ISSN: 1813-4890

Ground Surface Settlement Prediction for Rectangular Pipe Jacking Tunnels Based on Machine Learning Algorithms

Rong Hu [1], Yongsuo Li [1,2,*], Kunpeng Wei [1], Min Tan [1]

[1] School of Civil Engineering, Hunan City University, Yiyang 413002, China;

[2] Hunan Engineering Research Center of Structural Safety and Disaster Prevention for Urban Underground Infrastructure, Hunan City University, Yiyang 413002, China.

Abstract

This paper presents the prediction of surface settlement induced by the excavation of rectangular pipe jacking tunnels using four machine learning (ML) algorithms. The settlement database was derived from the West Extension Project of Liuye Avenue in Hunan Province, where 104 data indicators from the right side of the tunnel (including jacking force, excavation speed, grouting pressure, and settlement) were selected, with 80 of them serving as input parameters for the ML models. Hyperparameter tuning based on particle swarm optimization (PSO) was employed to effectively explore the optimal combinations and enhance prediction performance. The performance of the ML models was evaluated by comparing the mean squared error (MSE), mean absolute error (MAE), and coefficient of determination (R^2). The results indicate that the PSO-SVR model outperforms other models in terms of surface settlement prediction accuracy and generalization ability, with MSE, MAE, and R^2values of 0.294, 0.437, and 0.909, respectively.

Keywords

Machine Learning, Particle Swarm Optimization, Ground Surface Settlement, Rectangular Pipe Jacking.

1. Introduction

With the continuous advancement of urbanization in China, the development of urban underground space has been further promoted to alleviate the tension of urban land use and traffic congestion. The length of subway lines has even reached one-third of the total operating mileage[1]. During the tunnel excavation process, surface settlements induced by the tunneling may cause significant damage to existing structures. With the rapid development of artificial intelligence and the increasing interdisciplinarity of various fields, machine learning algorithms, as a method for studying the intrinsic relationships and patterns within data, have provided a more valuable approach to solving the problem of ground deformation prediction during shield tunneling construction.

In 1998, Shi[2] et al. used a BP neural network to predict deformation caused by tunnel excavation in Brasília and found that the error was reduced by half compared to conventional models. Since then, numerous scholars have attempted to use artificial neural network models, primarily BP neural networks, to predict ground deformation caused by shield tunneling, achieving promising results. Additionally, Support Vector Machines (SVM)often exhibit higher prediction accuracy than artificial neural networks when dealing with small datasets. Random Forest (RF), as an ensemble machine learning algorithm, has gained popularity in predicting ground deformation due to its high accuracy and ability to process large amounts of data quickly. Zhou[3] et al. established an intelligent model based on Random Forest to verify its applicability in risk prediction, assisting on-site engineers in determining safety risks.

Ramezanshirazi[4] et al. and others have also used machine learning algorithms to effectively predict ground surface settlement caused by tunneling. Mahmoodzadeh[5] et al. compared the efficiency and feasibility of various machine learning algorithms to study the accuracy of a single algorithm in predicting ground settlement, providing guidance for identifying the evolution of tunnel-induced settlement.

To better understand the application of machine learning in predicting ground deformation caused by tunneling, this study reviewed domestic and international literature from recent years and found that the algorithms most frequently used in previous studies are Random Forest (RF), Support Vector Machine (SVM), and Back Propagation Neural Network (BPNN). Subsequently, the superiority of certain algorithms was illustrated by comparing various algorithmic models. However, to date, there is still no specific method to determine which algorithm is the most suitable for predicting tunnel settlement. Generally, the performance of algorithmic models is improved by setting the hyperparameters of different algorithms. Therefore, it is worthwhile to study the performance comparison of various machine learning algorithms in the same case, given the lack of a machine learning algorithm for settlement prediction with robust training. Focusing on the prediction of tunnel settlement, this study combines the improved intelligent optimization algorithm (PSO) with Support Vector Regression (SVR), relying on the real-time monitoring data from the construction of the rectangular pipe jacking tunnel on the West Extension of Liuye Avenue to establish a PSO-SVR surface settlement prediction model. This model is compared and analyzed with traditional Random Forest, BP neural network, and SVM models to identify the optimal prediction model, in the hope of providing theoretical support and technical reference for the prediction of surface settlement in similar shield tunnels.

2. Machine Learning Algorithms

This work introduces four widely used artificial intelligence algorithms: Random Forest (RF), Support Vector Regression (SVR), and Back Propagation Neural Network (BPNN). Additionally, the Particle Swarm Optimization algorithm (PSO) is integrated into the Support Vector Regression (SVR)model to optimize its hyperparameters. These algorithms are extensively applied in underground engineering. Below, we focus on the improved PSO-SVR algorithm.

2.1. PSO-SVR Algorithm

The PSO algorithm is incorporated into the SVR model to optimize its hyperparameters, addressing the shortcomings brought by the randomness of initial hyperparameters. The flowchart of the algorithm is shown in Fig. 1. The PSO-SVR model process can be divided into the following three stages:

In the first stage, the SVR and PSO algorithms are initialized, and the optimization parameters of the SVR model are listed. For the SVR model, the kernel, regularization parameter C, and kernel coefficient epsilon are optimized. The particle swarm optimization algorithm defines key parameters such as population size, particle velocity and position, as well as the fitness function, in preparation for the hyperparameter optimization of the SVR model.

In the second stage, the assembled model is iteratively updated. According to the principles of the particle swarm optimization algorithm, the velocity and position of particles are updated using Equations (1) and (2). Subsequently, the combined model calculates the global minimum fitness of the particle swarm optimization algorithm and the error of the cross-validation samples, and updates the initial hyperparameters.

In the third stage, the optimal solution is output. The model determines whether the error meets the termination condition. If the termination condition is satisfied, the optimal hyperparameters determined by the global extremum position of the SVR model are output; otherwise, the iterative learning process will continue.

$$v_i^d = \omega \times v_i^d + c_1 \times rand_1^d \times (pBest_i^d - x_i^d) + c_2 \times rand_2^d \times (gBest^d - x_i^d) \quad (1)$$

$$x_i^d = x_i^d + \upsilon_i^d \quad (2)$$

In the formula: $i = 1, 2, \ldots, N$, N represents the total number of particles in the swarm. ω denotes the inertia weight, v_i represents the velocity of particle I, $rand()$ is a random number uniformly distributed between (0,1), x_i indicates the current position of particle I, c_1 and c_2 are the learning factors.

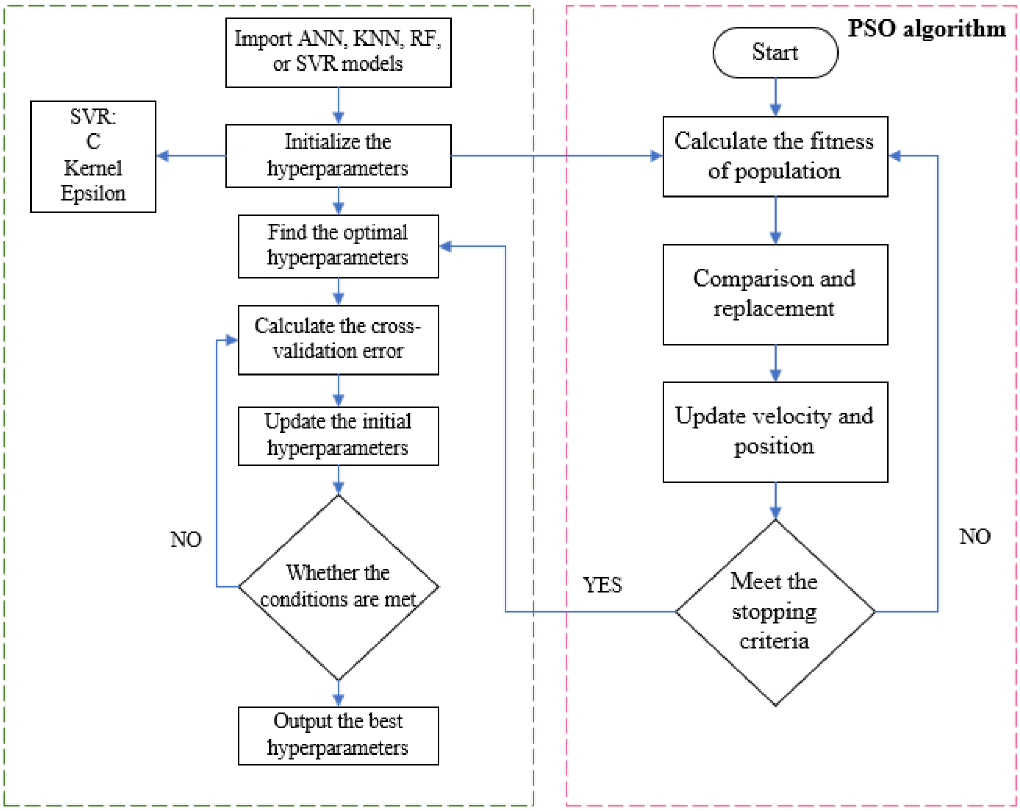

Fig.1 PSO-SVR flow chart

2.2. Hyperparameter selection

The hyperparameters for the four machine learning models are Table 1:

Table 1 Hyperparameter selection

Algorithmic models	Hyperparameter settings	Valid values
SVM	c	1
	gamma	1
RF	n_estimators	100
BP	learning_rate	0.03
	hidden_dim	20
	loss	0.001
PSO-SVR	kernel	1.169
	c	23.250
	epsiln	0.961

3. Engineering Case Analysis

The construction of Liuye Avenue adopts the manufacturing process of "oblique intersection, oblique construction, and vertical jacking." The frame bridge construction involves setting up working pits, back walls, slip plates, and on-site assembly of shield support frames. The frame bridge intersects with the Changzhang Expressway at an angle of 63.881°, with a cover soil thickness of approximately 2.000-2.683 m. During the construction process, the normal traffic on the expressway cannot be interrupted. However, the jacking construction will disturb the subgrade, causing settlement. Excessive settlement can affect the safety of normal traffic. Therefore, it is necessary to arrange monitoring points on the road surface. To ensure that the frame bridge meets the requirements during the construction process and to ensure construction safety, monitoring points are arranged at key points of the frame bridge and the steel shield.

4. Database Construction

4.1. Input Parameter Selection

The data used in this study originates from the construction monitoring data of Liuye Avenue in Changde City, Hunan Province. The tunnel was constructed using the pipe jacking method, a rapidly developing non-excavation tunnel construction method following the shield method. The database consists of two excavation variables and one grouting variable, totaling 26 datasets. Each dataset is updated in real-time by the excavation equipment and covers the parameter ranges of each variable. In this study, the first 80% (20 datasets) are used as the training set, and the remaining 20% (six datasets) are used as the prediction set. The specific dataset is shown in Table 2 below:

Table 2 Data Set

Variable	Parameter type	Data			
		Min.	Max.	Ave.	S.D.
Jacking force/KN	Input	15601.47	48504.561	34586.2	9814.75
Excavating velocity/(m/d)	Input	0.825	1.985	1.64	0.223
Grouting pressure/MPa	Input	0.063	0.416	0.27	0.079
Settlement value/mm	Output	-4	2	-0.23	1.245

4.2. Data Preprocessing

Considering the limited number of data sets but the large size of individual data, a normalization process is required. By using a linear equation to scale the data to a specific range, the normalization process can reduce errors caused by significant data differences during training, thereby increasing training efficiency and overall model accuracy. Additionally, normalization can render the data dimensionless, reducing differences between data points and facilitating better data analysis and comparison.

In this study, the traditional machine learning approach of using gradient descent to calculate the minimum fitness value was continued. Through an iterative process, the gradient information of the equation is updated in real time, allowing for a more accurate and effective search for the optimal solution of the parameters.

5. Model Evaluation

The error metrics of the models (as shown in Fig.2-5) indicate that the PSO-SVR model has an MSE of 0.294, MAE of 0.437, and R^2 of 0.909. Fig.2-3 shows that the PSO-SVR model has the lowest *MAE* and *MSE* values, indicating the best fit and strongest generalization ability. Among the three traditional models, the BP neural network demonstrates superior prediction performance but requires longer computation time and exhibits some overfitting. Overall, the PSO-SVR model outperforms others in accuracy and stability.

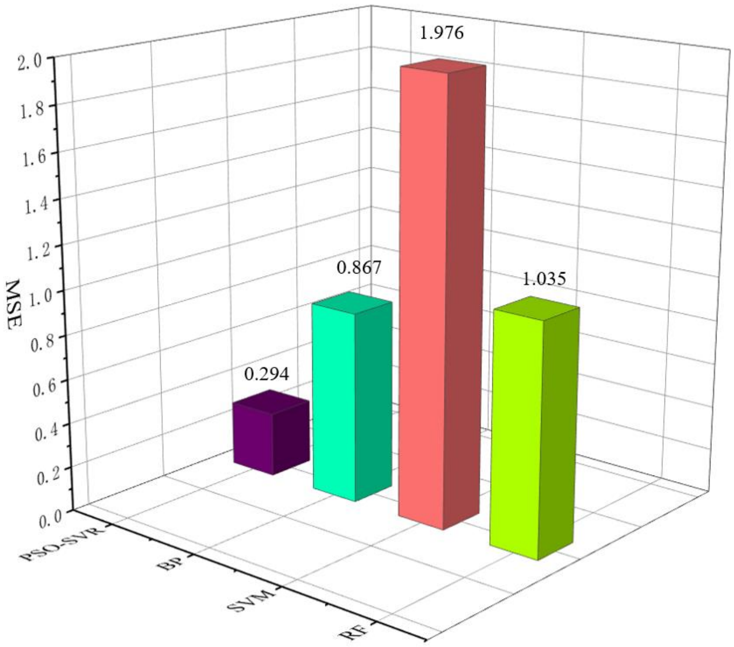

Fig. 2 Comparison of MSE error values

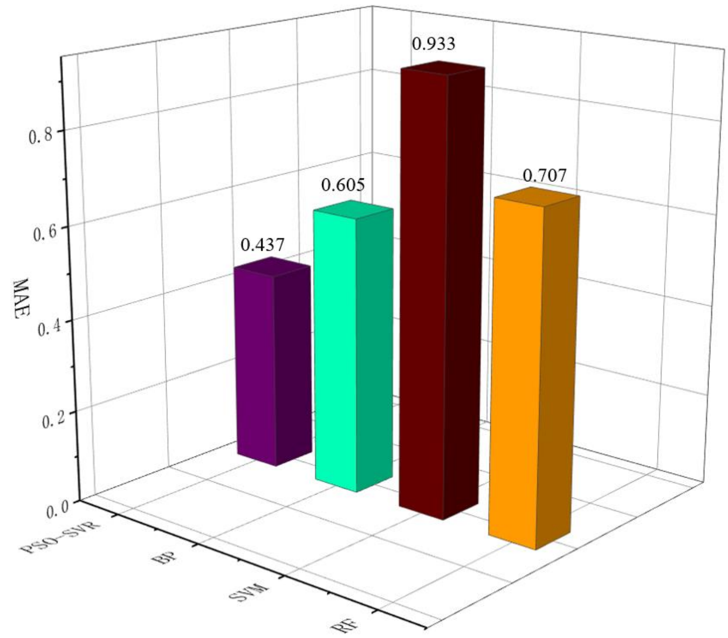

Fig. 3 Comparison of MAE error values

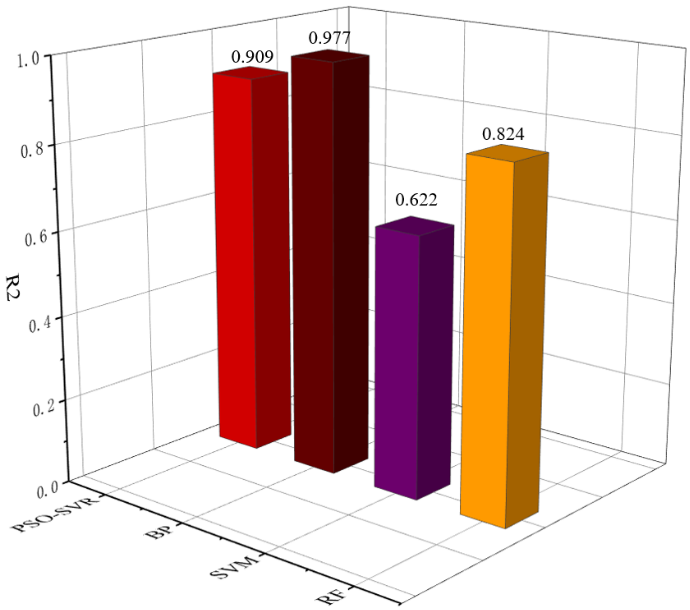

Fig. 4 Comparison of R2 error values

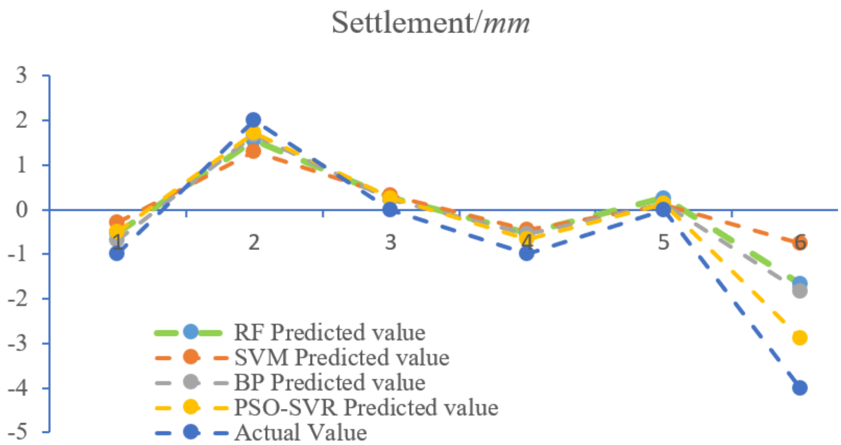

Fig. 5 Comparison chart of settlement predictions

6. Conclusion

In this study, a machine learning (ML)-based prediction framework for estimating surface settlement following tunnel excavation was developed. The methods for pre-processing raw data, error analysis, and hyperparameter selection were also introduced. Finally, based on the field records from the rectangular pipe jacking tunnel project on the West Extension of Liuye Avenue, a comprehensive comparison of the performance of four models was conducted, and the main conclusions are as follows:

1. The proposed intelligent prediction framework consists of three stages: database establishment, algorithm model construction, and model evaluation. The primary processes include the selection of input parameters, data set partitioning, data pre-processing, and optimization using the particle swarm optimization (PSO) algorithm. Through the analysis of 26 data sets from the rectangular pipe jacking tunnel on Liuye Avenue, it was found that the proposed ML algorithm prediction framework is rational and can provide a reference for predicting surface settlement in rectangular pipe jacking tunnels.

2. In this study, data normalization was performed to avoid large differences between features, which could otherwise cause some features to have a disproportionate influence on the model. This process also enhanced the convergence speed and accuracy of the model. Moreover, a

comparative analysis was conducted between models with optimized hyperparameters and those without optimization. The results indicated that the optimal model, within a relatively short period, outperformed other models in all metrics except for R^2, which was 0.068 lower than that of the back propagation neural network (BPNN). The model exhibited superior performance in predicting surface settlement. The support vector regression (SVR) algorithm improved by particle swarm optimization demonstrated higher accuracy than other models, with a reduction in error rates and enhanced generalization ability.

3. The interaction mechanism between the pipe jacking machine and the soil strata is complex. The intelligent model can effectively address the highly nonlinear issues of construction parameters and provide a reference for safe and efficient pipe jacking construction. Future research will thoroughly investigate the pipe-soil interaction mechanism and leverage more engineering data to conduct research on the application of artificial intelligence techniques in the field of tunnel settlement prediction.

Acknowledgements

The authors sincerely acknowledge the vital facilities and resources provided by general project of Hunan college students' innovation training program (No. S202311527035). We are also grateful for the valuable contributions and insightful suggestions from Professor Hu Da, whose guidance has significantly enhanced the quality of this research. Additionally, we appreciate the constructive comments from the anonymous reviewers, which have helped improve the clarity and robustness of this manuscript.

References

[1] Guo, S. Y. (2018). Development and prospects of tunnel and underground engineering technology in China. Modern Tunnel Technology , 55(S2), 1-14.

[2] Shi J, Ortigao J, Bai J. Modular neural networks for predicting settlements during tunneling[J]. Journal of Geotechnical and Geoenvironmental Engineering, 1998, 124(5): 389-395.

[3] Zhou Y, Li S, Zhou C, et al. Intelligent approach based on random forest for safety risk prediction of deep foundation pit in subway stations[J]. Journal of Computing in Civil Engineering, 2019, 33(1): 05018004.

[4] Ramezanshirazi M, Sebastiani D, Miliziano S. Artificial intelligence to predict maximum surface settlements induced by mechanized tunnelling[C]. Geotechnical Research for Land Protection and Development: Proceedings of CNRIG 2019 7, 2020: 490-499.

[5] Mahmoodzadeh A, Nejati H R, Mohammadi M, et al. Forecasting tunnel boring machine penetration rate using LSTM deep neural network optimized by grey wolf optimization algorithm[J]. Expert Systems with Applications, 2022, 209: 118303.