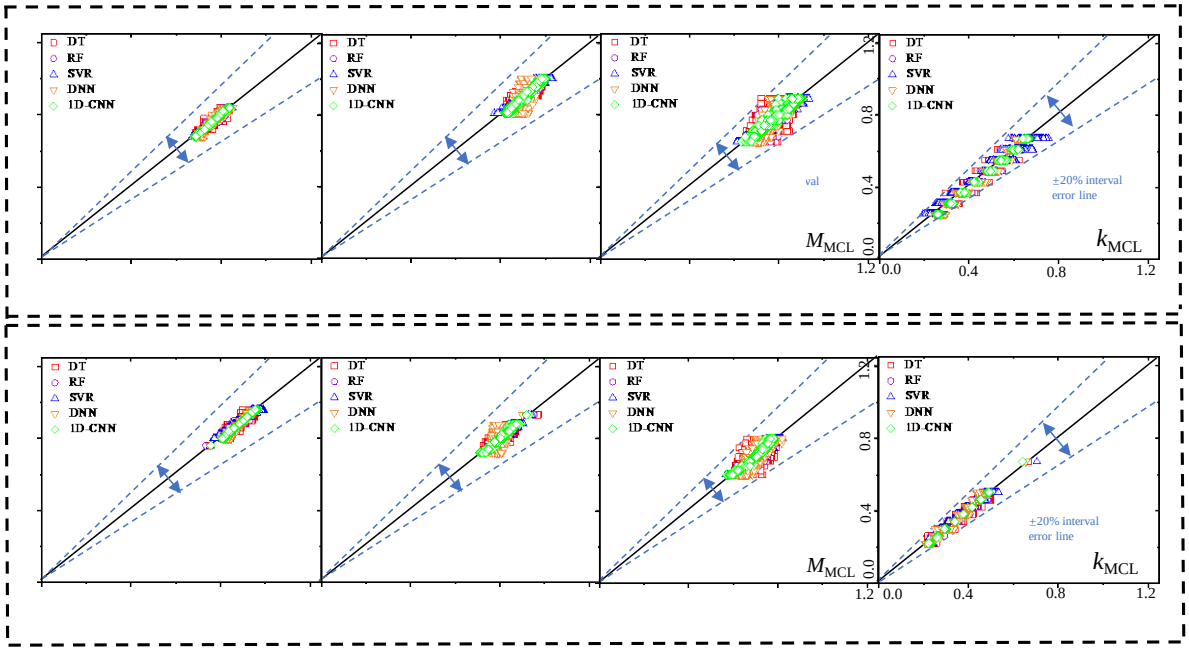

Fig. 8 Normalized performance of each ML model on the test datasets. (a) SP-A; (b) SP-B

Unedited

5.2 Validation of the parametric inversion models in real cases

To assess the models' generalization ability, real monitoring data from points SP-A and SP-B at the Xiamen Xiang'an Airport are employed for soil parameter inversion. The performance of each model on this real-world dataset is compared before and after Bayesian optimization, enabling a robust evaluation of their predictive capabilities and the effectiveness of the optimization in enhancing accuracy and reliability.

Real settlement data from monitoring points SP-A and SP-B are utilized as test datasets to input into each model and invert for soil parameters. The resulting soil parameter inversion for each model is detailed in Tables S11 and S12.

5.3 Prediction of settlement

To illustrate the differences in predicted settlement resulting from the inverted soil parameters more intuitively, the predicted soil parameters from each ML model are incorporated into a numerical model. This allows for the derivation of settlement curves at the SP-A and SP-B monitoring points, providing a clear visualization of the impact of the ML-predicted parameters on settlement behavior. Figs. 9 and 10 present a comparative analysis between the settlement curves computed from the inverted soil parameters of each model before and after Bayesian optimization, and the actual measured curves from the SP-A and SP-B datasets. It can be observed that the settlement curves are inconsistent. However, the differences between soil parameters are insignificant, indicating that the inverted soil parameters are the key factors affecting the development of soil settlement. Moreover, the soil parameters inverted by the models with Bayesian optimization exhibit greater accuracy, resulting in curves that closely align with the actual settlement monitoring curves.

Employing the MAE as an evaluation metric, the prediction errors for the settlement curves by each model before and after optimization are quantified, as depicted in Fig. 11. As one can see, the Bayesian optimization yields varying degrees of improvement across the five ML models. Among these, the DT model exhibits the strongest improvement in predictive accuracy; this is because of its simplistic structure, since minor parameter adjustments can significantly influence the model's generalization ability. Concurrently, the DT model's low complexity resulted in persistently subpar accuracy, even after refinement. The 1D-CNN model demonstrates low prediction error and achieves the highest predictive accuracy, following a comprehensive exploration of its potential through Bayesian optimization.

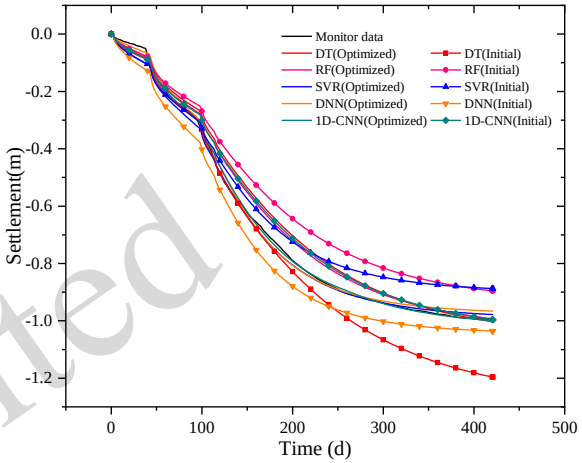

Fig. 9 A comparison between the settlement curves predicted by the five ML models before and after optimization at monitoring point SP-A, and the actual settlement curves

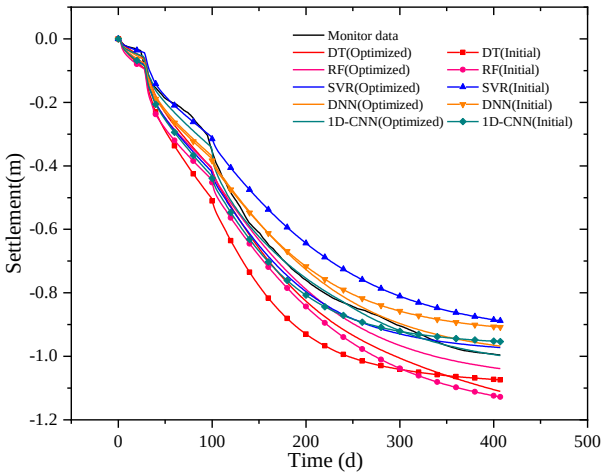

Fig. 10 A comparison between the settlement curves predicted by the five ML models before and after optimization at monitoring point SP-B, and the actual settlement curves

Unedited

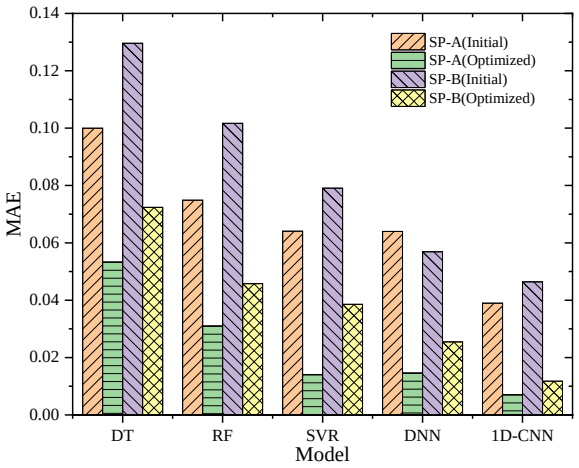

Fig. 11 The MAE of the five ML models before and after Bayesian optimization in predicting settlement curves

Tables S13 and S14 provide the MAE of settlement curves for various ML models before and after optimization on the SP-A and SP-B datasets. On the SP-A and SP-B datasets, the optimized DT model exhibits prediction errors of 0.05329 and 0.07238, respectively, indicating the poorest predictive accuracy. In comparison, the RF model demonstrates significantly lower prediction errors of 0.03101 and 0.04576, showcasing a substantial improvement in predictive accuracy. The SVR model's prediction error on the SP-A dataset is similar to that of the DNN model, but on the SP-B dataset, it exceeds that of the DNN. The optimized 1D-CNN model achieves the lowest MAE values, at 0.00706 and 0.01176, indicating its superior performance in soil parameter inversion and settlement prediction.

To select the most suitable optimized ML model for the inversion of soil parameters and settlement prediction, we employ four statistical evaluation criteria – namely MAPE, MSE, RMSE, and Residual plot – to further compare the settlement curves predicted by the Bayesian-optimized models. In contrast to other indicators, with which we analyze all settlement prediction outcomes, we use MAPE with the settlement forecasts from the final 100 days, aiming to assess the accuracy of different models' predictions in advanced stages of settlement development.

Tables 6 and 7 present the evaluation results using MAPE, MSE and RMSE. In both the SP-A and SP-B datasets, the 1D-CNN delivers the most precise predictions among the five models. While the SVR and DNN exhibit closely aligned results, their predictive accuracy falls short of the 1D-CNN. The MAPE values for settlement predictions over the subsequent 100 days by the SVR model stand at 0.01305 and 0.01577, respectively, which are both lower than the DNN's predictions. However, its MSE and RMSE surpass DNN across the entire dataset, which indicates that although the SVR has superior predictive accuracy in the latter portion of the predictions compared to DNN, it is more prone to significant outliers within the overall prediction curve. The RF model notably enhances prediction accuracy compared to the DT. However, its MSE and RMSE exceed those of SVR and DNN. The DT yields the least favorable results. Therefore, given the scale of the dataset employed in this study, among the five ML models, the 1D-CNN offers the most dependable outcomes for soil parameter inversion and settlement calculations.

Unedited

Table 6 The MAPE, MSE, and RMSE of the five optimized ML models on the SP-A dataset

Evaluation criteria	DT	RF	SVR	DNN	1D-CNN
MAPE	0.1498	0.013119	0.01305	0.02356	0.002607
MSE(m)	0.002824	0.001846	0.000319	0.000291	0.000113
RMSE(m)	0.05314	0.04297	0.01786	0.017059	0.010630

Table 7 The MAPE, MSE, and RMSE of the five optimized ML models on the SP-B dataset

Evaluation criteria	DT	RF	SVR	DNN	1D-CNN
MAPE	0.10433	0.04775	0.01577	0.02506	0.00481
MSE(m)	0.006147	0.002561	0.002102	0.000892	0.000242
RMSE(m)	0.078403	0.050609	0.04585	0.02986	0.015547

Figs. 12 and 13 depict the residuals between the predicted settlement values and the monitored values at SP-A and SP-B for the five optimized models. It is evident that the predictive residuals of the 1D-CNN

gradually diminish with increasing settlement time, trending toward convergence. Comparatively, the predictive efficacy of the DNN and SVR models is surpassed by that of the 1D-CNN. Moreover, the residuals of the DT and RF models are relatively high, indicating noteworthy disparities between the actual settlement curves and the model's projections. In summary, the performance of the 1D-CNN stands out. The 1D-CNN adeptly estimates the constitutive parameters of the soil layers, and accurately forecasts settlement progression, thereby aiding secure operation of the airport.

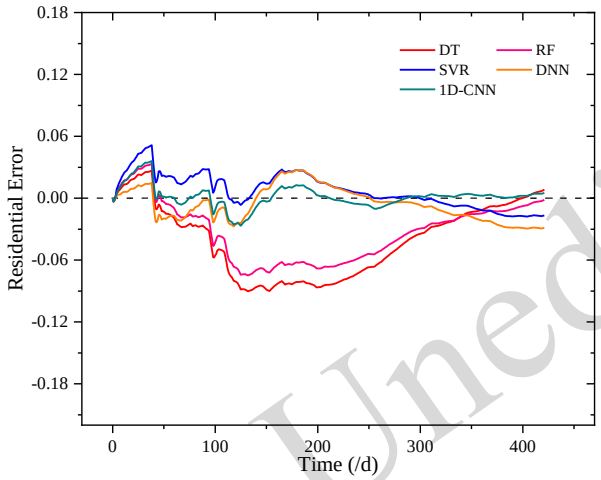

Fig. 12 The residual errors of the settlement prediction curves for the five optimized ML models at monitoring point SP-A

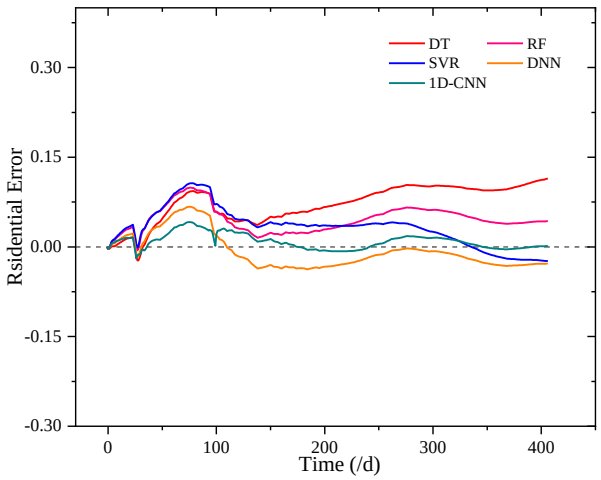

Fig. 13 The residual errors of the settlement prediction curves for the five optimized ML models at monitoring point SP-B

6 Discussion

It is important to make reasonable assumptions when using numerical methods to supplement ML training datasets in geotechnical applications, such as assumptions of subsoil stratification and soil homogenization, as these may introduce certain uncertainties in settlement prediction. Therefore, in this study we obtained prior knowledge from finite element simulations in order to train the ML models. The models' generalization abilities were tested through this process, and the final optimal ML model was determined based on observational data. As a result, systematic errors in the numerical methods were corrected.

Overall, ML models for soil parameter inversion based on Bayesian optimization demonstrate significant potential for airport construction and operation applications. Engineers might first obtain a massive dataset describing the development of runway layer settlement through numerical modelling under various conditions, such as different soil layers and soil consolidation. Subsequently, this dataset can be used to construct a 1D-CNN model, and Bayesian optimization may be employed to quickly obtain the optimal hyperparameter combination for the model. This model can acquire the constitutive parameters of soil layers at any position on the ground surface, and further predict settlement development.

7 Conclusions

In this study we introduce an enhanced ML approach for soil parameter inversion, leveraging Bayesian optimization to determine optimal hyperparameter configurations for ML models. A sensitivity analysis of hyperparameters is conducted to select appropriate search ranges. The TPE Bayesian optimization method is applied to optimize five types of ML models. The accuracy of soil parameter inversion and the impact on settlement prediction are compared using the optimized forms of these models. The models are trained and tested on finite element modeling synthetic datasets, and their generalization ability is validated using real monitoring data from two points on an airport runway. It is found that the soil parameters inverted by the optimized 1D-CNN are the most accurate among the different models. The proposed enhanced ML method can accurately esti-

Unedited

mate soil parameters, providing support for subsequent settlement predictions on runways.

Bayesian optimization is an efficient global optimization algorithm designed to identify optimal hyperparameter combinations. In the initial stages, the objective function decreases rapidly with increasing iterations, demonstrating Bayesian optimization's effectiveness in exploring a limited search space. During the optimization process, the TPE-based strategy not only minimizes the objective function but also explores new potential hyperparameter combinations to avoid local optima, ensuring robust convergence. When evaluating model performance on real-world datasets, significant improvements in both parameter inversion accuracy and settlement curve prediction accuracy were observed following Bayesian optimization. These findings highlight the efficient manner in which Bayesian optimization can enhance model performance, underscoring its utility in complex optimization tasks.

On test datasets, the ML models exhibit significant variability in their predictions of the parameter M_{MCL}, a critical factor influencing the development of airport soil settlement. Among the models, the 1D-CNN demonstrated the best predictive performance, highlighting its ability to capture the key parameters driving settlement behavior. This result underscores the power of the 1D-CNN in modeling complex geotechnical relationships, and its potential for reliable settlement prediction in practical applications.

The complexity of a given model architecture was found to notably influence its performance. When validated against real settlement monitoring data, the DT model exhibits the most substantial improvement following Bayesian optimization. However, due to its inherent simplicity, the overall performance of the DT model remained suboptimal. In contrast, the RF model, which integrates multiple DTs, demonstrates significantly higher prediction accuracy than the standalone DT model. But overall, the 1D-CNN, leveraging its unique convolutional architecture, achieves the best performance across both the training and test datasets, highlighting its ability to capture complex patterns and relationships in the data.

When the inverted soil parameters are input to ABAQUS software for reference settlement development calculations, the optimized 1D-CNN demonstrates the highest accuracy in soil parameter inversion. This enables the model to achieve settlement predictions that closely match the real monitoring sequences. Consequently, the 1D-CNN model, bolstered with Bayesian optimization as proposed in this study, provides a reliable and accurate approach for predicting soil parameters, thereby offering robust support for subsequent runway settlement predictions.

8 Limitations and Future Work

Hyperparameter optimization exhibits a trend of diminishing marginal returns, with the most significant performance improvements typically occurring early on. Considering computational costs and memory consumption, we set the number of trials for the Bayesian hyperparameter optimization to 100 to balance computational efficiency and search thoroughness. In the future, when experimental conditions permit, the number of trials for Bayesian optimization may be gradually increased to search for further improvements to model hyperparameters.

Given the uncertainty and complexity of the dataset, we constructed a 3-layer Deep Neural Network (DNN) to facilitate model convergence and reduce training time and costs. In the future, with sufficient computational resources, the number of hidden layers may be increased or more complex model architectures may be adopted, in order to seek greater accuracy in soil parameter inversion and settlement prediction tasks.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (Nos. 52378419 and 52478368). These supports are gratefully acknowledged.

Author Contributions

An-Feng Hu: Supervision, Writing - Review & Editing, Funding acquisition.

Chi Wang: Conceptualization, Methodology, Formal analysis, Writing - Original Draft.

Sen-Lin Xie : Data curation, Visualization, Writing - Review & Editing.

Zhi-Rong Xiao: Funding acquisition.

Tang Li: Data Curation, Writing - Review & Editing.

Ang Xu: Validation, Visualization

Conflict of interest

The authors declare that they have no known competing fi-

Unedited

nancial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The data generated in this study are available from the corresponding author upon reasonable request.

References

Abbas A, Vantassel JP, Cox BR, et al., 2023. A frequency-velocity cnn for developing near-surface 2d vs images from linear-array, active-source wavefield measurements. *Computers and Geotechnics*, 156:105305. https://doi.org/10.1016/j.compgeo.2023.105305

Asaoka A, 1978. Observational procedure of settlement prediction. *Soils and Foundations*, 18(4):87-101. https://doi.org/10.3208/sandf1972.18.4_87

Asteris PG, Skentou AD, Bardhan A, et al., 2021. Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. *Cement and Concrete Research*, 145 https://doi.org/10.1016/j.cemconres.2021.106449

Asteris PG, Karoglou M, Skentou AD, et al., 2024. Predicting uniaxial compressive strength of rocks using ann models: Incorporating porosity, compressional wave velocity, and schmidt hammer data. *Ultrasonics*, 141 https://doi.org/10.1016/j.ultras.2024.107347

Bardhan A, Ozcan NT, Asteris PG, et al., 2024. Hybrid ensemble paradigms for estimating tunnel boring machine penetration rate for the 10-km long bahce-nurdagi twin tunnels. *Engineering Applications of Artificial Intelligence*, 136 https://doi.org/10.1016/j.engappai.2024.108997

Benzaamia A, Ghrici M, Rebouh R, et al., 2024a. Predicting the compressive strength of cfrp-confined concrete using deep learning. *Engineering Structures*, 319 https://doi.org/10.1016/j.engstruct.2024.118801

Benzaamia A, Ghrici M, Rebouh R, et al., 2024b. Predicting the shear strength of rectangular rc beams strengthened with externally-bonded frp composites using constrained monotonic neural networks. *Engineering Structures*, 313 https://doi.org/10.1016/j.engstruct.2024.118192

Bergstra J, Bengio Y, 2012. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13:281-305.

Breiman L, 2001. Random forests. *Machine Learning*, 45:5-32. https://doi.org/10.1023/A:1010933404324

Chaaban M, Heider Y, Sun WC, et al., 2024. A machine-learning supported multi-scale lbm-tpm model of unsaturated, anisotropic, and deformable porous materials. *International Journal for Numerical and Analytical Methods in Geomechanics*, 48(4):889-910. https://doi.org/10.1002/nag.3668

Chen RP, Zhang P, Kang X, et al., 2019. Prediction of maximum surface settlement caused by earth pressure balance (epb) shield tunneling with ann methods. *Soils and Foundations*, 59(2):284-295. https://doi.org/10.1016/j.sandf.2018.11.005

Cheng Y, Wang JF, He Y, 2023. Prediction models of newmark sliding displacement of slopes using deep neural network and mixed-effect regression. *Computers and Geotechnics*, 156:105264. https://doi.org/10.1016/j.compgeo.2023.105264

Fayed HA, Atiya AF, 2019. Speed up grid-search for parameter selection of support vector machines. *Applied Soft Computing*, 80:202-210. https://doi.org/10.1016/j.asoc.2019.03.037

Glab K, Wehrmeyer G, Thewes M, et al., 2024. Predictive machine learning in earth pressure balanced tunnelling for main drive torque estimation of tunnel boring machines. *Tunnelling and Underground Space Technology*, 146:105642. https://doi.org/10.1016/j.tust.2024.105642

Gu X, Wang L, Ou Q, et al., 2023. Reliability assessment of rainfall-induced slope stability using chebyshev-galerkin-kl expansion and bayesian approach. *Canadian Geotechnical Journal*, 60(12):1909-1922. https://doi.org/10.1139/cgj-2022-0671

Hasanipanah M, Noorian-Bidgoli M, Jahed Armaghani D, et al., 2016. Feasibility of pso-ann model for predicting surface settlement caused by tunneling. *Engineering with Computers*, 32(4):705-715. https://doi.org/10.1007/s00366-016-0447-0

Hu AF, Li T, Chen Y, et al., 2021. Deep learning for preprocessing of measured settlement data. *Journal of Hunan University. Natural Sciences*, 48(9):43-51. https://doi.org/10.16339/j.cnki.hdxbzkb.2021.09.005

Hu AF, Xie SL, Li T, et al., 2023. Soil parameter inversion modeling using deep learning algorithms and its application to settlement prediction: A comparative study. *Acta Geotechnica*, 18(10):5597-5618. https://doi.org/10.1007/s11440-023-01935-z

Huang ZK, Zhang DM, Xie XC, 2022. A practical ann model for predicting the excavation-induced tunnel horizontal displacement in soft soils. *Underground Space*, 7(2):278-293. https://doi.org/10.1016/j.undsp.2021.07.009

Jin YF, Yin ZY, Zhou WH, et al., 2019a. Identifying parameters of advanced soil models using an enhanced transitional markov chain monte carlo method. *Acta Geotechnica*, 14(6):1925-1947. https://doi.org/10.1007/s11440-019-00847-1

Jin YF, Yin ZY, Zhou WH, et al., 2019b. Bayesian model selection for sand with generalization ability evaluation. *International Journal for Numerical and Analytical Methods in Geomechanics*, 43(14):2305-2327. https://doi.org/10.1002/nag.2979

Jones DR, 2001. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21(4):345-383. https://doi.org/10.1023/A:1012771025575

Lester AM, Kouretzis GP, Sloan SW, 2019. Finite element modelling of prefabricated vertical drains using 1d drainage elements with attached smear zones. *Computers*

Unedited

and Geotechnics, 107:235-254. https://doi.org/10.1016/j.compgeo.2018.09.007

Liu MP, Zhuang PZ, Lai FW, 2024. A bayesian optimization-genetic algorithm-based approach for automatic parameter calibration of soil models: Application to clay and sand model. *Computers and Geotechnics*, 176 https://doi.org/10.1016/j.compgeo.2024.106717

Ma ZC, He XH, Yan PC, et al., 2023. A fast and flexible algorithm for microstructure reconstruction combining simulated annealing and deep learning. *Computers and Geotechnics*, 164:105755. https://doi.org/10.1016/j.compgeo.2023.105755

Mahmoodzadeh A, Mohammadi M, Daraei A, et al., 2020. Forecasting maximum surface settlement caused by urban tunneling. *Automation in Construction*, 120:103375. https://doi.org/10.1016/j.autcon.2020.103375

Meng JJ, Mattsson H, Laue J, 2021. Three-dimensional slope stability predictions using artificial neural networks. *International Journal for Numerical and Analytical Methods in Geomechanics*, 45(13):1988-2000. https://doi.org/10.1002/nag.3252

Pan Y, Wu MZ, Zhang LM, et al., 2023. Time series clustering-enabled geological condition perception in tunnel boring machine excavation. *Automation in Construction*, 153:104954. https://doi.org/10.1016/j.autcon.2023.104954

Qi CC, Tang XL, 2018. Slope stability prediction using integrated metaheuristic and machine learning approaches: A comparative study. *Computers & Industrial Engineering*, 118:112-122. https://doi.org/10.1016/j.cie.2018.02.028

Snoek J, Larochelle H, Adams RP, 2012. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25

Sridharan A, Murthy N, Prakash K, 1987. Rectangular hyperbola method of consolidation analysis. *Ge´otechnique*, 37(3):355-368. https://doi.org/10.1680/geot.1987.37.3.355

Tang LB, Na SH, 2021. Comparison of machine learning methods for ground settlement prediction with different tunneling datasets. *Journal of Rock Mechanics and Geotechnical Engineering*, 13(6):1274-1289. https://doi.org/10.1016/j.jrmge.2021.08.006

Victoria AH, Maragatham G, 2021. Automatic tuning of hyperparameters using bayesian optimization. *Evolving Systems*, 12(1):217-223. https://doi.org/10.1007/s12530-020-09345-2

Wang JJ, Vanapalli S, 2024. A framework for estimating the matric suction in unsaturated soils using multiple artificial intelligence techniques. *International Journal for Numerical and Analytical Methods in Geomechanics*, https://doi.org/10.1002/nag.3755

Wang LQ, Wang L, Zhang WG, et al., 2024. Time series prediction of reservoir bank landslide failure probability considering the spatial variability of soil properties. *Journal of Rock Mechanics and Geotechnical Engineering*, https://doi.org/10.1016/j.jrmge.2023.11.040

Wijesinghe DR, Dyson A, You G, et al., 2022. Simultaneous slope design optimisation and stability assessment using a genetic algorithm and a fully automatic image-based analysis. *International Journal for Numerical and Analytical Methods in Geomechanics*, 46(15):2868-2892. https://doi.org/10.1002/nag.3431

Wu SB, Yang ZF, Ding XL, et al., 2020. Two decades of settlement of hong kong international airport measured with multi-temporal insar. *Remote Sensing of Environment*, 248:111976. https://doi.org/10.1016/j.rse.2020.111976

Xie SL, Hu AF, Xiao ZR, et al., 2024. Pinn-based approach to the consolidation analysis of visco-elastic soft soil around twin tunnels. *Tunnelling and Underground Space Technology*, 153 https://doi.org/10.1016/j.tust.2024.105981

Xie SL, Hu AF, Wang MH, et al., 2025. 1dcnn-based prediction methods for subsequent settlement of subgrade with limited monitoring data. *European Journal of Environmental and Civil Engineering*, https://doi.org/10.1080/19648189.2024.2416441

Yao YP, Huang J, Wang ND, et al., 2020. Prediction method of creep settlement considering abrupt factors. *Transportation Geotechnics*, 22:100304. https://doi.org/10.1016/j.trgeo.2019.100304

Ye XW, Jin T, Chen YM, 2022. Machine learning-based forecasting of soil settlement induced by shield tunneling construction. *Tunnelling and Underground Space Technology*, 124:104452. https://doi.org/10.1016/j.tust.2022.104452

Yin ZY, Jin YF, Shen JS, et al., 2018. Optimization techniques for identifying soil parameters in geotechnical engineering: Comparative study and enhancement. *International Journal for Numerical and Analytical Methods in Geomechanics*, 42(1):70-94. https://doi.org/10.1002/nag.2714

Zhang DM, Shen YM, Huang ZK, et al., 2022. Auto machine learning-based modelling and prediction of excavation-induced tunnel displacement. *Journal of Rock Mechanics and Geotechnical Engineering*, 14(4):1100-1114. https://doi.org/10.1016/j.jrmge.2022.03.005

Zhang LM, Wu XG, Ji WY, et al., 2017. Intelligent approach to estimation of tunnel-induced ground settlement using wavelet packet and support vector machines. *Journal of Computing in Civil Engineering*, 31(2):04016053. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000621

Zhang P, Yin ZY, Jin YF, et al., 2021a. Intelligent modelling of clay compressibility using hybrid meta-heuristic and machine learning algorithms. *Geoscience Frontiers*, 12(1):441-452. https://doi.org/10.1016/j.gsf.2020.02.014

Zhang Q, Su Q, Zhang ZY, et al., 2024a. Prediction of high-embankment settlement combining joint denoising

Unedited

technique and enhanced gwo-v-svr method. *Journal of Rock Mechanics and Geotechnical Engineering*, 16(1):317-332. https://doi.org/10.1016/j.jrmge.2023.06.018

Zhang RH, Li YQ, Goh AT, et al., 2021b. Analysis of ground surface settlement in anisotropic clays using extreme gradient boosting and random forest regression models. *Journal of Rock Mechanics and Geotechnical Engineering*, 13(6):1478-1484. https://doi.org/10.1016/j.jrmge.2021.08.001

Zhang WG, Wu CZ, Zhong HY, et al., 2021c. Prediction of undrained shear strength using extreme gradient boosting and random forest based on bayesian optimization. *Geoscience Frontiers*, 12(1):469-477. https://doi.org/10.1016/j.gsf.2020.03.007

Zhang WG, Liu SL, Wang LQ, et al., 2024b. The overall stability of a partially unstable reservoir bank slope to water fluctuation and rainfall based on bayesian theory. *Landslides*, 21(8):2021-2032. https://doi.org/10.1007/s10346-024-02250-8

Zhou Z, Chen Y, Liu ZZ, et al., 2020. Theoretical prediction model for deformations caused by construction of new tunnels undercrossing existing tunnels based on the equivalent layered method. *Computers and Geotechnics*, 123:103565. https://doi.org/10.1016/j.compgeo.2020.103565

Electronic supplementary materials

Eqs. S1-S6, Tables S1-S14, Figs. S1–S6

Unedited

中文概要

题　目：基于贝叶斯优化增强的机器学习方法在土体参数反演中的应用研究

作　者：胡安峰[1]，王池[1,2]，谢森林[1]，肖志荣[3]，李唐[1]，徐昂[1]

机　构：[1]浙江大学，滨海和城市岩土工程研究中心，中国杭州，310058；[2]浙江大学，软弱土与环境土工教育部重点实验室，中国杭州，310058；[3]浙江科技大学，土木与建筑工程学院，中国杭州，310023

目　的：土体参数标定困难是数值模拟方法在软土地基沉降预测中遇到的难题。本研究构建了 5 种机器学习模型，通过贝叶斯优化快速获取模型最优超参数配置，比选最优超参数配置下各模型在测试集和实际监测数据集上的泛化性能，旨在获取高精度、优性能的土体参数反演模型。

创新点：1. 结合数值模拟生成的高保真数据集和厦门翔安机场的实测数据构建数据集，兼顾数据规模和真实性；2. 通过贝叶斯优化，在有限试验次数内快速获取各模型的最优超参数配置。3. 建立 5 种基于沉降数据集的机器学习模型，多组试验比选获取最优超参数配置下用于土体参数反演的最优机器学习模型。

方　法：1. 采用修正剑桥模型，通过调整 FEM 脚本中土体参数的变化范围在 ABAQUS 中生成足量的用于训练机器学习模型的高保真数据集；2. 建立了 5 种基于沉降数据集的土体参数反演机器学习模型，并通过贝叶斯优化获取各模型的最优超参数配置，比较最优超参数配置下各模型在测试集下的预测精度；3. 为进一步验证所构建的机器学习模型在实测数据集上的泛化性能，以厦门翔安机场的实际沉降监测数据为基础，利用优化后的各机器学习模型反演沉降预测中所需的关键土体参数；4. 将反演的土体参数输入数值计算模型，对比沉降计算曲线和实际监测曲线的误差，比选获得最优土体参数反演机器学习模型。

结　论：1.贝叶斯优化是高效的全局优化算法，可快速获取到机器学习模型的最优超参数配置；2. 机器学习模型架构的复杂度会直接影响其土体参数的反演精度，集成了多个决策树的随机森林模型的预测精度较单一决策树模型有了极大提升；3. 一维卷积神经网络模型的预测精度最好，可以可靠、准确地反演土体参数，为沉降的精准预测提供了有力支持。

关键词：ABAQUS；贝叶斯优化；机器学习算法；参数反演；沉降预测