Available online at www.sciencedirect.com

ScienceDirect

Underground Space 24 (2025) 335–351

www.keaipublishing.com/undsp

Research Paper

Interlayer soil settlement prediction in the construction of under-crossing existing structures based on multi-parameter time series model

Boyu Jiang, Haibin Wei *, Dongsheng Wei, Zipeng Ma, Fuyu Wang

School of Transportation, Jilin University, Changchun 130022, China

Received 6 October 2024; received in revised form 6 April 2025; accepted 25 April 2025
Available online 11 August 2025

Abstract

Predicting surface settlement can identify potential risks associated in shield construction. However, in the construction of under-crossing existing structures, the surface settlement is minimal due to the high stiffness of the existing structure, making it unsuitable as a basis for risk assessment. Therefore, interlayer soil settlement was used as an evaluation index in this paper, which was predicted by the developed multi-parameter time series (MPTS) model. This model establishes new dataset, including time, effective stress ratio (ESR), mechanical fluctuation coefficient (MFC), and interlayer soil settlement, where ESR and MFC take into account the changing geological conditions. This study proposes a novel MPTS model, integrating grid search (GS), nonlinear particle swarm optimization (NPSO), and support vector regression (SVR) algorithms to predict interlayer soil settlement during under-crossing construction. It utilizes GS and NPSO to obtain the optimal hyperparameters for SVR. Sensitivity analysis based on MPTS model was used to identify important parameters and propose specific improvement measures. A real under-crossing tunnel project was adopted to verify the effectiveness of the MPTS. The results show that the new input parameters proposed in this paper reduce mean absolute error (MAE) by 20.3% and mean square error (MSE) by 46.7% of prediction results. Compared with the other three algorithms, GS-NPSO-SVR has better prediction performance. Through Sobol sensitivity analysis, previous settlement, ESR and MFC in fully weathered mudstone and moderately weathered mudstone are identified as the primary parameters affecting the interlayer soil settlement. The improvement measures based on analysis results reduce the accumulated settlement by 79.97%. The developed MPTS model can accurately predict the interlayer soil settlement and provide guidance for water stopping or reinforcement construction.

Keywords: Interlayer soil settlement; Under-crossing construction; Multi-parameter time series model; Sensitivity analysis

1 Introduction

As surface traffic congestion increases, the development of underground space has become a crucial solution to alleviate urban traffic pressures (Wu et al., 2023; Xue et al., 2015). Shield tunneling, with its advantages of rapid construction and minimal disruption to surface traffic, has been widely applied in underground tunnel projects (Maeda & Kushiyama, 2005; Ran et al., 2014). However, under-crossing construction induces soil settlement, driven by the interaction between structural loads and soil layers. Soil settlement caused by this construction technique can damage nearby roads and infrastructure. For example, during the construction of Foshan Rail Transit Line 2 in 2018, significant settlement caused road collapse, resulting in direct economic losses of 50 million yuan. Similarly, during the construction of Guangzhou Metro Line 13 in 2022, notable soil settlement led to extensive cracks in the building walls, seriously threatening the safety of the building. In order to reduce the damage or risk for shield

* Corresponding author.
E-mail address: weihb@jlu.edu.cn (H. Wei).
Peer review under the responsibility of Tongji University

https://doi.org/10.1016/j.undsp.2025.04.009

construction, it is crucial to accurately predict the development regularity of settlement through mathematical tools.

Over the past century, numerous approaches have been developed for predicting soil settlement induced by shield tunneling, including theoretical, experimental, and numerical methods. Theoretical method was first proposed by Peck (1969), and then improved by some scholars (Mair et al., 1996; Vorster et al., 2005). While theoretical methods provide computational convenience, they require multiple assumptions, which can lead to considerable differences between the calculated results and the actual settlement. Although experimental methods provide high prediction accuracy (Huang et al., 2021; Zhao et al., 2023), their applicability is limited to areas with similar geological conditions. Numerical methods can simulate complex construction processes (Do et al., 2021; Jallow et al., 2019). However, the results obtained are based on specific physical assumptions and rigid boundary conditions. Therefore, the predictive performance of previous three methods relies heavily on the rationality of the assumptions and accurate acquisition of the geological parameters. The interaction between shield tunnel and soil is complex, leading to poor usability of these methods.

In recent years, machine learning (ML) is widely used in many fields (Ashrafian et al., 2023; Barkhordari et al., 2023; Cavaleri et al., 2016; Ghanizadeh et al., 2023). ML has been applied to predict settlement induced by shield construction due to its powerful data analysis capabilities and multi-factor nonlinear coupling ability. Shen et al., (2023) employed time series clustering algorithms to analyze long-term longitudinal settlement during shield tunnel operation, uncovering settlement patterns across different periods and offering a fresh perspective on the temporal variation of settlement. Wen et al. (2023) utilized convolutional neural networks to forecast surface settlement caused by shield tunnels, highlighting the potential of deep learning for processing complex spatial data. Kovacevic et al. (2021) successfully predicted long-term vertical settlement in soft rock masses by integrating machine learning with theoretical derivation, demonstrating that hybrid models can leverage the strengths of both approaches to enhance prediction accuracy. Samaniego et al. (2020) proposed an energy based thermodynamically and variationally consistent ML approach. This approach integrates modeling and simulation, experimental data and uncertainties within a single framework, offering a novel perspective on applying machine learning to solve engineering problems. In summary, various machine learning methods have been developed to address settlement prediction problems. However, despite significant technical advancements, most studies rely on data from monitored surface structures, which may not be suitable for evaluating the impact of shield tunneling under-crossing existing structures. The reason is that existing structures typically exhibit high stiffness, meaning that even if significant soil deformation occurs, there is only minimal settlement observed in the existing structure. This limitation restricts our understanding of the true impact of shield tunnel construction and can lead to prediction biases. In addition, time series datasets typically include only two parameters: time and settlement (Huang et al., 2024; Yang & Jiang, 2022; Zhai, 2022; Z. Zhang et al., 2022), neglecting the effects of changing geological conditions, which leads to unsatisfactory prediction results. Therefore, future research should place greater emphasis on dataset construction, ensuring that they encompass a broader range of variables to enhance model interpretability and reliability.

In addition to establishing a reasonable dataset, hyperparameters are also important factors affecting the prediction accuracy of ML models. Many optimization methods have been developed to select hyperparameters for ML, such as least squares method (Shim & Hwang, 2010), grid search (Ngoc et al., 2021), and particle swarm optimization (Wang et al., 2023). However, each method has its limitations. Least squares method is sensitive to outliers in the data and has poor generalization capabilities. Grid search is not suitable for multi-parameter optimization due to its high computational cost. Particle swarm optimization converges slowly and is prone to getting stuck in local optima. To overcome these problems, researchers are exploring more efficient optimization strategies, such as Bayesian optimization (Hebbal et al., 2023), group optimization (Wong et al., 2023), and improved firefly algorithm (Bui et al., 2018), aiming to find a method that can guarantee computational efficiency and provide good generalization performance.

The paper is prepared as follows. The methodology is presented in Section 3, which consists of model principles, new dataset, and support vector regression combining grid search and nonlinear particle swarm optimization (GS-NPSO-SVR) fusion algorithm. The developed model is performed on a shield under-crossing project in Section 4. Three aspects are discussed: data acquisition and processing, model building and performance comparison, and model application. The conclusions are drawn in the final section.

2 Research significance

Based on the above discussion, there are three key issues in using ML to predict settlement caused by shield construction. Firstly, traditional predictions rely on surface settlement as a metric, which is inadequate for assessing the construction risk of shield tunneling under-crossing the existing structure. Secondly, conventional time-series settlement predictions use datasets that only include time and settlement data, neglecting changes in soil stress state and mechanical properties, which can significantly affect prediction accuracy. Finally, traditional SVR models select optimal hyperparameters through trial and error, leading to inefficiencies and lower precision. To address these issues, this study carried out the following innovations:

(1) Interlayer soil is located between shield tunnel and existing structure, which exhibits the most direct and obvious changes (He et al., 2023; Zhang et al., 2021). In this study, interlayer soil settlement is used as a predictor instead of surface settlement, which improves the accuracy and reliability in assessing the construction risk of shield tunnel under-crossing the existing structure.
(2) Two new parameters have been developed: effective stress ratio (ESR) and mechanical fluctuation coefficient (MFC). These parameters account for the impact of soil stress conditions and mechanical behavior on settlement, thereby improving prediction accuracy.
(3) A novel GS-NPSO-SVR fusion algorithm combining grid search (GS) with nonlinear particle swarm optimization (NPSO) has been introduced to automatically identify the best SVR hyperparameters. This approach increases computational efficiency and precision.

In addition, this paper also introduced the method to reduce the interlayer soil settlement based on multi-parameter time series (MPTS) model.

3 Methodology

3.1 Model principles

Figure 1 shows the principle and application of the MPTS model. The proposed methodology in this paper consists of two main components: MPTS model building and model application. During the model building stage, a monitoring section is selected where vertical soil pressure (VTP) boxes, pore water pressure (PWP) gauges, horizontal soil pressure (HTP) boxes, moisture content sensors, and settlement meters are installed. After the shield machine passes through this section, sensor data collection begins. The ESR can be obtained by VTP, PWP and HTP, the MFC can be obtained by moisture content, and the interlayer soil settlement can be obtained by settlement meters. A new dataset is developed, including time, ESR, MFC, and interlayer soil settlement. Here, ESR accounts for changes in soil stress state, while MFC reflects variations in soil mechanical properties. A novel GS-NPSO-SVR fusion algorithm is developed to train the dataset, which uses GS and NPSO to find the optimal hyperparameters for SVR, thereby improving computational efficiency

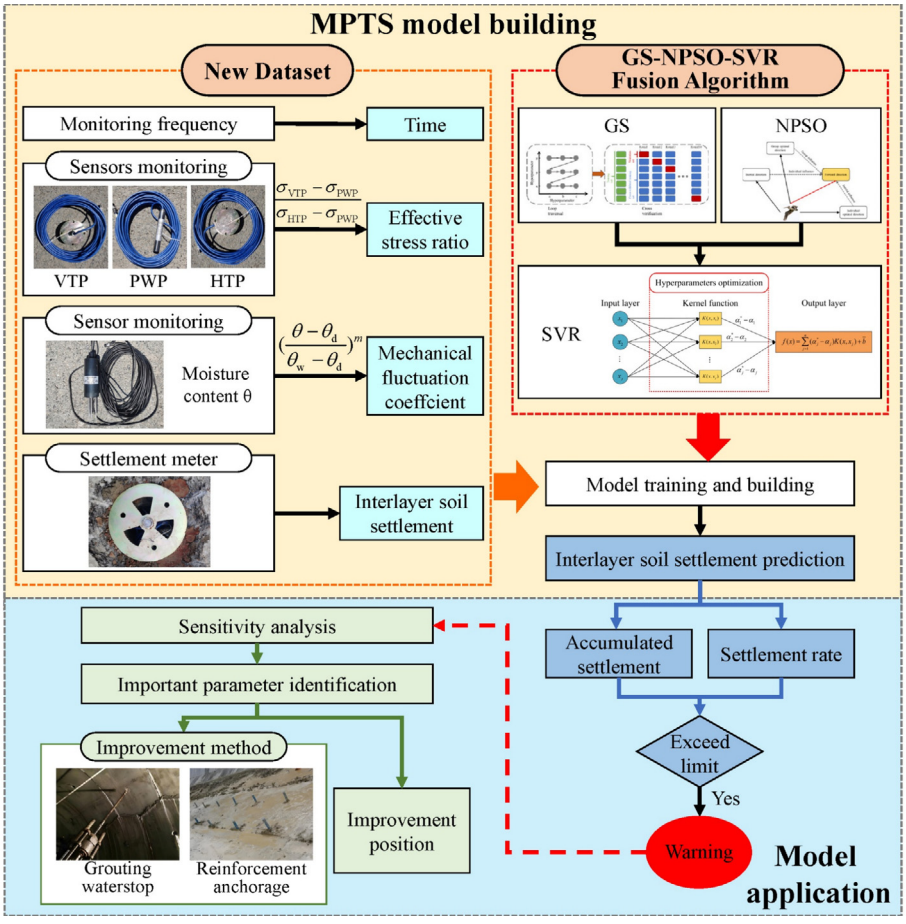


Fig. 1. Model principle and application.

and accuracy. During the model application stage, the trained MPTS model is used to predict interlayer soil settlement. If the accumulated settlement or the settlement rate exceeds the limiting threshold, the model automatically triggers an alert. Sensitivity analysis based on the MPTS model identifies parameters that have a significant correlation with interlayer soil settlement. When the influence coefficient of ESR is high, priority is given to reinforcement measures. For a high influence coefficient of MFC, water control measures are prioritized. Based on the sensitivity analysis results, soil layers with higher impact coefficients are selected for improvement.

3.2 New dataset

A complete and comprehensive dataset is the prerequisite for accurate prediction (Sarir et al., 2021; J. Zhou et al., 2020). Traditional time-series datasets consist of only time and settlement. Predictions based solely on these two parameters are highly uncertain, often leading to significant discrepancies between predicted and actual outcomes. Moreover, models built using just these parameters lack interpretability, making it difficult to identify the underlying causes of settlement development. From a mechanics perspective, the primary factors influencing soil settlement are the stress state and mechanical properties of the soil. For interlayer soils, the complex interactions between the shield machine and the surrounding soil lead to intricate changes in the stress state. Vertically, the soil is subjected to gravity from overlying structures and additional stresses induced by the shield machine's compression. Horizontally, lateral pressures from the surrounding soil act on it. The vertical force represents the maximum principal stress, which tends to disrupt soil structure, while the horizontal force, being the minimum principal stress, tends to preserve the soil structure. These forces are borne jointly by soil particles and pore water. The force carried by the soil particles is referred to as effective stress. The ratio of the maximum effective principal stress to the minimum effective principal stress is defined as the ESR. Numerous studies have demonstrated that ESR is a critical parameter influencing soil deformation (Chen et al., 2020; Da Silva et al., 2021; Riaz et al., 2021). Therefore, this study selects ESR as a model parameter to reflect the stress state of the soil.

In this study, ESR is obtained based on sensors. Soil pressure cells can measure the vertical and horizontal total pressures, while pore water pressure sensors can measure the pore water pressure. The effective stress acting on the soil is equal to the total pressure minus the pore water pressure, as shown in Eq. (1):

$$\begin{cases} \sigma'_{\mathrm{V}} = \sigma_{\mathrm{VTP}} - \sigma_{\mathrm{PWP}} \\ \sigma'_{\mathrm{H}} = \sigma_{\mathrm{HTP}} - \sigma_{\mathrm{PWP}} \end{cases}, \tag{1}$$

where σ'_{V} is the vertical effective stress, σ'_{H} is the horizontal effective stress, σ_{VTP} is the vertical total stress, σ_{HTP} is the horizontal total stress, and σ_{PWP} is the pore water pressure. Therefore, ESR can be expressed as follows:

$$\mathrm{ESR} = \frac{\sigma'_{\mathrm{V}}}{\sigma'_{\mathrm{H}}} = \frac{\sigma_{\mathrm{VTP}} - \sigma_{\mathrm{PWP}}}{\sigma_{\mathrm{HTP}} - \sigma_{\mathrm{PWP}}}. \tag{2}$$

Regarding the mechanical properties of soil, these are typically closely linked to the pore structure and moisture content. For interlayer soils, the pore structure remains largely unchanged after the passage of the shield machine. Therefore, this study focuses exclusively on the influence of moisture content on the mechanical properties of the soil. The mechanical behavior of soil encompasses both elastic and plastic behaviors. Elastic behavior is related to the elastic modulus (C. Zhang et al., 2022; Z. Zhou et al., 2020), while plastic behavior is associated with shear strength (Sharma & Bora, 2003; Spagnoli & Feinendegen, 2017). Moisture content is a crucial factor affecting these two parameters and can be monitored using sensors. Lu and Kaya (2014) proposed a power-law model to describe the elastic modulus and shear strength at different moisture contents, as shown in Eqs. (3) and (4):

$$E = E_{\mathrm{d}} + (E_{\mathrm{w}} - E_{\mathrm{d}})\left(\frac{\theta - \theta_{\mathrm{d}}}{\theta_{\mathrm{w}} - \theta_{\mathrm{d}}}\right)^{m}, \tag{3}$$

where E is the elastic modulus, θ is the moisture content, subscript d is dry state, w is wet state, and m is empirical fitting parameter, for clays and their weathered rocks $m = 0.56$ (Lu & Kaya, 2014).

Similarly, the shear modulus can be expressed as

$$G = G_{\mathrm{d}} + (G_{\mathrm{w}} - G_{\mathrm{d}})\left(\frac{\theta - \theta_{\mathrm{d}}}{\theta_{\mathrm{w}} - \theta_{\mathrm{d}}}\right)^{m}, \tag{4}$$

where G is the shear modulus. A dry state can be set at practically the driest environment of interest, and a wet state can be set at near saturated or saturated conditions. It can be seen from Eqs. (3) and (4) that the elastic modulus and shear strength are mainly affected by the power function $\left(\frac{\theta-\theta_{\mathrm{d}}}{\theta_{\mathrm{w}}-\theta_{\mathrm{d}}}\right)^{m}$, so it is defined as the MFC, which describes the change of mechanical properties.

3.3 GS-NPSO-SVR fusion algorithm

3.3.1 SVR

Vapnik (1964) proposed the method of support vector machine (SVM) in 1964 for linear data classification. In 1992, Boser et al. (1992) utilized kernel functions to develop a nonlinear SVM. SVR extends SVM from classification problems to regression problems by constructing classification hyperplanes or approximation function:

$$f(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\phi(\boldsymbol{x}) + b, \tag{5}$$

where $f(\boldsymbol{x})$ is the objective function, $\boldsymbol{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n]$ is the input data, $\boldsymbol{w}^{\mathrm{T}}$ is a linear weight vector, $\phi(\boldsymbol{x})$ is the mapping function, and b is the intercept.

By introducing slack variables ξ_i and ξ_i^*, the problem of solving for the optimal value can be transformed into the following form of a convex quadratic optimization problem:

$$\min F(\boldsymbol{w}, b, \xi_i, \xi_i^*) = \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*). \tag{6}$$

Among them:

$$\begin{cases} y_i - [\boldsymbol{w}^{\mathrm{T}}\phi(\boldsymbol{x}_i) + b] \le \varepsilon + \xi_i, \\ \boldsymbol{w}^{\mathrm{T}}\phi(\boldsymbol{x}_i) + b - y_i \le \varepsilon + \xi_i^*, \\ \xi_i \ge 0, \xi_i^* \ge 0, \end{cases} \tag{7}$$

where C is the penalty factor, which reflects the punishment degree to the samples whose training error is greater than b. ε is the loss function parameter, which represents the error degree of the regression function.

By introducing the Lagrange factor and taking the partial derivative of the parameter, the following dual form can be obtained:

$$\begin{aligned} \min R(\alpha^*, \alpha) = {} & \tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i^* - \alpha_i)\left(\alpha_j^* - \alpha_j\right)K(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ & +\varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) - \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i), \end{aligned} \tag{8}$$

where R is the new function formed after introducing Lagrange factors, α^* and α are the Lagrange factors, i and j are the sample numbers, and K is the kernel function.

According to Karush–Kuhn–Tucker conditions, the value of b can be obtained as follows:

$$b = y_i - \sum_{j=1}^{n}\left(\alpha_j^* - \alpha_j\right)K(\boldsymbol{x}_i, \boldsymbol{x}_j) \pm \varepsilon. \tag{9}$$

The final prediction model is as follows:

$$f(x) = \sum_{j=1}^{n}\left(\alpha_j^* - \alpha_j\right)K(\boldsymbol{x}_i, \boldsymbol{x}_j) + \bar{b}, \tag{10}$$

where $\bar{b}$ is the average of b.

In this study, SVR was chosen as the machine learning model for several key reasons:

(1) Nonlinear relationships. There exists a clear nonlinear relationship between the input parameters (time, ESR, MFC) and the output parameter (interlayer soil settlement). SVR excels in capturing such complex nonlinear relationships through the use of appropriate kernel functions, making it an ideal choice for our predictive modeling.
(2) Noise handling. The ESR and MFC values are obtained indirectly via sensor monitoring, which introduces potential instrument and reading errors, thereby adding noise to the dataset. SVR employs an insensitive loss function that effectively manages this noise, ensuring that the presence of noise does not unduly influence the performance or accuracy of the model.
(3) Suitability for small to medium datasets. The dataset for predicting interlayer soil settlement consists of several hundred to a few thousand samples, classifying it as a small to medium-sized dataset. SVR's relatively high training complexity makes it especially well-suited for such datasets, where it can provide robust predictions without requiring extensive computational resources.

Therefore, SVR was chosen to guarantee the reliability and accuracy of the prediction model. To develop an SVR model with robust predictive performance, it is essential to identify optimal hyperparameters. While PSO offers high efficiency in hyperparameter optimization, its effectiveness relies heavily on initial parameter selection and may prematurely converge to local optima. Therefore, GS was used to set the initial parameters for PSO. In addition, the parameters of PSO are changed nonlinearly to ensure a strong global search ability in the early iteration and a strong local convergence ability in the late iteration.

3.3.2 GS

Grid search is a method used to find hyperparameters for machine learning models that combines cross-validation and exhaustive search, as shown in Fig. 2. First, define a grid of parameter values that will be explored. Then, specify a set of values for each hyperparameter to be tuned, and use cross-validation to evaluate the model performance for every possible combination of these values. Finally, an exhaustive search within the specified range ensures that the best combination of hyperparameters within the predefined space is found.

3.3.3 NPSO

Particle swarm optimization (PSO) is an optimization technique used to find the best solution within a multidimensional search space. PSO was developed by Kennedy and Eberhart in 1995 (Kennedy & Eberhart, 1995), inspired by the social behavior of bird flocking. In PSO, each potential solution is treated as a particle. These particles fly through the search space, adjusting their positions based on personal best position (pbest) and global best position (gbest). Each particle has a velocity vector that guides its movement. The update rules for the particles' velocities and positions are influenced by both pbest and gbest, along with random factors that encourage exploration of the search space. The equation for updating position and velocity is as follows:

$$V_{t+1} = \omega V_t + c_1 r_1(p_{\mathrm{best}} - \chi_t) + c_2 r_2(g_{\mathrm{best}} - \chi_t), \tag{11}$$

$$\chi_{t+1} = \chi_t + V_{t+1}, \tag{12}$$

where ω is the inertia weight, which is used to reflect the ability to inherit the speed from the previous moment. t is the number of particle swarm iterations, V is the particle swarm velocity, χ is the particle swarm position, p_{best} is the individual extreme, g_{best} is the group extremum, c_1 is the self-learning factor, c_2 is the group learning factor; r_1 and r_2 are [0,1] random numbers.

PSO balances global search and local search by adjusting inertia weight and acceleration factors. An appropriate parameter combination can help the algorithm converge to

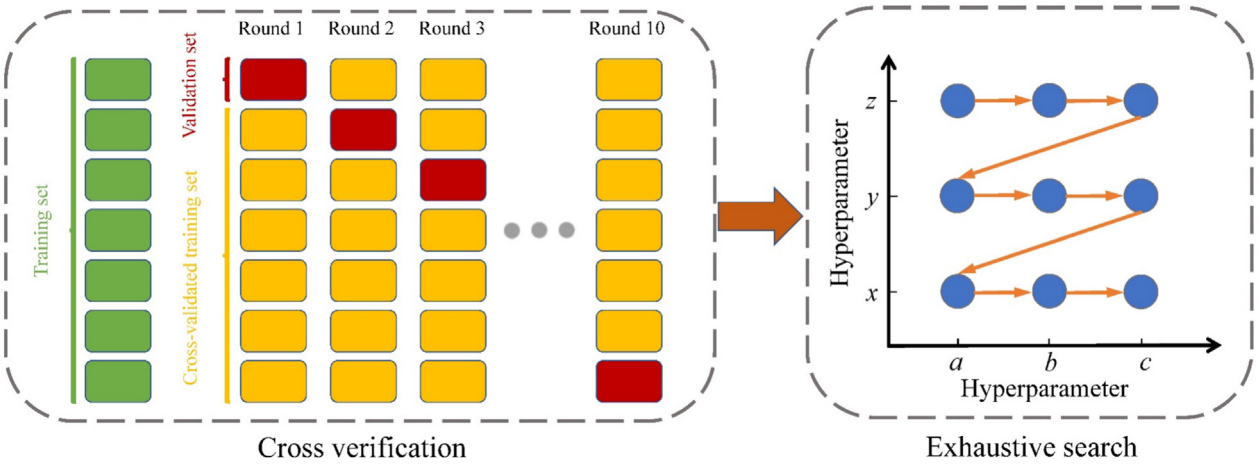


Fig. 2. GSM optimization process.

the global optimal solution more efficiently, while avoiding falling into the local optimal. In general, it is desired that the PSO has a strong search capability in the early iteration and a strong convergence capability in the late iteration. To this end, this paper proposes a nonlinear strategy which makes ω and c_1 large in the early stage and small in the late stage, and c_2 small in the early stage and large in the late stage. The specific improvement formula is as follows:

$$\omega = \lambda \tan[\lambda\pi(1 - t/m)], \tag{13}$$

$$c_1 = c_{1\max} - (c_{1\max} - c_{1\min})(t/m)^{0.2}, \tag{14}$$

$$c_2 = c_{2\min} + (c_{2\max} - c_{2\min})(t/m)^2, \tag{15}$$

where λ is the nonlinear coefficient, t is the number of current iterations, m is the maximum number of iterations, $c_{1\max}$ and $c_{1\min}$ represent the maximum and minimum values of c_1, $c_{2\max}$ and $c_{2\min}$ represent the maximum and minimum values of c_2.

3.3.4 Proposed fusion algorithm

This paper proposes a fusion algorithm, and its development flowchart is shown in Fig. 3. The GS-NPSO-SVR fusion algorithm uses GS and NPSO to find the optimal hyperparameters for SVR. Traditional PSO randomly initializes the positions and velocities of the particle swarm. However, this algorithm first uses GS to determine a preliminary range of optimal hyperparameters. Then, the preliminary optimization results are utilized to initialize the particle swarm's positions and velocities. During the iterations, the inertia weight and acceleration factors are updated nonlinearly. These improvements enhance the algorithm's prediction accuracy, computational efficiency, and stability, while also preventing the model from converging to local optima.

The predictive performance of SVR is primarily influenced by two hyperparameters: the penalty coefficient C and the kernel parameter γ. To identify optimal C and γ, the GS-NPSO hybrid method employs a two-stage process. First, a hyperparameter space is defined to generate all possible combinations. Each combination is evaluated using K-fold cross-validation, and the initial optimal parameter range is identified based on performance metrics. The initial best range is then used to configure particle swarm positions and velocities. Each particle's fitness is evaluated via an objective function. Individual and global best values are recorded, and termination criteria are checked. If unmet, particle velocities and positions are nonlinearly updated using Eqs. (13)–(15). The updated particles undergo renewed fitness evaluations, followed by iterative updates to individual and global best values until termination criteria are satisfied. This iterative process converges to an optimal hyperparameter combination. An appropriate C prevents overfitting or underfitting, while a suitable γ captures both local and global data patterns. Therefore, the prediction performance of SVR is improved by the above method.

To prevent overfitting or underfitting of the model, attention should be paid to the following points:

(1) Selecting an appropriate regularization parameter. Choosing the right regularization parameter is critical for balancing model complexity. A higher value can lead to overfitting by making the model too sensitive to the training data, while a lower value can result in underfitting by producing a model that is too generalized. Cross-validation techniques should be employed to find an optimal balance.
(2) Choosing an appropriate epsilon parameter: The epsilon parameter plays a key role in determining how tolerant the model is to prediction errors. Too small an epsilon can cause the model to overfit by focusing excessively on minor variations in the training data. Conversely, too large an epsilon can lead to underfitting by ignoring important details. Utilizing grid search or visualization tools can help identify the most suitable epsilon value by observing how well the model generalizes across different error tolerances.

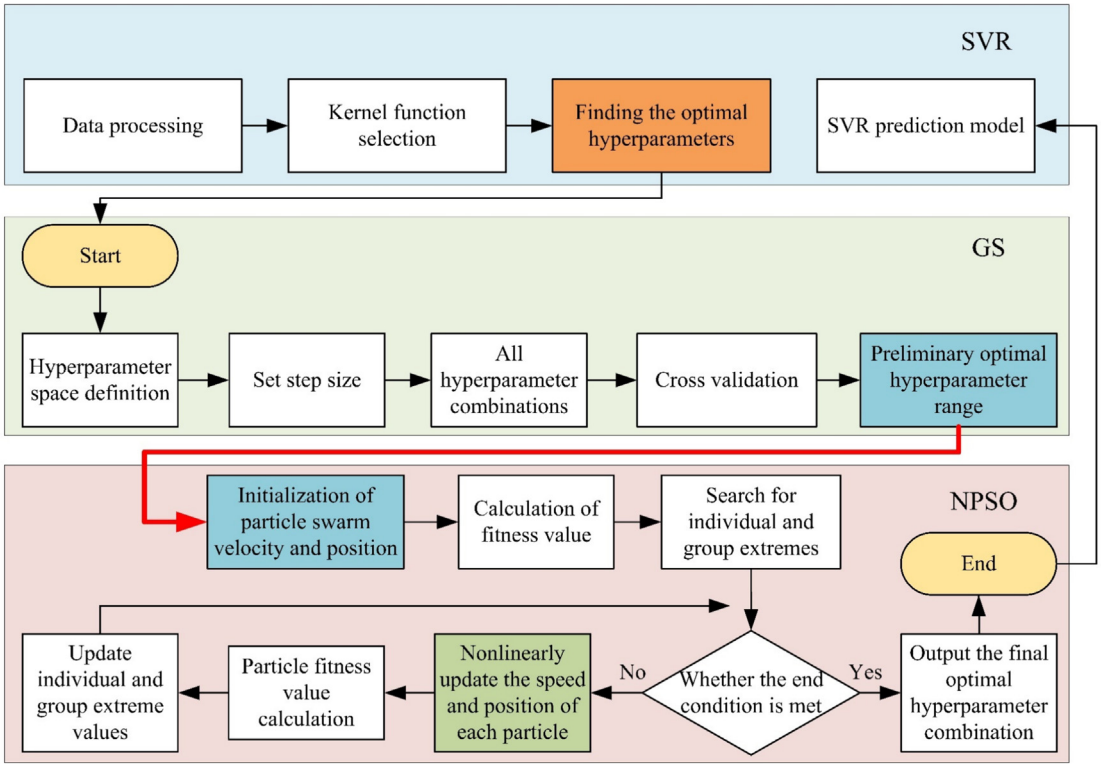


Fig. 3. Flowchart of GS-NPSO-SVR fusion algorithm.

(3) Constructing a reasonable dataset. Ensuring a robust dataset is essential for training a reliable model. A larger dataset allows the model to capture true underlying patterns rather than random noise, reducing the risk of overfitting. Additionally, normalizing all input data ensures that features with different scales do not disproportionately influence the model's predictions, helping to prevent underfitting.

3.4 Evaluation indicators

Five evaluation metrics are used to evaluate the predictive performance of the algorithms, including coefficient of determination (R^2), mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE), and Theil's inequality coefficient (TIC). R^2 is a metric used to evaluate the goodness of fit, with values closer to one indicating better performance. The other four metrics describe prediction errors from different perspectives, with smaller values indicating better performance. The formulas for these metrics are as follows:

$$R^2 = \frac{\left(n\sum_{i=1}^{n}\widehat{y}_i y_i - \sum_{i=1}^{n}\widehat{y}_i\sum_{i=1}^{n}y_i\right)^2}{\left[n\sum_{i=1}^{n}\widehat{y}_i^{\,2} - \left(\sum_{i=1}^{n}\widehat{y}_i\right)^2\right]\left[n\sum_{i=1}^{n}y_i^2 - \left(\sum_{i=1}^{n}y_i\right)^2\right]}, \tag{16}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\widehat{y}_i - y_i|, \tag{17}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2, \tag{18}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left(\left|\frac{y_i - \widehat{y}_i}{y_i}\right|\right) \times 100\%, \tag{19}$$

$$\mathrm{TIC} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{y}_i)^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i)^2}}, \tag{20}$$

where n is the number of samples in the data, $\widehat{y}_i$ is the predicted value, and y_i is the true value.

4 Case study

4.1 Project overview

In this study, the MPTS model was used to predict the interlayer soil settlement in an under-crossing existing tunnel project. The project is located from Feiyue Square Station to Eurasia Shopping Mall Station in Changchun Metro Line 6. It uses earth pressure balance shield machine. The test section is located at mileage 17 315. The existing tunnel is 10 m high and 28.4 m wide, which is under-crossed by a shield tunnel. The shield tunnel is composed of segments with an outer diameter of 6.2 m and an inner diameter of 5.5 m. The interlayer soil settlement is located between the existing tunnel and the shield

tunnel, which consists of silty clay (SC), fully weathered mudstone (FWM) and moderately weathered mudstone (MWM). The physical parameters of these soils are shown in Table 1. It has been shown that the soil on the shield axis exhibits the most obvious changes (Deng, 2011; Feng et al., 2022), so it is used as the prediction target. To obtain settlement data, the settlement meter was buried in the interlayer soil through a borehole. To determine ESR and MFC, moisture content sensors, vertical and horizontal soil pressure cells, and pore water pressure gauges were embedded in each soil, as shown in Fig. 4. After the installation, the original soil was backfilled and compacted.

In this paper, moisture content sensors were produced by Beijing Spark Chuangyi Electronic Technology Co., Ltd., with a range of 0–100% and a measurement accuracy of 2%. The working temperature is −30 to 70 °C, and the response time is 1 s. The soil pressure cells and pore water pressure gauges were produced by Changsha Xianghao Electronic Technology Co., Ltd., with a range of 0.6 MPa and a measurement accuracy of 0.001 MPa. The working temperature is −20 to 70 °C, and the response time is 3 s. They can accurately collect moisture content and stress states of interlayer soil.

4.2 Data acquisition and processing

The interlayer soil on the test section was monitored once a day for 103 days. The first 82 days were used as the training set, and the last 21 days as the testing set. The monitoring results of these sensors are shown in Fig. 5. According to Eqs. (2)–(4), the ESR and MFC were calculated based on the monitoring data.

When addressing missing data, the first step involves categorizing the missingness type as either random or non-random. If non-random missingness is identified, potential equipment or sensor malfunctions should be investigated, with damaged components requiring replacement. For random missingness, short-term gaps can be filled using linear interpolation, while long-term gaps require prediction through machine learning models. In addition, the plausibility of filled values should be verified through cross-validation or physical constraint checks. In handling outliers, a sliding window approach is first applied to detect anomalies by calculating the window-specific mean and standard deviation, marking values exceeding $\mu \pm 3\sigma$ standard deviations as outliers, where μ is the sample mean and σ is the standard deviation. Isolated

Table 1
Physical properties.

Name	Water content ω_w (%)	Natural unit Weight γ_s (kN/m^3)	Void ratio e	Liquid limit W_L (%)	Plastic limit W_P (%)	Poisson's ratio λ
SC	19.9	19.0	0.79	28.5	16.5	0.30
FWM	18.3	19.4	0.72	28.8	18.1	0.28
MWM	15.7	19.9	0.62	27.5	17.0	0.28

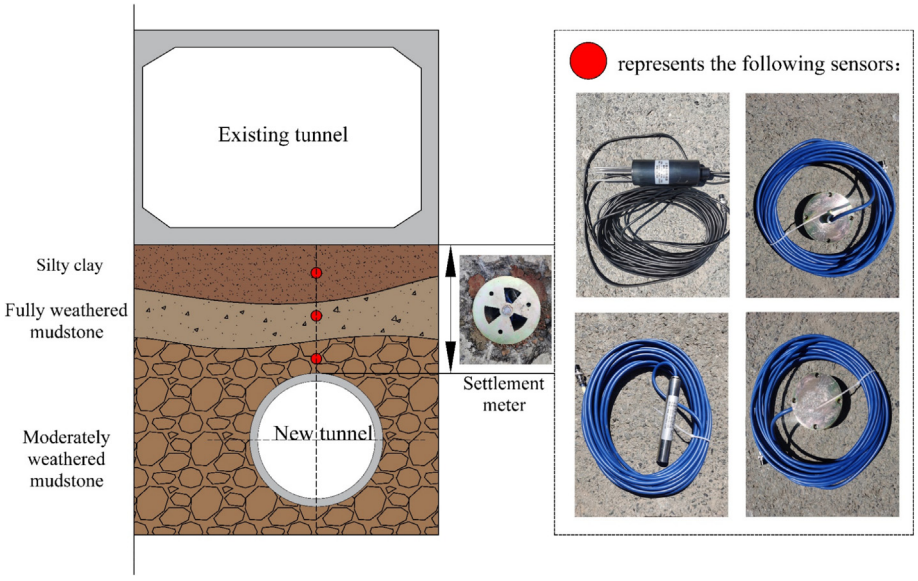


Fig. 4. Geological conditions and sensors.

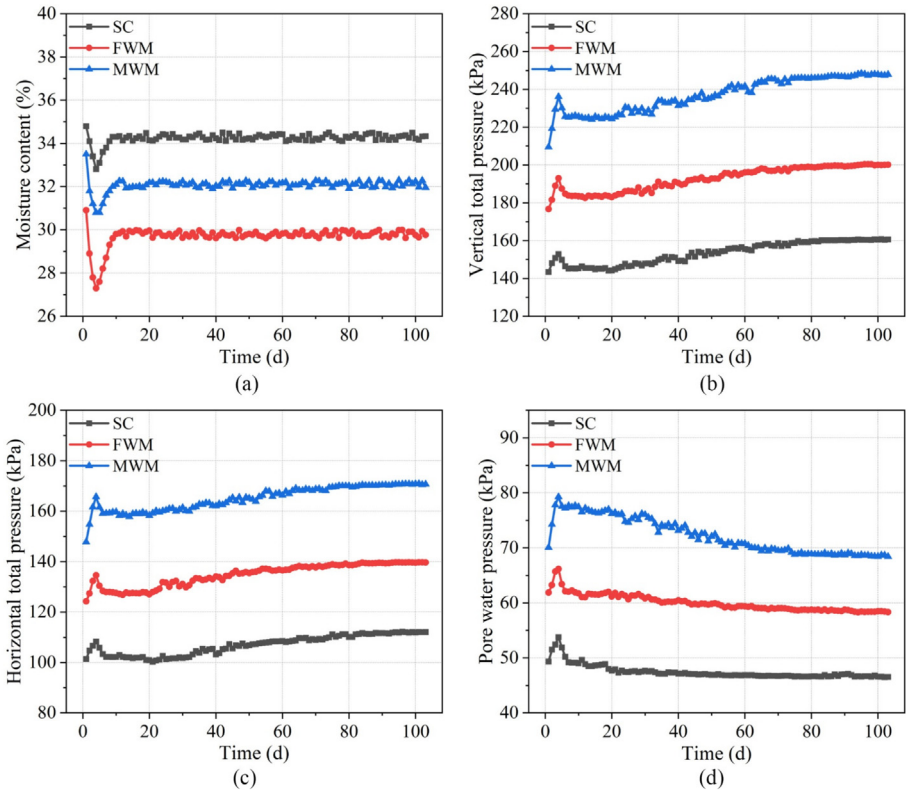


Fig. 5. Sensor monitoring results. (a) Moisture content monitoring results, (b) VTP monitoring results, (c) HTP monitoring results, and (d) PWP monitoring results.

outliers can be replaced with the mean of adjacent data points, whereas consecutive anomalies require prediction and substitution using machine learning models. The corrected outliers also need to be verified for rationality.

To improve the performance of the model, the whole dataset should be normalized to make the value of each parameter map into the interval [−1,1]. The specific calculation equation is presented:

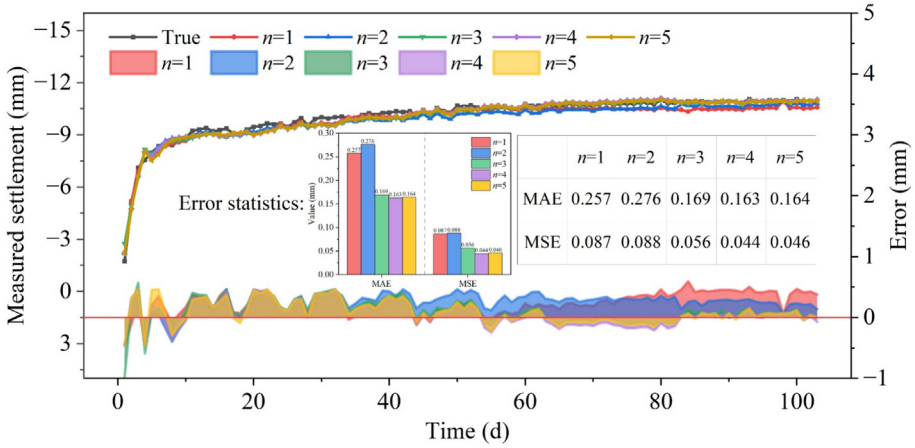
$$x' = 2 \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1, \tag{21}$$

where x' is the normalized data, x is the initial sample data, $x_{\min}$ is the minimum value of the corresponding parameter, and $x_{\max}$ is the maximum value of the corresponding parameter. A portion of the processed data is shown in Table 2.

Another key parameter to be determined is the time step. Time step indicates how many previous days of data should be used for predicting, which is closely related to the accuracy of the results. In this study, the prediction performance of the model was tested when the time step $n = 1$, 2, 3, 4, 5, as shown in Fig. 6. It can be seen that the model's prediction performance is poor when $n = 1$, 2, but improves when $n = 3$, 4, 5. The error of the model is minimal when $n = 4$. Therefore, this study determined that the time step is 4.

Table 2
Original data.

Time (d)	ESR			MFC			Settlement (mm)
	SC	FWM	MWM	SC	FWM	MWM	
1	0.204	0.766	0.350	0.983	0.994	0.995	1.000
2	0.277	0.837	0.496	0.297	−0.093	−0.251	0.285
3	0.329	0.933	0.656	−0.401	−0.718	−0.702	−0.154
4	0.348	0.996	0.773	−1.009	−1.009	−1.007	−0.250
5	0.317	0.947	0.809	−0.704	−0.834	−1.007	−0.341
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
99	−0.822	−0.879	−0.390	0.680	0.462	−0.004	−0.971
100	−0.858	−0.794	−0.489	0.416	0.377	0.078	−1.000
101	−0.851	−0.878	−0.551	0.361	0.324	−0.121	−0.972
102	−0.950	−0.835	−0.448	0.507	0.479	0.111	−0.979
103	−0.907	−0.836	−0.439	0.524	0.383	−0.136	−0.979



	n=1	n=2	n=3	n=4	n=5
MAE	0.257	0.276	0.169	0.163	0.164
MSE	0.087	0.088	0.056	0.044	0.046

Fig. 6. Determination of time step.

4.3 Model building and performance comparison

4.3.1 Hyperparameter optimization and model building

In SVR model, the RBF kernel is one of the most widely used nonlinear functions (Pang et al., 2011; Shariati et al., 2022). For RBF kernel functions, the kernel hyperparameter γ plays a critical role in the distribution of data mapped to a high-dimensional space, and its expression is as follows:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\gamma|\boldsymbol{x}_i - \boldsymbol{x}_j|^2\right). \tag{22}$$

In addition, the penalty factor C plays a crucial role in balancing the model's generalization ability and complexity. In this paper, GS-NPSO was employed to optimize hyperparameters C and γ. First, GS was utilized to narrow down the range of hyperparameters and provide initial conditions for NPSO. The hyperparameter space was set as $[2^{-10}, 2^{10}]$ with a step size of $2^{0.5}$, and all possible combinations were subjected to cross-validation. The results are shown in Fig. 7. MSE was used as an indicator to evaluate prediction accuracy. It can be seen from Fig. 7(c) and (d) that the preliminary optimal ranges of C and γ were $[2^3,2^6]$ and $[2^{-10},2^{-7}]$, respectively.

Then, the initial position of the particle swarm was set in this range and the initial velocity was 10% of the optimal range. The particle swarm size was 30, and the particle dimension was 2. The maximum number of iterations was 400, and the loss function was 0.01. The nonlinear coefficient $\lambda = 0.45$ and the acceleration factors $c_{1\max} = c_{2\max} = 1.5$, $c_{1\min} = c_{2\min} = 0.1$. All the parameters required for the model are shown in Table 3.

The nonlinear updates of inertia weight and acceleration factors are shown in Fig. 8(a). To evaluate the hyperparameter optimization capability of GS-NPSO, the standard PSO ($\omega = 0.9, c_1 = c_2 = 1.5$) was used for comparison as shown in Fig. 8(b). The MSE decreases with the number of iterations. The MSE of PSO eventually stabilized at 0.033 46 after 201 iterations, while the MSE of GS-NPSO stabilized at 0.027 49 after only 70 iterations. Compared with PSO, GS-NPSO improves prediction accuracy by 17.8% and prediction speed by 65.2%. The results show that GS-NPSO has better prediction performance. The hyperparameter optimization results ($c = 985.0684, \gamma = 0.000\,977$) through GS-NPSO were used to build SVR model.

4.3.2 Effects of ESR and MFC on prediction accuracy

ESR and MFC are the new parameters proposed in this paper to represent changing geological conditions. In order to show their contribution to improving prediction accuracy, the prediction results using the new input parameters were compared with those using only time and interlayer soil settlement as inputs, as shown in Fig. 9. It can be seen that the prediction results with ESR and MFC are closer to the true value. The maximum error between the predicted and measured settlements was 0.466 mm, with an average error of 0.181 mm. Compared with the prediction results without ESR and MFC, the MAE and MSE are reduced by 20.3% and 46.7%, respectively. The results confirm the effectiveness of the new input parameters in enhancing prediction accuracy.

4.3.3 Fusion algorithm performance comparison

In order to further evaluate the performance of GS-NPSO-SVR fusion algorithm, grid search support vector regression (GS-SVR), particle swarm optimization support vector regression (PSO-SVR) and back propagation neural network (BPNN) were trained on the same data set. The R^2, MAE, MSE, MAPE and TIC of each algorithm were calculated, as shown in the Fig. 10. The scatter plot on the left side shows the prediction results for both training and testing sets, and the radar plot on the right side presents the prediction performance of the algorithms. The deviation of all algorithms in the prediction results is within 20%, indicating that all four machine learning models have acceptable prediction accuracy (Asteris et al., 2024a, 2024b). The evaluation metrics of GS-NPSO-SVR algorithm are as follows: $R^2 = 0.990$, MAE = 0.098, MSE = 0.022, MAPE = 1.125, and TIC = 0.008 in the training set, and $R^2 = 0.875$, MAE = 0.038,

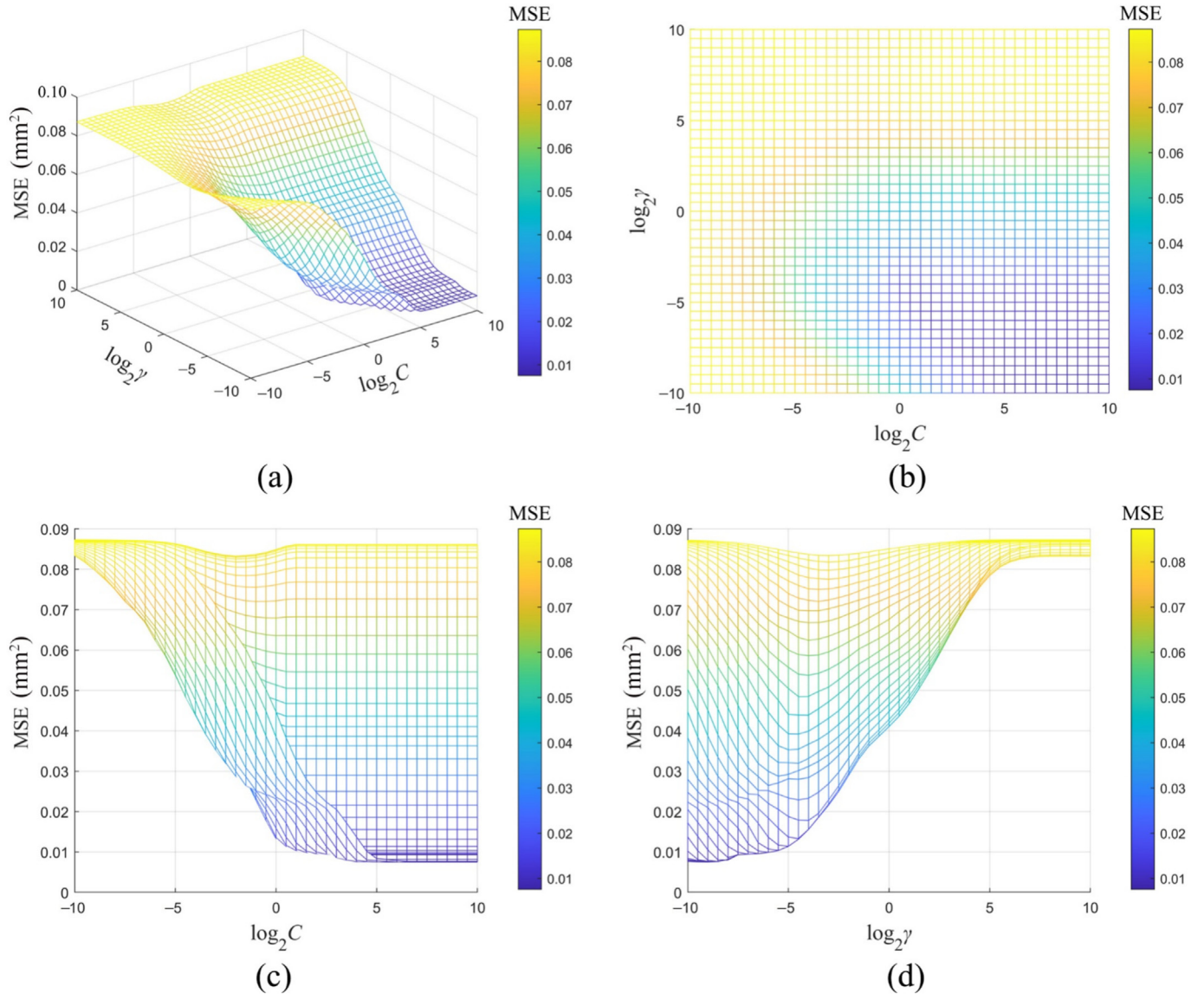


Fig. 7. MSE of C and γ. (a) Azimuth −37.5°, vertical elevation 30°, (b) azimuth 0°, vertical elevation 90°, (c) azimuth 0°, vertical elevation 0°, and (d) azimuth 90°, vertical elevation 0°.

Table 3
Parameter value.

Parameters	C	γ	Particle swarm size	Particle dimension	Maximum number of iterations	Loss function
Value	985.0684	0.000 977	30	2	400	0.01

MSE = 0.002, MAPE = 0.343, and TIC = 0.002 in the testing set. In both the training and testing sets, the R^2 of GS-NPSO-SVR is higher than that of other algorithms, while the four error-related metrics (MAE, MSE, MAPE, and TIC) are lower. It means that GS-NPSO-SVR performs better in target parameter estimation. In addition, the R^2 on the testing set is lower than that on the training set, but it performs better on other parameters. This result is expected since there is a rapid development period of settlement in the training set, leading to a large error.

4.4 Model application

The MPTS model was used to predict the accumulated settlement and settlement rate, as shown in Fig. 11. This project requires that the accumulated settlement to be less than 5 mm and the settlement rate to be less than 1 mm/d. It can be seen that normal construction cannot meet the settlement requirements. Therefore, this model made a warning. Then, the sensitivity analysis based on MPIS model was used to determine the reinforcement measures.

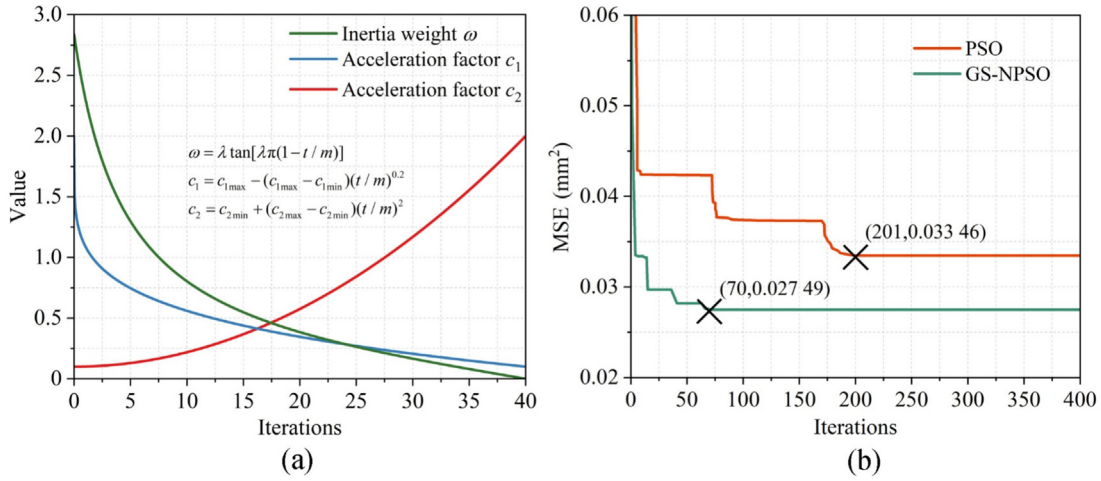


Fig. 8. Acceleration effect. (a) Nonlinear updates of inertia weight and acceleration factors, and (b) optimization comparison of PSO and GS-NPSO.

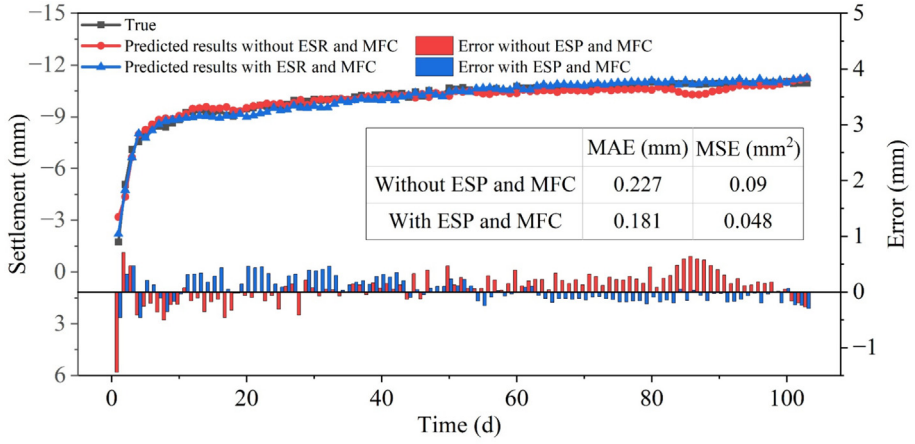


Fig. 9. Effects of ESR and MFC on prediction accuracy.

Sobol method a commonly used sensitivity analysis method, which evaluates the importance of input parameters to predicted outcomes based on variance decomposition. Its equation is as follows:

$$A = \int f^2(\boldsymbol{X})\mathrm{d}\boldsymbol{X} - f_0^2, \tag{23}$$

$$A_i = \int f_i^2(\boldsymbol{x}_i)\mathrm{d}\boldsymbol{x}_i, \tag{24}$$

$$A_{i_1,i_2,\cdots i_s} = \int\int\cdots\int f_{i_1,i_2,\cdots,i_s}^2(\boldsymbol{x}_{i_1},\boldsymbol{x}_{i_2},\cdots,\boldsymbol{x}_{i_s})\mathrm{d}\boldsymbol{x}_{i_1}\mathrm{d}\boldsymbol{x}_{i_2}\cdots\mathrm{d}\boldsymbol{x}_{i_s}, \tag{25}$$

where A is the total variance, which represents the influence of all input parameters on the model output, A_i is the partial variance, which represents the influence of an individual input parameter on the model output, and $A_{i_1,i_2,\cdots i_s}$ is the partial variance, which represents the influence of the interaction between multiple input parameters on the model output.

The sensitivity coefficient is the ratio of partial variance to total variance, as follows:

$$S_{i_1,i_2,\cdots,i_s} = \frac{A_{i_1,i_2,\cdots,i_s}}{A}, \tag{26}$$

where S_i is the first-order sensitivity coefficient, indicating the influence of $\boldsymbol{X}_i$ on the model output. To obtain the impact of all parameters related to $\boldsymbol{X}_i$, the corresponding opposite event should be considered: the influence of all other factors except $\boldsymbol{X}_i$ on the model output. Then, the global sensitivity coefficient can be expressed as follows:

$$S_i^T = 1 - \frac{A_{-i}}{A}, \tag{27}$$

where S_i^T is the global sensitivity coefficient of $\boldsymbol{X}_i$.

In this study, Sobol method was employed analyze the correlation between the input parameters and the interlayer soil settlement. The global sensitivity is shown in Fig. 12(a), and the first-order sensitivity is shown in Fig. 12(b). Among all the input parameters, previous settlement (PS) exhibits the highest global and first-order sensitivity. It

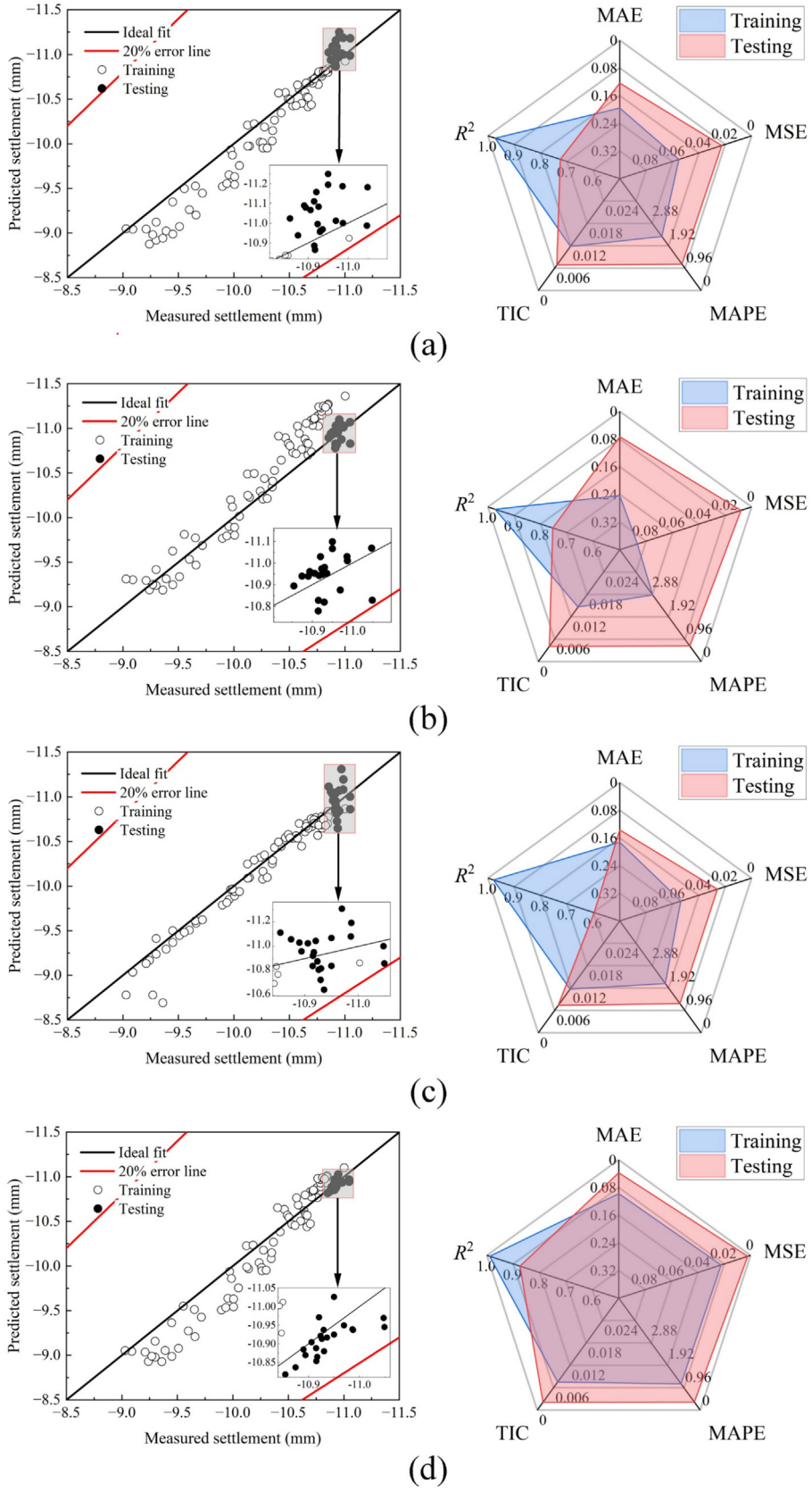


Fig. 10. Prediction results and model performance. (a) GS-SVR, (b) PSO-SVR, (c) BPNN, and (d) GS-NPSO-SVR.

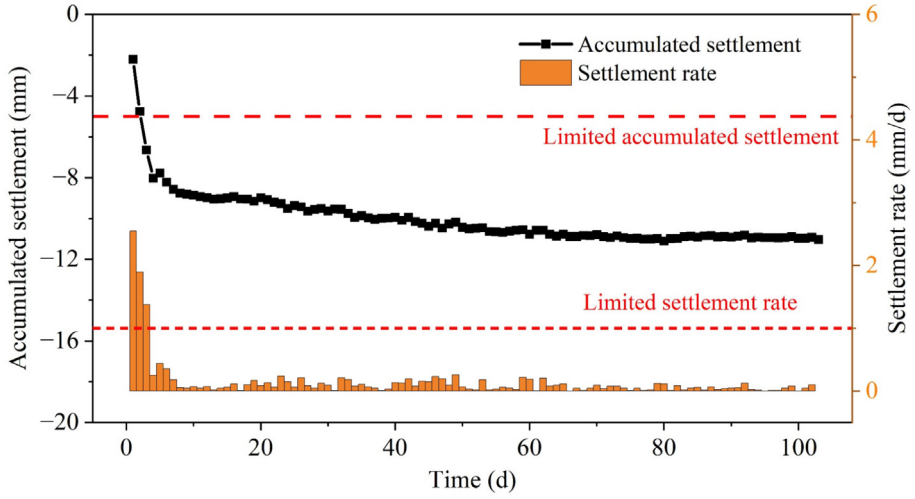


Fig. 11. Accumulated settlement and settlement rate.

means that regular monitoring remains an effective measure for evaluating settlement development. As for the ESR and MFC developed in this paper, they have large influence coefficients in the FWM and MWM and small influence coefficients in the SC. The reason is that FWM and MWM are significantly influenced by construction activities, leading to notable changes in stress and moisture content. Conversely, silty clay exhibits little variation due to its long distance from the excavation face.

According to the sensitivity analysis results, implementing reinforcement or water-stopping measures for FWM and MWM can help reduce interlayer soil settlement. Therefore, FWM and MWM on another section with similar geological conditions were improved by grouting. The MPTS model was used to predict the interlayer soil settlement, which was compared with the unreinforced section. The 500-day prediction results are shown in Fig. 13. It can be seen that the settlement with reinforcement is significantly smaller than that without reinforcement. The maximum accumulated settlement after reinforcement is 2.52 mm, which is 79.97% less than that without reinforcement, and the maximum settlement rate is 0.94 mm/d, both of which are less than the limit value, indicating the effectiveness of the MPTS model-based reinforcement measures.

To further investigate the time-varying characteristics of interlayer soil settlement, this study compares the settlement behavior with and without improvement based on machine learning predictions. Using the settlement at day 500 as the reference total settlement, Fig. 14(a) and (b) illustrates the daily and cumulative settlement proportions for the first 10 days and from day 10 to day 100, respectively. It can be seen from Fig. 14 that the settlement mainly occurred within a few days after the passage of the shield machine, and the daily settlement ratio gradually decreased with time. For reinforced soil, the cumulative settlement quickly approaches nearly 100% within the first 10 days. In contrast, for unreinforced soil, the cumulative settlement progresses more slowly, reaching 71.3% of the total settlement by day 10. As shown in Fig. 14(b), the daily settlement ratio significantly decreases compared to the first 10 days. The cumulative settlement for reinforced soil has stabilized, whereas for unreinforced soil, it continues to increase gradually. By day 100, the cumulative settlement for unreinforced soil amounts to 88.6% of the total settlement.

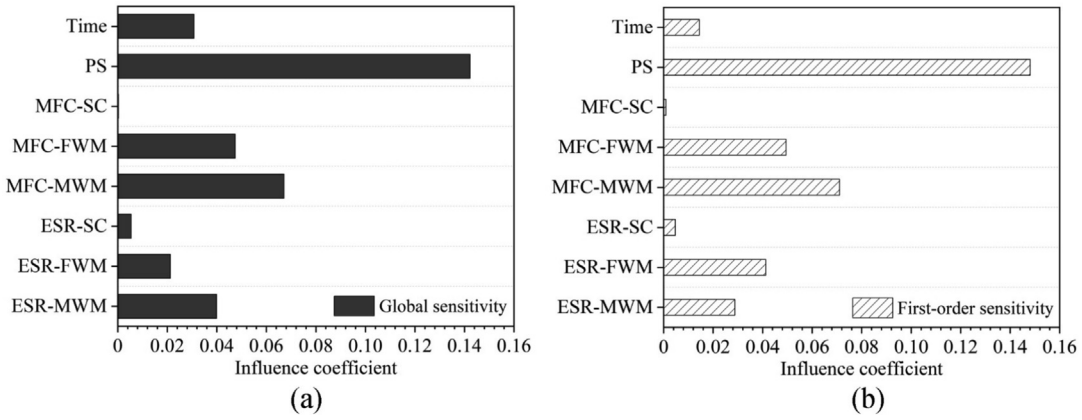


Fig. 12. Sensitivity analysis results. (a) Global sensitivity, and (b) first-order sensitivity.

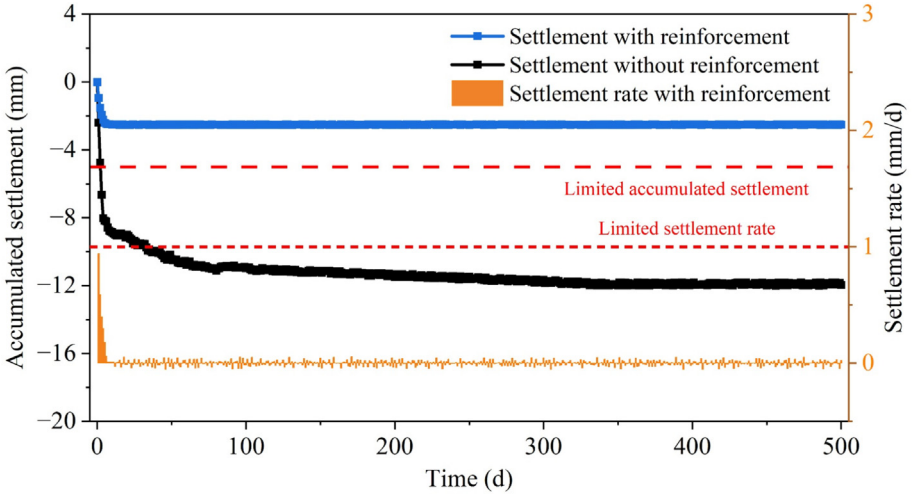


Fig. 13. Reinforcement effect based on MPTS model.

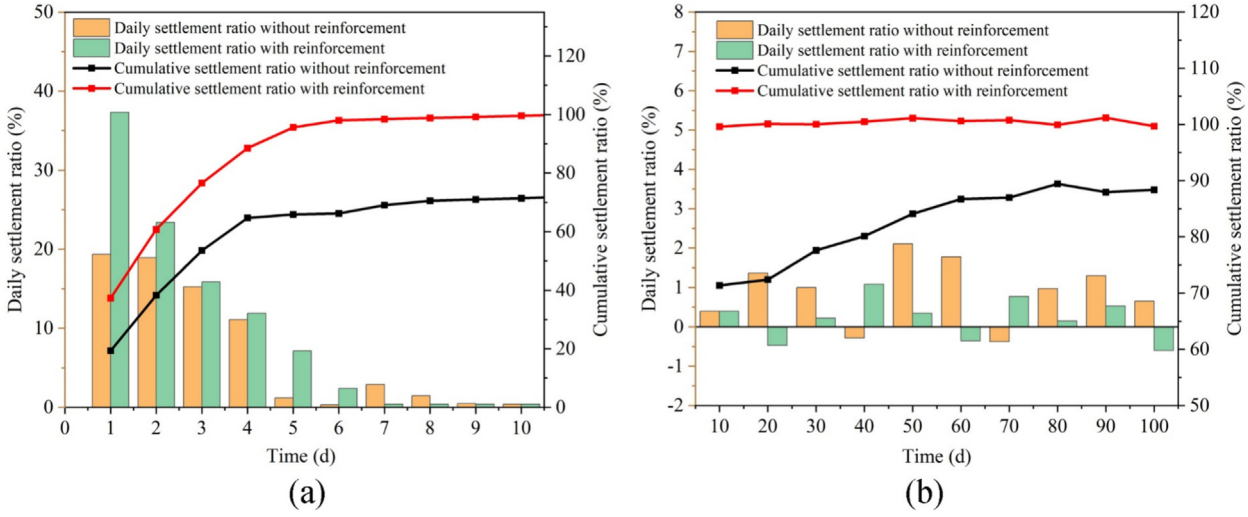


Fig. 14. Daily and cumulative settlement. (a) Settlement ratio of 1–10 days, and (b) settlement ratio of 10–100 days.

According to the information presented in Fig. 14, the settlement of reinforced soil can be divided into two stages: a rapid accumulation stage followed by a stabilization stage. During the rapid accumulation stage, settlement increases quickly before settling into a stable pattern. In contrast, the settlement of unreinforced soil experiences three distinct stage s: an initial rapid accumulation stage, a subsequent gradual development stage, and finally a stabilization stage. Settlement increases swiftly during the rapid phase, slows down during the gradual stage, and then stabilizes. The phenomenon can be attributed to the fact that the settlement caused by shield construction consists of two components: one is short-term settlement due to the gap left by the tail of the shield machine, which occurs during the rapid accumulation stage. The other is long-term settlement resulting from the sustained dissipation of pore water pressure, which takes place during the gradual development stage. Reinforced soil, having its pores filled with grout that solidifies, does not experience seepage consolidation. Therefore, it does not have a gradual development stage. In contrast, for unreinforced soil, approximately 70% of the settlement originates from the gap left by the shield machine's tail, while about 30% results from seepage consolidation.

5 Conclusions

This paper proposed a multi-parameter time series (MPTS) model to predict the interlayer soil settlement in the construction of existing structures under-crossed by a shield tunnel. The MPTS model developed new input parameters and a fusion algorithm. Sobol sensitivity analysis was used to identify important parameters and propose improvement measures. The model can be used to analyze potential risks and guide water stopping or reinforcement construction. The following conclusions can be obtained:

(1) The new input parameters include time, ESR, MFC and interlayer soil settlement. Among them, ESR and MFC take into account changing geological con-

ditions. Compared with the conventional input parameters, the introduction of ESR and MFC reduced the MAE and MSE by 20.3% and 46.7%, respectively.

(2) The novel GS-NPSO-SVR fusion algorithm combines grid search and nonlinear particle swarm optimization to find the optimal hyperparameters for SVR. Five performance indicators were applied to evaluate the predictive ability of GS-SVR, PSO-SVR, BPNN, and GS-NPSO-SVR. The GS-NPSO-SVR exhibited the best performance, with $R^2 = 0.990$, MAE = 0.098, MSE = 0.022, MAPE = 1.125, and TIC = 0.008 in the training set and $R^2 = 0.875$, MAE = 0.038, MSE = 0.002, MAPE = 0.343, and TIC = 0.002 in the testing set. The GS-NPSO-SVR model demonstrated high prediction accuracy and robustness in complex undercrossing conditions.

(3) Sobol sensitivity analysis identified that previous settlement exhibited the highest global and first-order sensitivity. In addition, MFC and ESR in FWM and MWM also had large influence coefficients, demonstrating that FWM and MWM should be prioritized for reinforcement. For these two types of soil, the influence coefficient of MFC was greater than the influence coefficient of ESR, indicating that water stopping measures should be prioritized over reinforcement measures.

(4) A project of an existing tunnel under-crossed by a shield tunnel was used as a case study to verify the effectiveness of MPTS. The maximum error between the predicted and measured settlements was 0.466 mm, with an average error of 0.181 mm. The results indicate that the developed MPTS model exhibits high accuracy in predicting interlayer soil settlement. In addition, the accumulated settlement after reinforcement was reduced by 79.97%, indicating that the improvement measures based on MPTS model can effectively reduce the interlayer soil settlement.

In this paper, MFC was introduced to evaluate the mechanical properties of soil. However, this study focused solely on the impact of moisture content on mechanical behavior and did not account for certain aspects of soil heterogeneity, such as uneven moisture distribution caused by rainfall and variations in pore structure. Therefore, the influence of rainfall-induced moisture distribution and changes in pore structure on mechanical behavior are important issues that warrant attention and further investigation in future research.

Data availability

All raw data used in the training can be downloaded from the following link: https://www.researchgate.net/publication/387974212_Data_availability?channel=doi&linkId=6785bea71afb4e11f5e51bd5&showFulltext=true.

CRediT authorship contribution statement

Boyu Jiang: Methodology, Writing – original draft, Formal analysis. **Haibin Wei:** Project administration, Funding acquisition. **Dongsheng Wei:** Resources, Software. **Zipeng Ma:** Resources, Writing – review & editing. **Fuyu Wang:** Validation, Visualization.

Declaration of competing interest

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

Acknowledgement

Authors acknowledge that this work was supported by the National Natural Science Foundation of China (Grant No. 51578263); China Railway 22 Bureau Group Track Engineering Co., Ltd. is gratefully acknowledged for generously providing the experimental platform. The anonymous reviewers' constructive comments are also cordially appreciated.

References

Ashrafian, A., Panahi, E., Salehi, S., Karoglou, M., & Asteris, P. G. (2023). Mapping the strength of agro-ecological lightweight concrete containing oil palm by-product using artificial intelligence techniques. *Structures*, 48, 1209–1229.

Asteris, P. G., Karoglou, M., Skentou, A. D., VaSoncelos, G., He, M., Bakolas, A., Zhou, J., & Armaghani, D. J. (2024a). Predicting uniaxial compressive strength of rocks using ann models: Incorporating porosity, compressional wave velocity, and schmidt hammer data. *Ultrasonics*, 141.

Asteris, P. G., Tsavdaridis, K. D., Lemonis, M. E., Ferreira, F. P. V., Le, T.-T., Gantes, C. J., & Formisano, A. (2024b). Ai-powered gui for prediction of axial compression capacity in concrete-filled steel tube columns. *Neural Computing & Applications*, 36(35), 22429–22459.

Barkhordari, M. S., Armaghani, D. J., & Asteris, P. G. (2023). Structural damage identification using ensemble deep convolutional neural network models. *Cmes-Computer Modeling in Engineering & Sciences*, 134, 835–855.

Boser, B., Guyon, I., & Vapnik, V. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, 144–152.

Bui, D. K., Nguyen, T., Chou, J. S., Nguyen-Xuan, H., & Ngo, T. D. (2018). A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete. *Construction and Building Materials*, 180, 320–333.

Cavaleri, L., Chatzarakis, G. E., Di Trapani, F., Douvika, M. G., Foskolos, F. M., Fotos, A., Giovanis, D. G., Karypidis, D. F., Livieratos, S., Roinos, K., Tsaris, A. K., Vaxevanidis, N. M., Vougioukas, E., & Asteris, P. G. (2016). Surface roughness prediction of electro-discharge machined components using artificial neural networks. *Proceedings of the 5th International Conference on Integrity-Reliability-Failure (IRF)*. Universidade do Porto, Faculdade de Engenharia.

Chen, C., Zhang, L. M., Pei, L., & Wu, Z. Y. (2020). Soil deformations induced by particle removal under complex stress states. *Journal of Geotechnical and Geoenvironmental Engineering*, 146(9), 04020085.

Da Silva, M. F., Pereira Ribeiro, M. M., Furlan, A. P., & Pessa Fabbri, G. T. (2021). Effect of compaction water content and stress ratio on permanent deformation of a subgrade lateritic soil. *Transportation Geotechnics*, 26, 100443.

Deng, P.X. (2011) Research on the subway travel demand based on multiply social & economic factors. [Master's thesis, Central South University]. (in Chinese).
Do, N. A., Dias, D., Vu, T. T., & Dang, V. K. (2021). Impact of the shield machine's performance parameters on the tunnel lining behaviour and settlements. *Environmental Earth Sciences*, 80(16), 507.
Feng, X., Wang, P., Liu, S., Wei, H., Miao, Y., & Bu, S. (2022). Mechanism and law analysis on ground settlement caused by shield excavation of small-radius curved tunnel. *Rock Mechanics and Rock Engineering*, 55(66), 3473–3488.
Ghanizadeh, A. R., Ghanizadeh, A., Asteris, P. G., Fakharian, P., & Armaghani, D. J. (2023). Developing bearing capacity model for geogrid-reinforced stone columns improved soft clay utilizing mars-ebs hybrid method. *Transportation Geotechnics*, 38, 100906.
He, J., Yang, Z., & Zhang, X. (2023). Effect of undercrossing shield tunnels excavation on existing rectangular pipe-jacking tunnels. *Applied Sciences-Basel*, 13(22), 12235.
Hebbal, A., Balesdent, M., Brevault, L., Melab, N., & Talbi, E.-G. (2023). Deep gaussian process for multi-objective bayesian optimization. *Optimization and Engineering*, 24(3), 1809–1848.
Huang, S., Huang, J., Kelly, R., Jones, M., & Kamruzzaman, A. H. M. (2024). Settlement prediction of the ballina embankment, australia, considering creep. *Journal of Geotechnical and Geoenvironmental Engineering*, 150(5), 04024025.
Huang, Z., Zhang, H., Long, Z., Qiu, W., Meng, G., & Zhu, L. (2021). Field test optimization of shield tunnelling parameters undercrossing an existing high-speed railway tunnel: A case study. *Geotechnical and Geological Engineering*, 39(2), 1381–1398.
Jallow, A., Ou, C.-Y., & Lim, A. (2019). Three-dimensional numerical study of long-term settlement induced in shield tunneling. *Tunnelling and Underground Space Technology*, 88, 221–236.
Kennedy, J., & Eberhart, R. (1995) Particle swarm optimization. *IEEE International Conference on Neural Networks - Conference Proceedings, NEW YORK.*
Kovacevic, M. S., Bacic, M., Gavin, K., & Stipanovic, I. (2021). Assessment of long-term deformation of a tunnel in soft rock by utilizing particle swarm optimized neural network. *Tunnelling and Underground Space Technology*, 110, 103838.
Lu, N., & Kaya, M. (2014). Power law for elastic moduli of unsaturated soil. *Journal of Geotechnical and Geoenvironmental Engineering*, 140(1), 46–56.
Maeda, M., & Kushiyama, K. (2005). Use of compact shield tunneling method in urban underground construction. *Tunnelling and Underground Space Technology*, 20(2), 159–166.
Mair, R. J., Taylor, R. N., & Burland, J. B. (1996). Prediction of ground movements and assessment of risk of building damage due to bored tunnelling. Rotterdam: Geotechnical aspects of underground construction in soft ground.
Ngoc, T. T., Dai, L. V., & Thuyen, C. M. (2021). Support vector regression based on grid search method of hyperparameters for load forecasting. *Acta Polytechnica Hungarica*, 18(2), 143–158.
Pang, H.-X., Dong, W.-D., Xu, Z.-H., Feng, H.-J., Li, Q., & Chen, Y.-T. (2011). Novel linear search for support vector machine parameter selection. *Journal of Zhejiang University-Science C-Computers & Electronics*, 12(11), 885–896.
Peck, R. B. (1969). Deep excavations and tunneling in soft ground. *Proceedings of the International Conference on Soil Mechanics and Foundation Engineering*. .
Ran, L., Ye, X. W., Ming, G., & Dong, X. B. (2014). Structural monitoring of metro infrastructure during shield tunneling construction. *The Scientific World Journal*, 2014, 784690.
Riaz, S., Kikumoto, M., Basharat, M., & Putra, A. D. (2021). Wetting induced deformation of soils triggering landslides in pakistan. *Geotechnical and Geological Engineering*, 39(8), 5633–5649.
Samaniego, E., Anitescu, C., Goswami, S., Nguyen-Thanh, V. M., Guo, H., Hamdia, K., Zhuang, X., & Rabczuk, T. (2020). An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Computer Methods in Applied Mechanics and Engineering*, 362, 112790.
Sarir, P., Chen, J., Asteris, P. G., Armaghani, D. J., & Tahir, M. M. (2021). Developing gep tree-based, neuro-swarm, and whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns. *Engineering with Computers*, 37(1), 1–19.
Shariati, M., Mafipour, M. S., Ghahremani, B., Azarhomayun, F., Ahmadi, M., Trung, N. T., & Shariati, A. (2022). A novel hybrid extreme learning machine-grey wolf optimizer (elm-gwo) model to predict compressive strength of concrete with partial replacements for cement. *Engineering with Computers*, 38(1), 757–779.
Sharma, B., & Bora, P. K. (2003). Plastic limit, liquid limit and undrained shear strength of soil-reappraisal. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(8), 774–777.
Shen, Y., Zhang, D., Wang, R., Li, J., & Huang, Z. (2023). Sbd-k-medoids-based long-term settlement analysis of shield tunnel. *Transportation Geotechnics*, 42.
Shim, J., & Hwang, C. (2010). Support vector quantile regression with weighted quadratic loss function. *Communications for Statistical Applications and Methods*, 17, 183–191.
Spagnoli, G., & Feinendegen, M. (2017). Relationship between measured plastic limit and plastic limit estimated from undrained shear strength, water content ratio and liquidity index. *Clay Minerals*, 52(4), 509–519.
Vapnik (1964). A note one class of perceptrons. *Automation and Remote Control*.
Vorster, T. E. B., Klar, A., Soga, K., & Mair, R. J. (2005). Estimating the effects of tunneling on existing pipelines. *Journal of Geotechnical and Geoenvironmental Engineering*, 131(11), 1399–1410.
Wang, T., Cao, X., Li, Y., Zhai, Y., & Ye, G. (2023). Real-time anomaly data detection method based on mixed kernel function pso-svr. *Aip Advances*, 13(6).
Wen, Z., Guo, L. M., Meng, S. P., Rong, X. L., & Shi, Y. H. (2023). Iot-assisted feature learning for surface settlement prediction caused by shield tunnelling. *Computer Communications*, 203, 276–288.
Wong, L. N. Y., Liu, Z. H., Tse, K. K. C., Cheung, S. H., & Yu, L. Q. (2023). A computational algorithm for calculating fracture index of core runs. *Rock Mechanics and Rock Engineering*, 56(9), 6905–6918.
Wu, W., Yun, Y., Sun, Y., & Zhang, G. (2023). *The geography of subway development and commuting behavior*. Tunnelling and Underground Space Technology (pp. 133). .
Xue, X., Zhang, R., Zhang, X., Yang, R. J., & Li, H. (2015). Environmental and social challenges for urban subway construction: An empirical study in China. *International Journal of Project Management*, 33(3), 576–588.
Yang, C., & Jiang, Z. (2022). A discrete-time model-based method for predicting settlement of geotechnical foundations in buildings. *Mobile Information Systems*, 2022(1), 5631634.
Zhai, J. (2022). *Machine-learning-based road soft soil foundation treatment and settlement prediction method*. Scientific Programming.
Zhang, C., Hu, S., & Lu, N. (2022a). Unified elastic modulus characteristic curve equation for variably saturated soils. *Journal of Geotechnical and Geoenvironmental Engineering*, 148(1), 04021171.
Zhang, D.-M., Chen, S., Wang, R.-C., Zhang, D.-M., & Li, B.-J. (2021). Behaviour of a large-diameter shield tunnel through multi-layered strata. *Tunnelling and Underground Space Technology*, 116, 104062.
Zhang, Z. Y., Xu, R. Q., Wu, X., & Wang, J. C. (2022b). Ann-based dynamic prediction of daily ground settlement of foundation pit considering time-dependent influence factors. *Applied Sciences-Basel*, 12(13), 6324.
Zhao, F., Lu, X., Shi, H., Liu, B., Liu, S., Dai, K., & Fan, Y. (2023). Study on stratified settlement and weak reflectivity fiber grating monitoring of shield tunnel crossing composite strata. *Applied Sciences-Basel*, 13(3), 1769.
Zhou, J., Asteris, P. G., Armaghani, D. J., & Binh Thai, P. (2020a). Prediction of ground vibration induced by blasting operations through the use of the bayesian network and random forest models. *Soil Dynamics and Earthquake Engineering*, 139, 106390.
Zhou, Z., Liu, Z.-Z., Yang, H., Gao, W.-Y., & Zhang, C.-C. (2020b). Freeze-thaw damage mechanism of elastic modulus of soil-rock mixtures at different confining pressures. *Journal of Central South University*, 27(2), 554–565.