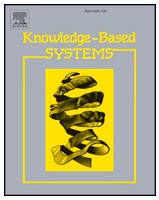
Knowledge-Based Systems 286 (2024) 111413

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys



Multisource information fusion for real-time optimization of shield construction parameters

Hongyu Chen [a], Jun Liu [b], Geoffrey Qiping Shen [a], Luis Martínez [c], Muhammet Deveci [d,e,f,*], Zhen-Song Chen [g], Yang Liu [h,i]

[a] Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[b] School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China
[c] Department of Computer Science, University of Jaén, Jaén 23071, Spain
[d] Department of Industrial Engineering, Turkish Naval Academy, National Defence University, 34942 Tuzla, Istanbul, Turkey
[e] The Bartlett School of Sustainable Construction, University College London, 1-19 Torrington Place, London WC1E 7HB, UK
[f] Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon
[g] School of Civil Engineering, Wuhan University, Wuhan 430072, China
[h] ZhongNan Hospital of Wuhan University, Wuhan University, Wuhan 430071, China
[i] Economics and Management School of Wuhan University, Wuhan 430072, China

ARTICLE INFO

Keywords:
Shield construction parameter
Multiobjective optimization
Surface settlement
Cutter wear
Advance speed
BO-RF
NSGA-III

ABSTRACT

This paper introduces a hybrid intelligent framework that combines Bayesian optimization (BO), a random forest (RF) model, and the nondominated sorting genetic algorithm-III (NSGA-III) for the optimization and control of tunnel shield construction parameters. The BO-RF method establishes a nonlinear mapping function between the input variables and three targets, surface settlement, cutter wear, and advance speed, serving as the fitness function for NSGA-III. Model interpretability analysis is conducted using Shapley Additive ExPlanations (SHAP). A multiobjective intelligent optimization model is formulated with NSGA-III, targeting surface settlement, cutter wear, and advance speed. A case study validates the applicability and effectiveness of this approach, leading to the following conclusions: (1) The BO-RF algorithm yields highly accurate prediction results, with R^2 values ranging from 0.930 to 0.938, RMSE ranging from 0.138 to 0.172, and MAE ranging from 0.112 to 0.138 for the three targets. (2) The optimization results for surface settlement, cutter wear, and advance speed are outstanding, with an average improvement of 12.56 %. The simultaneous adjustment of the three shield construction parameters leads to the best optimization results, with an average improvement of 19.67 %. (3) The energy consumption of the shield drive system decreases by an average of 10.70 %, and the optimization improvement for the first three objectives decreases by an average of 1.82 %, 1.46 %, and 2.23 %, respectively. By introducing the integrated BO-RF-NSGA-III algorithm, this study contributes to the field of tunnel engineering optimization management.

1. Introduction

In recent years, China's rapid urbanization has increased the demand for underground spaces, increasing the prevalence of tunnel construction [1]. Shield machines have become the most widely used large construction devices in urban tunnel construction due to their advantages of not affecting ground traffic, high technical and economic advantages, high degree of automation, and construction not being affected by climatic conditions. Given the high complexity and variability of geological conditions, uncertainty surrounding machinery and equipment, and other factors, such as human operation errors, it can be challenging to mitigate safety risks [2], reduce project costs [3], and adhere to construction timelines [4]. Surface settlement is an inevitable issue during shield construction and can cause significant harm to surrounding structures and underground facilities, posing safety risks [5]. The significant wear of the cutter frequently necessitates a halt in the tunneling process for blade replacement, leading to extended construction periods and a reduction in construction efficiency. In addition, opening storehouses and changing cutters can easily result in engineering accidents and increase construction risk and cost, especially in karst tunnels, for which the wear problem is often prominent. The

* Corresponding author at: Department of Industrial Engineering, Turkish Naval Academy, National Defence University, 34942 Tuzla, Istanbul, Turkey
E-mail address: muhammetdeveci@gmail.com (M. Deveci).

https://doi.org/10.1016/j.knosys.2024.111413
Received 15 November 2023; Received in revised form 8 January 2024; Accepted 17 January 2024
Available online 18 January 2024

Nomenclature

GM	grouting amount, m^3 x_1
CS	cutterhead speed, $rad \cdot min^{-1}$ x_2
GT	gross thrust, kN x_3
GP	grouting pressure, bar x_4
CT	cutterhead torque, $kN \cdot m^{-1}$ x_5
CEP	chamber earth pressure, bar x_6
SM	shield muck, m^3 x_7
FA	foam agent, m^3 x_8
BD	burial depth, m x_9
CSR	cover-span ratio x_{10}
IFA	internal friction angle, x_{11}
SC	soil cohesion, kPa x_{12}
SS	surface settlement, mm f_1
CW	cutter wear, mm f_2
AS	advance speed, $mm \cdot min^{-1}$ f_3
ECS	energy consumption of the shield drive system, $kW \cdot h$ f_4
BO	Bayesian optimization
RF	random forest
NSGA-III	nondominated sorting genetic algorithm-III
SHAP	Shapley additive explanations

advance speed is also crucial; excessive speed can lead to increased surface settlement due to rapid stress release and pore water pressure changes, while overly slow progress affects the construction timeline [6]. In view of the above problems, it is essential to consider the construction period, cost and safety and to employ effective methods for the real-time prediction and control of surface settlement, cutter wear reduction, and advance speed optimization, thereby facilitating efficient, economic and safe tunnel construction.

Scholars have been studying surface settlement, cutter wear and advance speed in shield construction. The methods used are usually empirical formula methods, model test methods and numerical simulation methods. For empirical methods, Gui et al. [7] employed an empirical equation that assumes a Gaussian distribution curve for the settlement profile of tunnels to estimate surface settlement arising from double-O-tube tunnel shield construction. Wang et al. [8] established a hob compound wear test device model based on the working parameters of a tunnel boring machine (TBM) and calculated the degree of hob wear considering the cutter wear mechanism. In model test methods, Hu et al. [9] employed a miniature shield machine model to simulate shield construction in pebble-rich earth pressure balance (EPB) tunnels and investigated the resulting surface settlement. Sun et al. [10] constructed a hob compound wear test model based on TBM working principles and established a cutter wear prediction method. In numerical simulation methods, Melis et al. [11] developed a numerical model to simulate surface settlement resulting from shield machines using Madrid Metro engineering data and compared the results to previous monitoring data. Fang et al. [12] coupled 3D rigid body dynamics (RBD) with the discrete element method (DEM) to simulate hob cutting processes and study hob performance.

While each of the above three methods can be used to obtain surface settlement, cutter wear and advance speed, they have respective limitations. For example, the empirical method obtains surface settlement by intuitive fitting based on field observations, which heavily relies on empirical relevance and experience and requires some assumptions. However, at complex construction sites, empirical methods cannot be flexibly adjusted, resulting in a limited prediction scope [13]. Model tests have the disadvantages of high cost, long duration, and failure to consider all the details of the excavation process [14]. Numerical simulations are very accurate and can be adjusted to specific site conditions, complex construction procedures and nonlinear soil behavior; however, the process of establishing and verifying models is time-consuming. Obtaining accurate shield construction parameters for numerical simulations, particularly in unpredictable geological environments, presents significant challenges [15]. The above methods are usually not comprehensive enough, resulting in poor prediction accuracy for surface settlement, cutter wear and advance speed.

With the rapid advancement of artificial intelligence, machine learning (ML) has found applications across diverse domains in civil engineering and many other industries [16,17]. ML, by learning from extensive datasets, captures the nonlinear mapping relationship between outcomes and characteristic variables, taking into account important construction factors. Additionally, ML methods maximize the use of available data and do not ignore information. Therefore, ML methods, such as back-propagation neural network (BPNN), support vector machine (SVM), and gradient boosting decision tree (GBDT), are applied in underground engineering prediction [18–20]. However, BPNN a certain network structure and a significant amount of data to achieve accurate predictions [21]. When the data availability is limited, the core functions of SVM are limited, and insensitive parameters and penalty weight parameters are difficult to determine [22]. When abundant data are available, GBDT must presort the features before iterations are performed, resulting in high spatial complexity and a long execution time [23].

To address the limitations of the previously mentioned ML algorithms, Breiman [24] proposed an RF intelligent algorithm in 2001. The algorithm combines the advantages of a decision tree and bagging. Different decision trees can be generated through the parallel training of different hosts with high efficiency, and all decision trees are combined in bagging to avoid the overfitting problems caused by single decision trees. Therefore, RF has gained widespread application in the realm of underground space construction [25,26]. Previous studies have demonstrated that prediction models based on RF exhibit robust generalization capabilities and high accuracy, establishing nonlinear correlations between input parameters and burial depth, surface settlement, cutter wear, and advance speed.

Shield tunneling-induced surface settlement is a significant safety concern in tunnel construction. Numerous studies have sought to enhance safety by optimizing critical shield construction parameters to minimize surface settlement [27]. In practical engineering projects, focusing solely on surface settlement control is inadequate; factors such as construction period and cost must also be considered. The control of cutter wear during tunneling is also a key problem in construction, especially in tunnel engineering under poor geological conditions. In the tunneling process, especially in the karst area of Guiyang, considering karst treatment and other factors, the cutter is seriously worn. Problems such as clamping and the need to change cutters due to wear not only reduce the tunneling efficiency and increase the construction period and cost but also have certain safety risks [28]. The advance speed of the TBM reflects the tunneling efficiency, which directly influences the duration and cost of the project [29]. Therefore, the control of surface settlement, cutter wear, and advance speed is crucial in shield engineering. However, these objectives are often conflicting, and achieving optimal values simultaneously is challenging. To address this, a multiobjective optimization framework is formulated to identify the optimal shield construction parameters and optimize surface settlement, cutter wear, and advance speed.

In multiobjective optimization problems, the primary objective is to identify a solution set that optimizes multiple objectives, known as the Pareto solution set [30]. The Pareto solution set is composed of a series of points distributed on the Pareto frontier. Scholars have widely used multiobjective optimization technology in various stages of tunnel engineering. In the design phase, engineers are required to balance multiple factors, including structural safety, construction cost, construction period, and construction feasibility. These objectives are often in conflict; for example, improving structural safety may increase costs, while shortening the construction period may reduce the quality of

construction. Liu et al. [31] introduced an intelligent decision-making approach based on multiobjective optimization, focusing on improving the advance speed and the cutter life of TBM. Guo et al. [32] employed the NSGA-II algorithm to optimize tunnel cost, construction time and comfort level. During the construction phase, tunnel projects require multiple decisions that directly affect the progress and cost of the project and may involve multiple competing objectives. For example, accelerating a construction schedule may lead to increased construction risk, while adopting complex construction methods may increase costs while improving construction efficiency. Fan et al. [33] introduced a constrained multiobjective optimization model aimed at minimizing the construction cycle, construction energy consumption, and construction cost of TBM. Zhang et al. [34] devised a method that combined the eXtreme gradient boosting (XGBoost) and NSGA-II algorithms to reduce the ultimate support pressure and surface deformation during tunnel excavation. In terms of subway maintenance, Guo et al. [35] applied NSGA-III to subway evacuation problems to optimize the evacuation time, density and cost.

Common multiobjective optimization algorithms include NSGA-II, NSGA-III, MOEA/D and MOMPA [36–39]. Among them, NSGA-III incorporates fast nondominated sorting, distance allocation, and reference point technology to maintain diversity in multiobjective optimization and explore the full scope of global optimal solutions, making it effective for handling three or more objectives. In addition, NSGA-III requires the establishment of an objective optimization function. Some studies have used ML algorithms to establish an objective optimization function. Traditionally, the determination of functional relationships between variables involves the use of empirical formulas or simulation software. However, these methods have limitations. Empirical formulas lack flexibility for adjustments based on engineering situations, and simulation software calculations are time-consuming, reducing efficiency [40–42]. Several scholars have suggested using ML methods such as BPNN, SVM, and others to construct nonlinear prediction models, establish fitness functions, and perform optimization [43–45]. In this paper, the objective optimization function of NSGA-III is established based on the BO-RF algorithm, optimizing this function to determine the best construction parameter decision scheme.

The main research questions addressed in this study are as follows: (1) How can the nonlinear relationships between the input parameters and surface settlement, cutter wear and advance speed be established? (2) How can the impact of each input parameter on the surface settlement, cutter wear and advance speed be evaluated, and what are the key optimization parameters? (3) How can the optimal shield construction parameters be determined when the surface settlement, cutter wear and advance speed are optimized simultaneously? This paper introduces a BO-RF-NSGA-III multiobjective intelligent optimization framework. A high-precision prediction model for surface settlement, cutter wear, and advance speed is established using the BO-RF algorithm based on actual monitoring data. Subsequently, the Shapley Additive ExPlanations (SHAP) method is used to determine the key optimization parameters affecting the surface settlement, cutter wear and advance speed. Finally, the trained prediction models serve as the objective optimization functions for NSGA-III, which is then employed to optimize the three objectives according to the proposed optimization principles. The research results can be summarized as follows: (1) A high-precision prediction model for surface settlement, cutter wear, and advance speed is obtained using the BO-RF algorithm and subsequently optimized as the fitness function. (2) The SHAP algorithm is employed to determine the importance of each input parameter to surface settlement, cutter wear, and advance speed, leading to the identification of key optimization parameters. (3) A BO-RF-NSGA-III hybrid approach is developed to predict and optimize the surface settlement, cutter wear and advance speed. Based on the proposed optimization principle, the impact of surface settlement on the surrounding environment is reduced, the construction cost caused by cutter wear is reduced, and the tunneling efficiency is improved.

The paper is structured as follows. Section 2 outlines the process of constructing a multiobjective optimization model using the BO-RF algorithm. Section 3 demonstrates the application of our proposed method by predicting surface settlement, cutter wear, and advance speed for a section of the Guiyang Rail Transit Line 3 and optimizing the shield construction parameters. Section 4 extends the methodology to optimize the energy consumption of the shield machine drive system. Finally, Section 5 provides a summary of the paper and offers potential directions for future research.

2. Methodology

This paper introduces a hybrid optimization method that combines BO-RF and NSGA-III to optimize the surface settlement, cutter wear and advance speed and determine the best shield construction parameter decision scheme, which provides guidance for practical construction. The specific process is visually represented in Fig. 1.

2.1. Data collection and processing

The EPB shield machine utilized in this project is outfitted with an extensive array of sensors dedicated to recording operational conditions throughout the excavation process. The raw data generated by these sensors are gathered by the TBM's data acquisition system and are presented in the form of CSV files, with the data recorded corresponding to changes in the number of rings. These files encompass abundant information, including time, ring number, and shield operating parameters. For this study, the average values of the operating parameters for each ring are computed. Although the information is comprehensive, relatively few geological parameters are available, and the value of cutter wear cannot be automatically recorded, requiring manual measurement. Therefore, the data require preprocessing to extract valuable information suitable for model development. The data processing steps encompass data cleaning, geological parameter interpolation, cutter wear calculation, data normalization and Pearson correlation analysis.

(1) Data cleaning

The shield construction parameter data obtained from the EPB-TBM data acquisition device comprise states from the propulsion and shutdown stages. Numerous records with a parameter value of 0 are generated during the TBM shutdown process. To construct the dataset, it is necessary to delete the blank records to improve the validity of the data. Generally, zero-value data and data at the start and close stages can be eliminated through state discrimination. The state discrimination function is defined as follows:

$$D = d(v)\cdot d(n)\cdot d(F)\cdot d(T) \tag{1}$$

$$d(x) = \begin{cases} 1, if\, x \neq 0 \\ 0, if\, x = 0 \end{cases} \tag{2}$$

where v represents the advance speed, n represents the cutterhead speed, F represents the gross thrust, and T represents the cutter torque. $d(x)$ is the function used to determine whether to retain records, and D is the result of multiplication of the $d(x)$ function. If any attribute in the input dataset contains a zero or invalid value, $d(x)$ is set to 0, causing D to equal 0, at which point the corresponding parameters should be removed. In addition, since we obtained the average value of each tunnel ring, the negative effects of outliers can be mitigated.

(1) Geological parameter interpolation

Before the shield was constructed, core samples were obtained with a drilling machine, and the burial depth data were recorded. These core samples are subsequently sent to a geotechnical mechanics laboratory

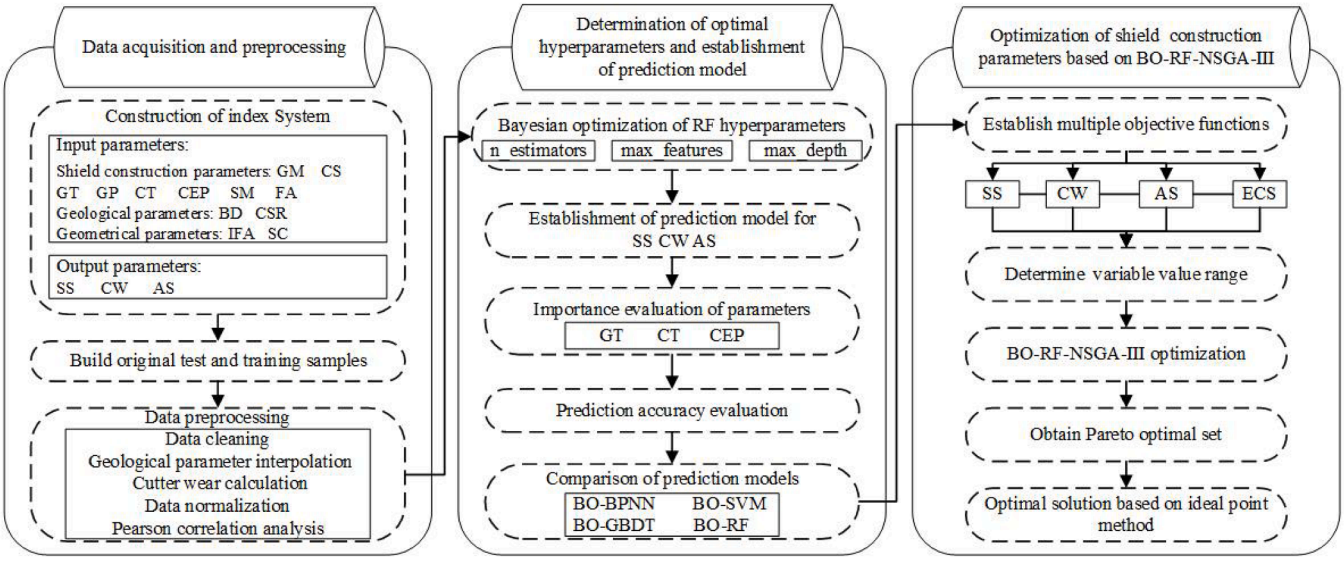


Fig. 1. Flow chart of the optimization of shield construction parameters based on the BO-RF-NSGA-III algorithm.

for testing to obtain geological information about the construction route. This information included the location and depth of the borehole, the depths of the various formations, and the petrophysical and mechanical parameters of the formations. However, due to project budget constraints, it is often not possible to obtain sufficient borehole data, resulting in relatively limited geological information. The original geological distribution data reveal discrete geological depth information, but construction with a shield machine is a continuous process, and borehole-based geological distribution data may not represent the real continuous geological distribution.

To overcome this limitation, the ordinary kriging interpolation method is used to interpolate geological data to more accurately estimate the continuous geological distribution during the construction process of TBM [46]. First, laboratory tests were carried out on drilled core samples from the construction area, and the geological parameters of these samples were obtained. The common kriging method was subsequently applied to interpolate the burial depth data for the various strata and obtain the continuous geological distribution within the construction interval. Finally, the excavation stratum of each ring was determined according to the interpolated and complete map of the geological data, and the soil parameters, including the internal friction angle and soil cohesion, of each ring tunnel were obtained. These parameters are crucial for shield design and the formulation of tunneling strategies.

The ordinary kriging method requires spatial assumptions, such as the properties at every point in space are uniform and every point is characterized by the same mathematical expectation and variance. The purpose of ordinary kriging interpolation is to find a set of coefficients λ_i for a point (x, y) in space with attribute values as follows:

$$\widehat{z}_0 = \sum_{i=1}^{n} \lambda_i z_i \tag{3}$$

where $\widehat{z}_0$ represents the estimated value at point (x_0, y_0), namely, $z_0 = z(x_0, y_0)$, and z_i denotes the known attribute value.

Kriging interpolation is an unbiased estimation method that requires unbiased constraints: $E(\sum_{i=1}^{n}\lambda_i z_i - z_o) = 0$. The estimated deviation J, where $J = Var(\sum_{i=1}^{n}\lambda_i z_i - z_o) = Var(\sum_{i=1}^{n}\lambda_i z_i) - 2Cov(\sum_{i=1}^{n}\lambda_i z_i, z_o) + Cov(z_i, z_o) = \sum_{i=1}^{n}\sum_{j=0}\lambda_i\lambda_j C_{ij} - 2\sum_{i=1}^{n}\lambda_i C_{io} + C_{oo}$, is set as the optimization target, where C_{ij} represents $Cov(z_i, z_j)$. The semivariance is $r_{ij} = \sigma^2 - C_{ij}$, and the error formula is used to obtain $J = 2\sum_{i=1}^{n}\lambda_i r_{io} - \sum_{i=1}^{n}\sum_{j=0}\lambda_i\lambda_j r_{ij} - \gamma_{oo}$. To minimize the error J, we need to find the point at which the partial derivative of the error formula with respect to λ_i is 0. Using the Lagrange multiplier method, the objective function Y is constructed as $Y = J + 2\phi(\sum_{i=1}^{n}\lambda_i - 1)$, where ϕ is the Lagrange multiplier. There is a set of parameters $\phi, \lambda_1, \lambda_2, \cdots, \lambda_n$ that minimize the error function J under the constraint of unbiased estimation, and the corresponding system of equations is presented in matrix form as Eq. (4). The inverse of this matrix is the result of kriging interpolation.

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} & 1 \\ r_{21} & r_{22} & \cdots & r_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ -\phi \end{bmatrix} = \begin{bmatrix} r_{1o} \\ r_{2o} \\ \vdots \\ r_{no} \\ 1 \end{bmatrix} \tag{4}$$

(1) Cutter wear calculation

Because cutter wear cannot be automatically recorded by TBM, we obtained cutter wear data measured by construction personnel using special calipers. However, due to the potential risks of the frequent opening of the warehouse for manual measurement, relatively few cutter wear values were manually obtained. To solve the problem of an insufficient data volume, an empirical formula was used to calculate the cutter wear for each ring [47]. This formula is shown in Eq. (5). The measured values were compared with the cutter wear calculated based on the empirical formula used in the paper, and the gap was not large. Therefore, the empirical formula was used to calculate the degree of cutter wear in this paper.

$$q = \frac{n_d \cdot F_n^4 \cdot n \cdot 2\pi r \cdot l}{k \cdot \sigma_c \cdot T^2 \cdot v} \tag{5}$$

where n_d is a dynamic factor, F_n is the gross thrust, n is the cutterhead speed, r is the cutter radius, l is the tunnel ring length, k is the empirical coefficient between friction energy and cutter wear, σ_c is the unconfined compressive strength, T is the cutterhead torque, and v is the advance speed.

(1) Data normalization

To mitigate the impact of dimensionality differences among indicators on prediction efficiency and accuracy, the data from the sample set need to be normalized and scaled to the same degree. During normalization, the data samples for different targets are scaled to the

interval [-1, 1] [48].

$$y = (y_{\max} - y_{\min}) \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} + y_{\min} \tag{6}$$

where $y_{\max}$ and $y_{\min}$ denote the normalized maximum and minimum values of the input variable x, respectively. The default values for $y_{\max}$ and $y_{\min}$ are 1 and -1, respectively. The input variable x is normalized using the corresponding maximum and minimum values ($x_{\max}$ and $x_{\min}$, respectively).

(1) Pearson correlation analysis

Before building a ML model, if there is feature redundancy in the data sample, it may adversely affect the training speed and prediction accuracy of the model. Therefore, before model construction, feature selection must be performed, and the correlations between various features must be analyzed. To measure the correlations between different variables in the raw data, the Pearson correlation coefficient is commonly used, and the corresponding formula is shown in (7).

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2}\sqrt{N\sum y_i^2 - (\sum y_i)^2}} \tag{7}$$

where x_i and y_i are the i values of X and Y, respectively; N is the number of samples; and R is the Pearson correlation coefficient. The Pearson correlation coefficient assesses the linear correlation between feature variables on a scale from -1 to 1. Values nearing -1 or 1 suggest strong correlation, whereas values near 0 imply the absence of a linear relationship between the variables.

2.2. BO-RF model development

2.2.1. Random forest

In 2001, Breiman and Cutler proposed the RF algorithm, which is based on a classification tree [24]. The RF process is as follows. First, random sampling with return is performed for the sample set, resulting in the creation of a new training set comprising an equivalent number of samples. Then, individual decision trees are associated with each new dataset, and regression trees are constructed based on these data to establish an RF model. Each tree is trained by selecting a random number of features at each node of the regression tree. Finally, the test set data are fed into the trained RF model for prediction. Each regression tree produces a predicted value, and the average of these predicted values across all the regression trees yields the final RF-predicted value. The prediction result of an RF is expressed as follows [49]:

$$f_i(x) = \frac{1}{k}\sum_{i=1}^{k} h_i(x) \tag{8}$$

where $f_i(x)$ represents the predicted value of the RF regression model and $h_i(x)$ denotes the predicted value of each regression tree.

2.2.2. Bayesian optimization algorithm

When employing machine learning to predict a target, determining the optimal algorithm hyperparameters is essential. The Bayesian optimization (BO) algorithm has emerged as the preferred choice for optimizing model hyperparameters due to its superior performance when compared with alternative global optimization algorithms [50]. BO primarily comprises two components: a probabilistic surrogate model and an acquisition function.

The probabilistic surrogate model represents the probability model of the unknown objective function and is constantly modified through an iterative increase in information to improve accuracy. Gaussian processes are commonly used in probabilistic surrogate models due to their flexibility and ability to effectively describe an unknown objective function [51].

$$f(x) \sim Gaussianprocess(\mu(x), k(x, x^{'})) \tag{9}$$

where $\mu(x)$ denotes the mathematical expectation of $f(x)$, set to 0, and $k(x, x^{'})$ represents the covariance function of x.

The acquisition function searches for the next sample point with the most potential from the candidate set according to the mean and variance of $f(x)$ by sampling in the region where the global optimal solution is most likely to occur and the regions that have not been sampled according to the posterior distribution. This paper adopts the upper confidence bound (UCB) as the chosen acquisition function, and its expression is as follows [52]:

$$x_{t+1} = \text{argmax}\mu_t(x) + \sqrt{\beta_t}\sigma_t(x) \tag{10}$$

where β_t denotes a constant that balances exploration and development.

2.2.3. Model accuracy evaluation

The predictive accuracy of the BO-RF model is evaluated through three key evaluation metrics: the root mean square error (RMSE), mean absolute error (MAE) and coefficient of determination (R^2). The RMSE and MAE gauge the disparities between the predicted and actual values, while the R^2 reflects the degree of dispersion in the sample. The three evaluation indices are calculated based on Eqs. (11)–(13) [53]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2} \tag{11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(f_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2} \tag{13}$$

where n is the sample number, f_i is the predicted value, y_i is the observed value, and $\bar{y}_i$ is the average of the actual observed value. The optimal prediction model is ascertained by choosing the model with the lowest RMSE and MAE values and the highest R^2 value. Low RMSE and MAE values, along with a high R^2, signify superior model performance.

2.3. Shapley additive interpretation

As ML algorithms become increasingly complex, the resulting black box effect results in model prediction results that lack sufficient credibility. SHAP interpretation is a method that was proposed by Lundburg and Lee and is a model-independent global interpretation method; that is, any ML model can be globally interpreted [54]. This method employs the idea of game theory to assess feature impact on model outcomes through the computation of each feature's Shapley value-based marginal contribution. This contribution can be either positive, indicating enhanced predictions, or negative, implying less effective predictions. Additionally, greater feature significance within the model corresponds to higher absolute contribution values.

Let the ith sample be x_i, the jth feature of the ith sample be x_{ij}, the predicted value of the model for the sample be y_i, the baseline of the entire model (the mean value of the prediction results when the model has no input features) be y_{base}, and $f(x_{ij})$ be the SHAP value of the jth feature of the ith sample. Then, the SHAP value is obtained with the following formula:

$$y_i = y_{base} + f(x_{i1}) + f(x_{i2}) + \cdots + f(x_{ij}) \tag{14}$$

The formula for $f(x_{ij})$ is as follows:

$$f(x_{ij}) = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [v(S \cup \{j\}) - v(S)] \tag{15}$$

where F is the complete set of all influential factors in sample x_i, S is the subset formed by any number of influential factors in sample x_i, $v(S)$ is the contribution generated by the joint action of the influential factors included in subset S, and $v(S \cup \{j\}) - v(S)$ is the contribution of influential factor j to the joint action.

The cumulative mean of the influential factors j for all samples is calculated as the SHAP value of influential factor j. The formula is as follows:

$$f(x_j) = \sum_{i=1}^{M} f(x_{ij}) \tag{16}$$

2.4. Establishment of the NSGA-III multiobjective optimization model

2.4.1. NSGA-III

Srinivas and Deb introduced NSGA as a solution to multiobjective optimization problems by enhancing the fitness value of the GA with Goldberg's concept [55]. NSGA-II, which builds on the NSGA, uses an elite retention strategy and shared parameters to enhance the optimization capability and reduce computational complexity. NSGA-II uses a crowding distance method to select the nearest-neighbor domain points around individuals and improve the population diversity at the same dominance level. Deb proposed NSGA-III by modifying the selection mechanism of NSGA-II and introducing a reference point-based selection method [56]. By utilizing distributed reference points, this approach effectively preserves population diversity in scenarios involving high-dimensional objectives. Consequently, the NSGA-III has found widespread application in various multiobjective optimization problems [57].

The multiobjective optimization based on NSGA-III involves the following steps:

Step 1: Initial population generation. Begin by randomly generating an initial population P of size N. Conduct fast nondominated sorting on this initial population, and apply genetic algorithm operations such as selection, crossover, and mutation to produce a child population Q of size N.

Step 2: Population consolidation and sequencing. Merge the parent population P and the child population Q into a new population R with a size of 2N. The new population R is classified into different levels according to dominance relationships, and individuals are subsequently selected layer by layer from different dominance levels and added to the next generation of offspring.

Step 3: Adaptive normalization. Perform adaptive normalization on individuals, and construct a set of reference points. Determine the number of reference points according to Eq. (17).

$$H = \left(\frac{M + p - 1}{p} \right) \tag{17}$$

where M is the number of optimization targets, p is the number of each dimension partition, and H is the number of reference points.

Step 4: Target value normalization. The optimization target is normalized, which helps reduce the influence of target values with different dimensions on the algorithm. To determine the desired point z in population S_t, the minimum value $z_i^{\min}$ for each target plane is first identified. Next, each target is converted to make the desired point of the converted target S_t a zero vector by subtracting the desired point $z_i^{\min}$ from the target value f_i. The converted target is then labeled $f_i'(x)$. Then, M extreme points of these targets are used to generate M-dimensional hyperplanes, and the intercept a_i of the ith axis is calculated. Finally, the objective function is normalized.

$$f_i'(x) = f_i(x) - z_i^{\min} \tag{18}$$

$$f_i^n(x) = \frac{f_i'(x)}{a_i - z_i^{\min}} \tag{19}$$

Step 5: Association between individuals and the reference point. Connect the origin and reference points to create a reference line. Calculate the vertical distance between each individual and the reference line. Select the individual with the smallest vertical distance from the reference point for association analysis. Based on the minimum number of habitats, the remaining individuals are selected, and together with the individuals previously selected by fast nondominated sorting, a new generation of the population is constructed, with the size of the new population equal to N. Repeat this process until convergence conditions are met, ultimately obtaining a Pareto optimal solution set.

The Bayesian algorithm, RF algorithm and NSGA-III are the three main algorithms of the BO-RF-NSGA-III hybrid framework, and the pseudocodes are shown in Table 1.

2.4.2. Ideal point method

The NSGA-III algorithm produces a set of Pareto optimal decisions that satisfy the requirements of multiobjective optimization. However, practical tunnel construction typically requires a single decision scheme to guide the actual tunnel excavation. To select the optimal solution from the Pareto optimal solution set, this paper adopts the ideal point method.

The ideal point method entails determining the distance between the ideal point and the optimal trade-off solutions in the Pareto front. These solutions can then be ranked based on their respective distances. The distance formula for the four-objective optimization scheme used in this study is expressed in Eq. (20) [32]:

Table 1
BO-RF-NSGA-III pseudocodes.

Algorithm 1: Bayesian algorithm
Input: X_t, N(maximum iterations)
Output: x_*
for t=1, 2,..., N do
  Find x_* by optimizing the acquisition function over the Gaussian process: $x_* = \text{argmax}_x u(x|X_{1:t-1})$
  Obtain a new sample$(x_*, f(x_*))$
  Augment the data $X_{1:t} = \{X_{1:t-1}, (x_*, f(x_*)\}$ and update the Gaussian process
end

Algorithm 2: Random forest algorithm
Input: N (training data), $n_{\min}$ (minimum node), B (number of decision trees), m (variable number)
Output: $\hat{f}_{\text{rf}}^B(x)$
for m=1, 2,..., B do
  Obtain a bootstrap sample Z of size N from the training data
  Establish a random forest tree T_b based on the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size $n_{\min}$ is reached.
  i. Select m variables at random from the p variables.
  ii. Pick the best variable/split-point among the m options.
  iii. Split the node into two daughter nodes.
  Output the ensemble of trees $\{T_b\}_1^B$.
end
$\hat{f}_{\text{rf}}^B(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x)$

Algorithm 3: NSGA-III optimization algorithm
Input: P(population size), *tmax*(maximum number of iterations), *selec*(selection probability), *crosp*(crossover probability) ad *mutp*(mutation probability)
Output: Pareto-optimal solution
for t=1, 2,..., *tmax* do
  Calculate the function value for each target in the population
  Implement three genetic operations: selection, crossover and mutation
  Produce a progeny population Q through the above operations
  Merge progeny population Q with the parent population to produce a new population$R$
  Using a series of operations, such as reference point-based selection, fast nondominated sorting, standardization, relevance analysis, and correspondence analysis, the optimal individual $P$is selected from population R
end

$$D_i = \sqrt{\left[\frac{x_i - x_{Epoint}}{x_{Epoint}}\right]^2 + \left[\frac{y_i - y_{Epoint}}{y_{Epoint}}\right]^2 + \left[\frac{z_i - z_{Epoint}}{z_{Epoint}}\right]^2} \quad (20)$$

where D_i is the distance between the average of all points and the ideal point, (x_i, y_i, z_i) are the coordinates of the optimal Pareto frontier point, $(x_{Epoint}, y_{Epoint}, z_{Epoint})$ are the coordinates of the ideal point, and the point corresponding to the smallest distance is the optimal point; that is,

$$D_{opt} = \min(D_i) \quad (21)$$

3. Case study

3.1. Project background

This study focuses on the interval tunnel between Zaojiaojing Station and Taiciqiao Station (ZT) of part 7 of the Guiyang Rail Transit Line 3 project. The project location is shown in Fig. 2. The ZT tunnel spans from YDK25+894.583~YDK27+295.528, covering a total length of 1400.945 m. The tunnel features a maximum longitudinal slope of 34.0‰, and its depth ranges between 9.58 m and 28.27 m. It predominantly traverses moderately weathered dolomite and local moderately weathered argillaceous dolomite. A geological cross-sectional profile derived from the interpolation and consolidation of borehole data is presented in Fig. 3. This study selects data from approximately 200 rings, primarily situated in moderately weathered dolomite with a compressive strength ranging from 45 to 78 MPa, a soil cohesion of around 190 kPa, and an internal friction angle of about 26°. The development of karst fissures in this region is quite heterogeneous, with some dolomite sections being separated by water and others containing water, reflecting complex hydrogeological conditions.

Considering the geological conditions of the tunnel section and the operational demands in the field, a 6470 mm-diameter EPB-TBM was chosen for tunnel construction. The cutterhead of this shield machine was equipped with 41 cutting hobs, namely, 35 single-edge hobs, 6 double-edge hobs, 12 side scrapers, 43 scrapers, and 1 superdigging knife. The cutterhead and cutter layout of the shield machine are shown in Fig. 4. However, the construction process of shield is extremely complicated, and many construction techniques are used. The stability of karst strata in Guiyang is poor, which easily results in large-scale soil subsidence, collapse and other geological disasters. Ensuring the safety of tunnel construction necessitates stringent demands for controlling surface deformation caused by tunnel excavation in karst areas. In karst areas, the cutters of shield machines are prone to severe wear. Especially in the context of cemented and broken rocks in weathered rock formations, the hob easily cuts, but because the rock blocks are glued to each other, the rock blocks are not squeezed downward but are squeezed to both sides of the blade, resulting in increased tool penetration. This further intensifies the wear of the blade, leading to faster wear of the hob blade and even sharpening. During construction, the shield machine is prone to encounter unstable strata in karst areas. To avoid safety risks such as ground collapse, the advance speed of the TBM is limited to ensure construction safety. Moreover, drainage and surrounding rock treatment measures in karst areas also affect the excavation speed.

Under karst geological conditions, surface settlement, cutter wear and advance speed are important factors in the construction process, and appropriate measures should be taken to address the corresponding challenges. Fig. 5(a) shows the internal construction of the shield tunnel, and Fig. 5(b) displays the shield machine operating system interface.

3.2. Optimization target prediction based on the BO-RF model

3.2.1. Input parameter selection and data processing

The input parameters were determined by analyzing the shield construction process and referring to many relevant studies. Reasonable control of the grouting amount and pressure is crucial for ensuring soil stability during shield machine operation. An insufficient grouting amount or pressure can cause inadequate clearance filling at the shield tail, leading to significant surface settlement. Conversely, excessive amounts of heavy metals can lead to soil crushing and slurry waste [58]. An increase in cutterhead speed can lead to heightened cutter wear and aggravate the soil disturbance [59]. In the process of excavation, the gross thrust is mainly adjusted to overcome the frictional resistance exerted by the soil and the resistance of the cutterhead. An increase in gross thrust corresponds to greater disturbance to the soil and increased cutter wear consumption [60]. The cutterhead torque is used to ensure normal advancement of the TBM and is one of the key shield construction parameters [61]. Maintaining an equilibrium between the pressure within the soil bin and the cutting soil pressure on the tunnel face is crucial for minimizing surface settlement. Thus, the chamber earth pressure is a critical factor for controlling surface settlement [62]. In EPB shield tunnelling, the excavated soil and water mixture, commonly known as shield muck, serves as the supporting medium in the excavation chamber. The muck is then discharged from the chamber through a screw conveyor. It is essential to achieve a balance between the discharged and excavated soil volumes to establish a stable support pressure and ensure the stability of the tunnel face [63]. Foam agents are widely used soil conditioners that can protect the cutter and reduce the cutterhead torque by adjusting the soil fluidity [64].

The tunnel burial depth significantly affects surface settlement during shield construction. A deeper tunnel generally corresponds to reduced vertical surface settlement and an increased for the width of the settlement tank [65]. The cover-span ratio, closely tied to tunnel depth, is a key variable in estimating surface settlement [66]. The internal

Fig. 2. Plan location of section 7 of Guiyang rail transit line 3.

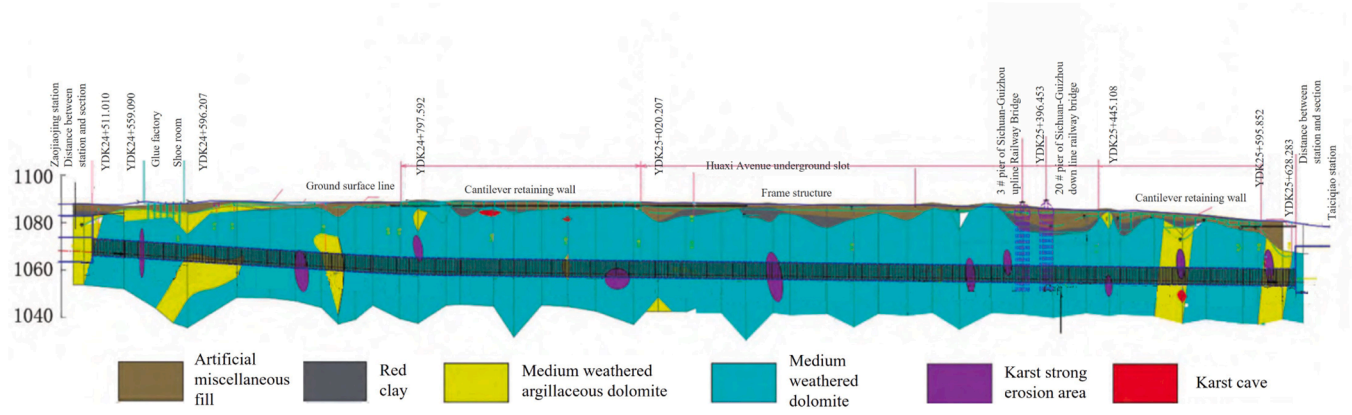


Fig. 3. Geological profile map from Zaojiao well station to Taiziqiao station.

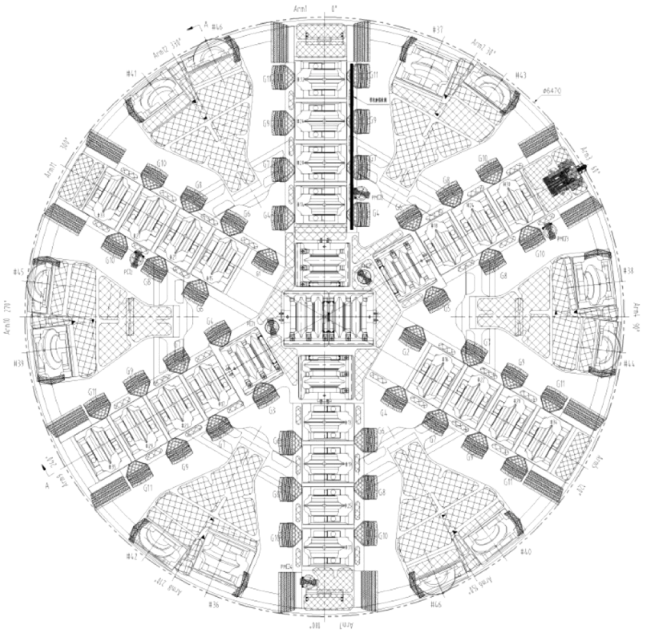


Fig. 4. Cutterhead front view.

friction angle and soil cohesion are crucial in influencing various aspects of shield construction. Increased values of these parameters translate to heightened soil-to-soil and soil-to-equipment friction, resulting in greater shield thrust and cutterhead torque. This, in turn, leads to slower advancement, increased excavation difficulty, accelerated cutterhead wear, and a decrease in surface settlement [67].

Based on the above mechanism analysis, an input index system related to surface settlement, cutter wear and advance speed is proposed based on the grouting amount (x_1), cutterhead speed (x_2), gross thrust (x_3), grouting pressure (x_4), cutterhead torque (x_5), chamber earth pressure (x_6), shield muck (x_7), foam agent (x_8), burial depth (x_9), cover-span ratio (x_{10}), internal friction angle (x_{11}), and soil cohesion (x_{12}). The grouting amount (x_1), cutterhead speed (x_2), gross thrust (x_3), grouting pressure (x_4), chamber earth pressure (x_6) and foam agent (x_8) are the active control parameters of the TBM and are adjusted by the operator according to the engineering requirements to directly adjust the operation of the TBM. The cutterhead torque (x_5) and shield muck (x_7) are passive control parameters of the TBM. These parameters are affected by the mechanical characteristics and working conditions of the shield machine and cannot be directly controlled by the operator; however, these parameters affect the adjustment of active control parameters and the tunnel driving effect. Table 2 shows the statistical information for the input and output parameters of the preprocessed 200-421 ring, and the corresponding distribution diagram is depicted in Fig. 6.

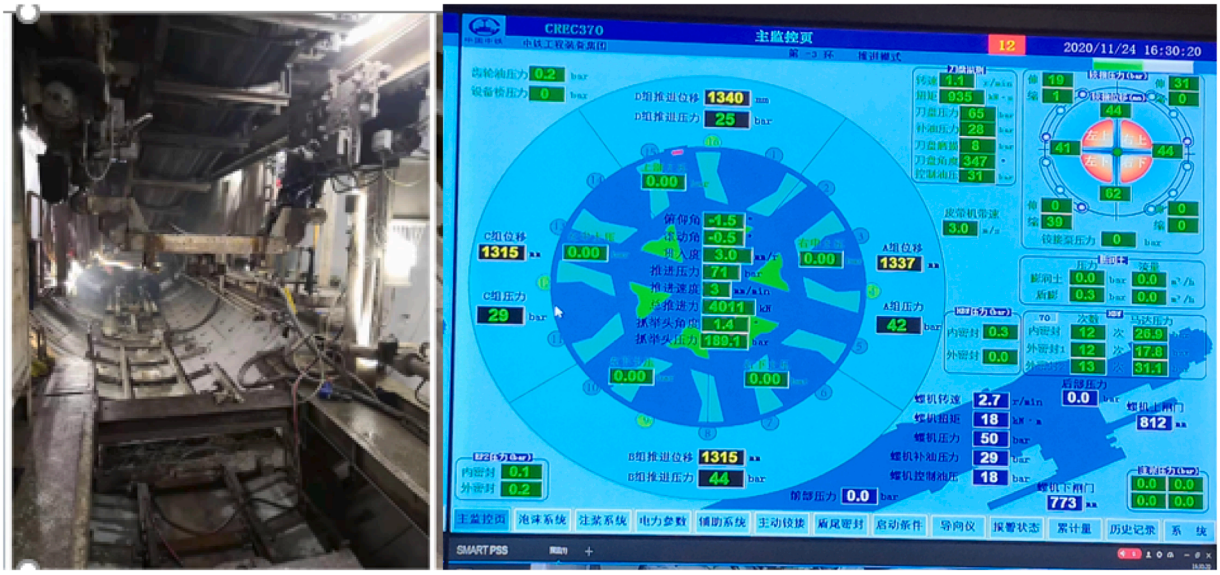
After the obtained data are preprocessed, the Pearson correlation coefficient is employed to analyze the relationship between the parameters, ensuring the dataset's quality. The correlation coefficient matrix, depicted in Fig. 7, reveals that the absolute values of the Pearson correlation coefficients between the 12 input parameters and surface settlement, cutter wear, and advance speed are all less than 0.5. This suggests the absence of obvious linear relationships between the input parameters and these variables, reinforcing that a mathematical regression model is unsuitable for predicting surface settlement, cutter wear, or advance speed. The RF algorithm accurately captures nonlinear mapping relationships between input variables and control targets to obtain prediction results. In addition, the absolute value of the Pearson correlation coefficient for the 12 input parameters is not greater than 0.5, indicating that the correlation among the input parameters is not strong, the parameters are relatively independent, and the degree of redundancy of the input parameters is weak.

3.2.2. *Model hyperparameter optimization*

The RF prediction model requires adjustments to three primary hyperparameters: the number of decision trees (n_estimators), the maximum depth of the decision trees (max_depth) and the number of random features (max_features) [68]. While high hyperparameter values generally improve model accuracy, there is a heightened risk of overfitting. Therefore, it is essential to carefully tune the hyperparameters to optimize the predictive performance of the model.

Hyperparameter optimization in ML often involves the use of methods such as grid search, random search, and Bayesian optimization algorithms. However, grid search method becomes computationally intensive when dealing with numerous parameters since it systematically explores all hyperparameter combinations, increasing susceptibility to dimensional constraints. Because the random search method seeks the optimal hyperparameter combination in an infinite search space, each search iteration is independent of the others, resulting in a high variance in results. In this paper, Bayesian optimization is employed to optimize the hyperparameters of the RF algorithm, which can solve complex parameter search problems with only a small number of iterations and obtain optimal hyperparameter values. To fully evaluate the model performance and reduce fluctuations due to unreasonable data partitioning, a quintuple cross-validation method is adopted. The entire dataset is divided into five parts, four of which are used for training the model and one for model evaluation. This process is iterated five times to ensure the inclusion of all data in the model evaluation, providing a comprehensive evaluation of the proposed model. The mean squared error (MSE) serves as the loss function during model training.

To obtain the best prediction result, max_features is set to auto by



(a) (b)

Fig. 5. Shield construction of Metro Line 3: (a) the internal construction of the shield tunnel; (b) the shield machine operating system interface.

Table 2
Statistics for the input and output parameters.

Parameter type	Variable	Data			Unit
		Min	Max	Ave	
Input	Grouting amount (x_1)	3.4	11.6	5.7	m^3
	Cutterhead speed (x_2)	1.2	2.4	1.7	$rad \cdot min^{-1}$
	Gross thrust (x_3)	7000	16000	11548	kN
	Grouting pressure (x_4)	0.92	2.46	1.90	bar
	Cutterhead torque (x_5)	1218	3150	2091	kN·m
	Chamber earth pressure (x_6)	0.6	1.8	1.2	bar
	Shield muck (x_7)	45	82	54	m^3
	Foam agent (x_8)	22.22	44.44	33.36	L
	Burial depth (x_9)	15.58	18.27	16.92	m
	Cover-span ratio (x_{10})	2.41	2.82	2.62	-
	Internal friction angle (x_{11})	20	33	26	°
	Soil cohesion (x_{12})	150	230	190	kPa
Output	Surface settlement (f_1)	-32.55	-0.22	-3.49	mm
	Cutter wear (f_2)	0.13	0.23	0.17	mm
	Advance speed (f_3)	12.96	28.17	22.48	$mm \cdot min^{-1}$

default, max_depth and n_estimators are optimized and adjusted, and the evaluation index of training performance is the MSE. max_depth and n_estimators are set to initial ranges of [7,10] and [0, 100], respectively. The optimization process of max_depth and n_estimators is illustrated in Fig. 8, and the hyperparameter optimization results are presented in Table 3.

3.2.3. Analysis of the predicted results

Among the 221 ring data obtained from monitoring, 177 ring data are selected as the training set for building the prediction model for surface settlement, cutter wear, and advance speed. The remaining 44 ring data constitute the test set to validate the constructed model. Evaluation metrics including R^2, RMSE, and MAE are computed for the predicted results using Eqs. (11)–(13), and the results are visually presented in Fig. 9, with a detailed summary of prediction accuracy provided in Table 4.

As depicted in Fig. 9(a), the BO-RF model effectively captures the intricate relationship between the input parameters and surface settlement within the training sample set. The surface settlement results predicted using the training set closely align with the actual observed values. Following the training with the RF algorithm, the model is applied to make predictions using the test set. The results in Fig. 9(a) provide a more intuitive representation, demonstrating the good fitting effect of the RF prediction model for surface settlement on the test set data. The model exhibits a strong capacity to accurately predict values.

The outcomes presented in Table 4 demonstrate that the RF algorithm-based models for predicting surface settlement, cutter wear, and advance speed yield highly effective prediction results. Cheng et al. [69] employed an RF model for predicting surface settlement caused by shield construction and achieved good prediction performance. Wu et al. [70] showed that the RF algorithm is an accurate predictor of cutter wear through a comparison of prediction performance. Zhou et al. [71] utilized the RF algorithm to accurately predict the advance speed of TBM under hard rock conditions.

To assess its accuracy and validity, the proposed RF prediction model is compared with alternative models, including BPNN, SVM, GBDT, and LSTM models, as summarized in Table 5. Notably, the LSTM prediction model demonstrates the highest accuracy in this study, with the RF prediction model following closely. However, the LSTM model takes twice as long as the RF model to obtain a prediction result, and the RF algorithm has an advantage in terms of computational efficiency. Therefore, this paper adopts the BO-RF algorithm to predict the surface settlement, cutter wear and advance speed and introduces the trained regression function as the optimized fitness function. Chen et al. [26] showcased the superior performance of the RF algorithm compared to BPNN and SVM in predicting surface settlement using six ML algorithms. Tang et al. [72] demonstrated the efficiency of using an RF model with suitable parameters for quickly and effectively predicting surface settlement.

3.3. Model interpretability analysis based on SHAP

Given the inherent correlations among shield parameters, it is impractical to qualitatively analyze the influence of a single parameter on surface settlement, cutter wear, and advance speed. To address this issue, the SHAP method is applied for a comprehensive analysis, offering insights into the contribution of each construction parameter to the

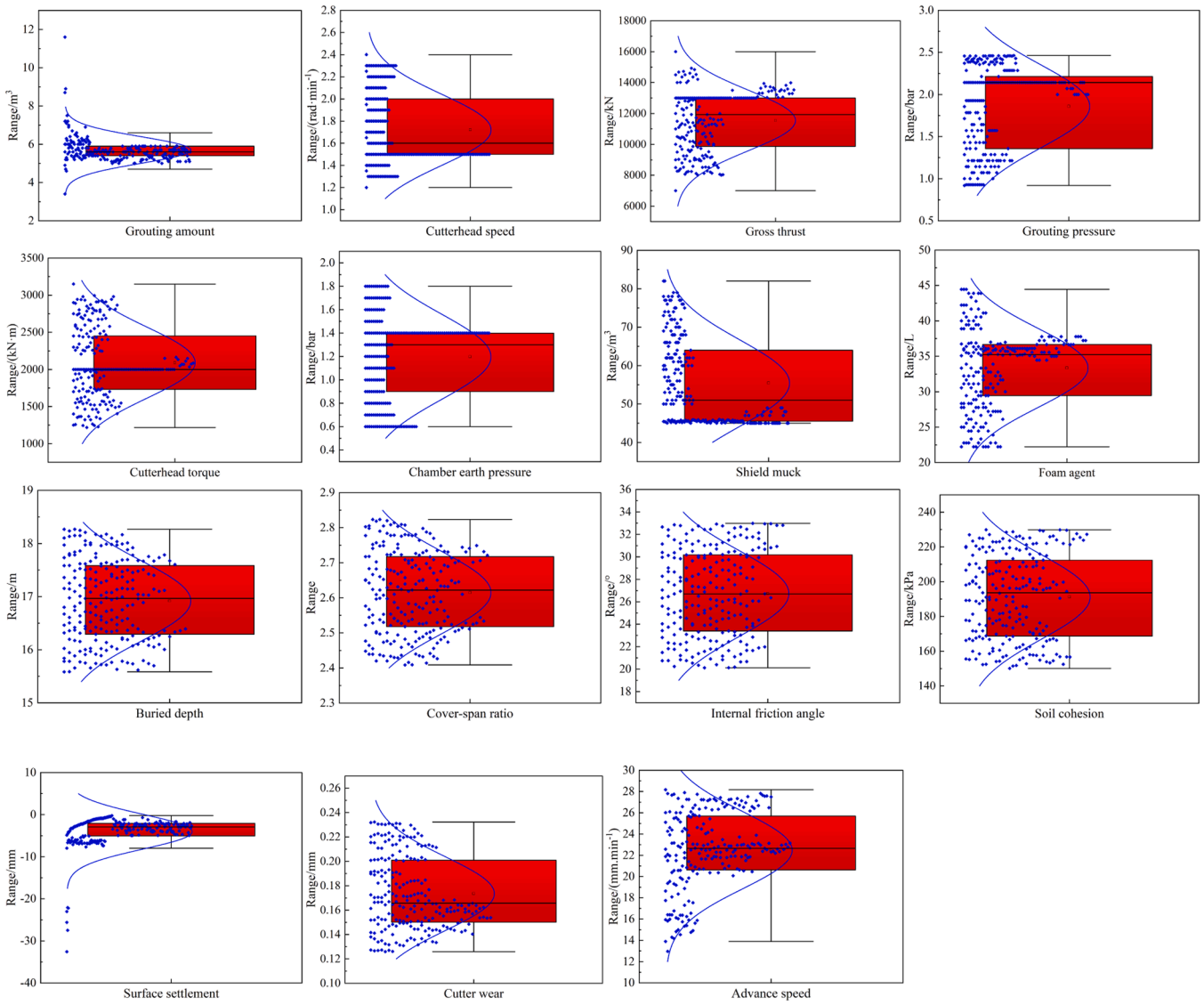


Fig. 6. The distribution of each index.

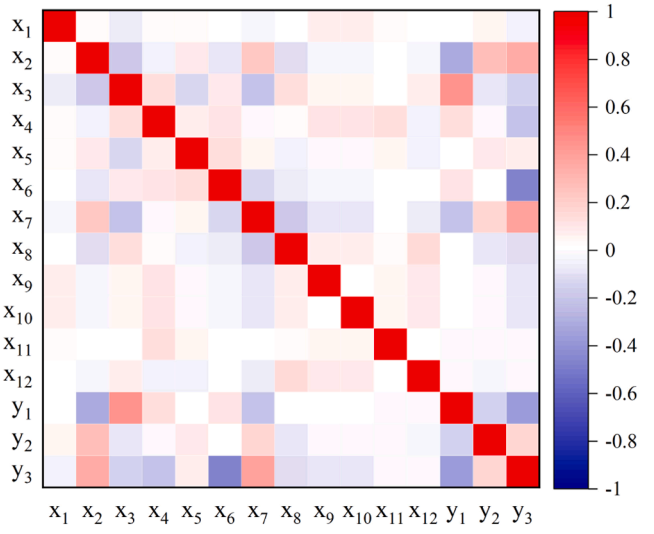


Fig. 7. Matrix of correlation coefficients between different pairs of parameters.

predicted results. This analysis encompasses both positive and negative contributions, enabling the identification of important construction parameters that significantly impact the control objectives. The SHAP analysis results of the input parameters for surface settlement, cutter wear and advance speed are presented in Fig. 10. These findings serve as a valuable foundation for decision-making in shield construction management and control, providing guidance on the key parameters that warrant close attention.

Fig. 10 illustrates the order of influence of the input parameters on surface settlement as follows: chamber earth pressure, gross thrust, grouting amount, cutterhead speed, grouting pressure, shield muck, cutterhead torque, burial depth, cover-span ratio, foam agent, soil cohesion, and internal friction angle. The order of influence of the input parameters on cutter wear is as follows: chamber earth pressure, cutterhead speed, cutterhead torque, gross thrust, foam agent, shield muck, internal friction angle, grouting amount, soil cohesion, grouting pressure, burial depth, and cover-span ratio. The order of influence of the input parameters on advance speed is as follows: chamber earth pressure, gross thrust, shield muck, cutterhead speed, foam agent, grouting pressure, soil cohesion, cutterhead torque, grouting amount, internal friction angle, burial depth and cover-span ratio. The results show the following:

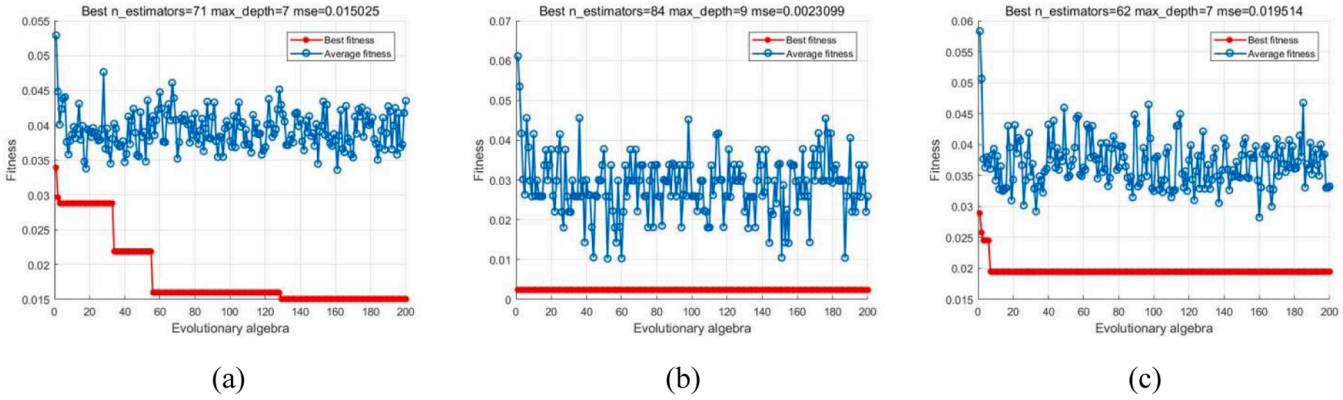


Fig. 8. The results of the RF algorithm: (a) surface settlement; (b) cutter wear; (c) advance speed.

Table 3
Results of the Bayesian optimization of RF hyperparameters.

Optimization objective	Hyperparameters	Optimal range	Optimized result	MSE
Surface settlement	n_estimators	(0,100)	71	0.0150
	max_depth	(7,10)	7	
Cutter wear	n_estimators	(0,100)	84	0.0023
	max_depth	(7,10)	9	
Advance speed	n_estimators	(0,100)	62	0.0195
	max_depth	(7,10)	7	

(1) The chamber earth pressure is the main shield construction parameter that greatly impacts the surface settlement, cutter wear and advance speed. Excessive pressure in the soil chamber can cause the soil in front of the TBM to be squeezed, resulting in surface uplift, while insufficient pressure can lead to the soil moving toward the soil chamber, causing surface settlement. Both scenarios can cause different degrees of stratum deformation. Ling et al. [14] emphasized the significance of chamber earth pressure as a key factor influencing surface settlement. Zhang et al. [73] emphasized the high relative importance of chamber earth pressure in relation to surface settlement.

(2) Shield construction parameters such as the gross thrust, cutterhead speed, cutterhead torque, grouting amount, grouting pressure and shield muck amount have similar effects on the three targets. The gross thrust serves as the driving force for TBM advancement. As the gross thrust increases, the disturbance to the soil also intensifies, leading to increased surface settlement

Table 4
Target prediction accuracy.

Forecast target	Dataset type	Prediction accuracy		
		R^2	RMSE	MAE
Surface settlement	Training set	0.947	0.162	0.131
	Test set	0.930	0.172	0.138
Cutter wear	Training set	0.963	0.111	0.081
	Test set	0.931	0.153	0.114
Advance speed	Training set	0.959	0.122	0.103
	Test set	0.938	0.138	0.112

Table 5
Prediction performance of different models.

Forecast target	Prediction accuracy	Prediction model				
		BO-BPNN	BO-SVM	BO-GBDT	BO-RF	BO-LSTM
Surface settlement	R^2	0.845	0.902	0.872	0.930	0.961
	RMSE	0.275	0.237	0.206	0.172	0.143
	MAE	0.234	0.203	0.173	0.138	0.116
Cutter wear	R^2	0.901	0.943	0.927	0.931	0.970
	RMSE	0.175	0.153	0.150	0.153	0.102
	MAE	0.142	0.132	0.126	0.114	0.062
Advance speed	R^2	0.862	0.896	0.910	0.938	0.967
	RMSE	0.220	0.211	0.205	0.138	0.136
	MAE	0.171	0.169	0.161	0.112	0.111

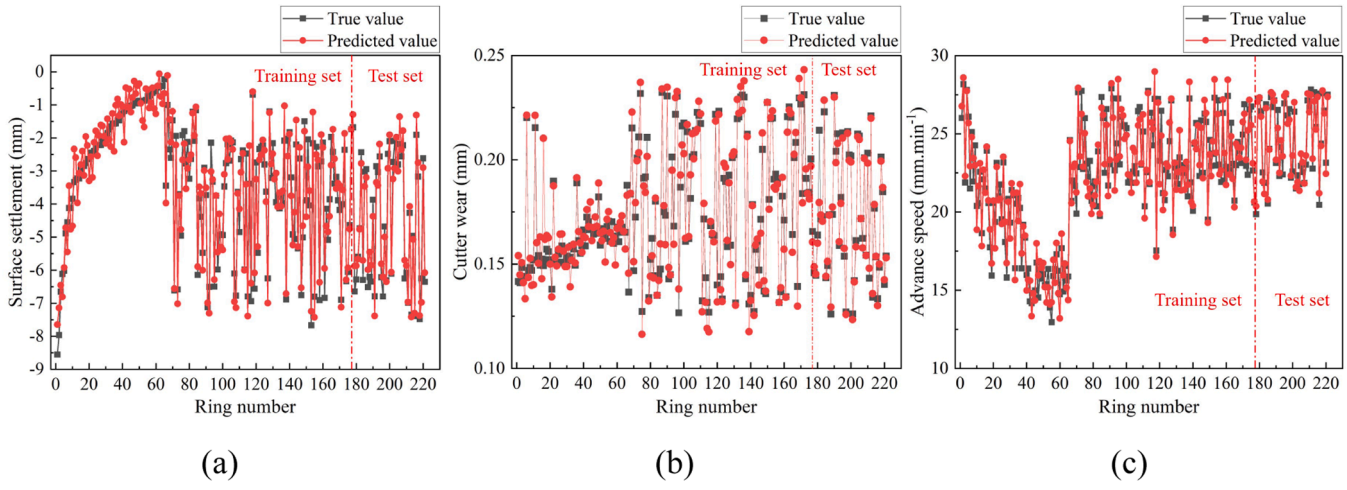


Fig. 9. Prediction results: (a) surface settlement; (b) cutter wear; (c) advance speed.

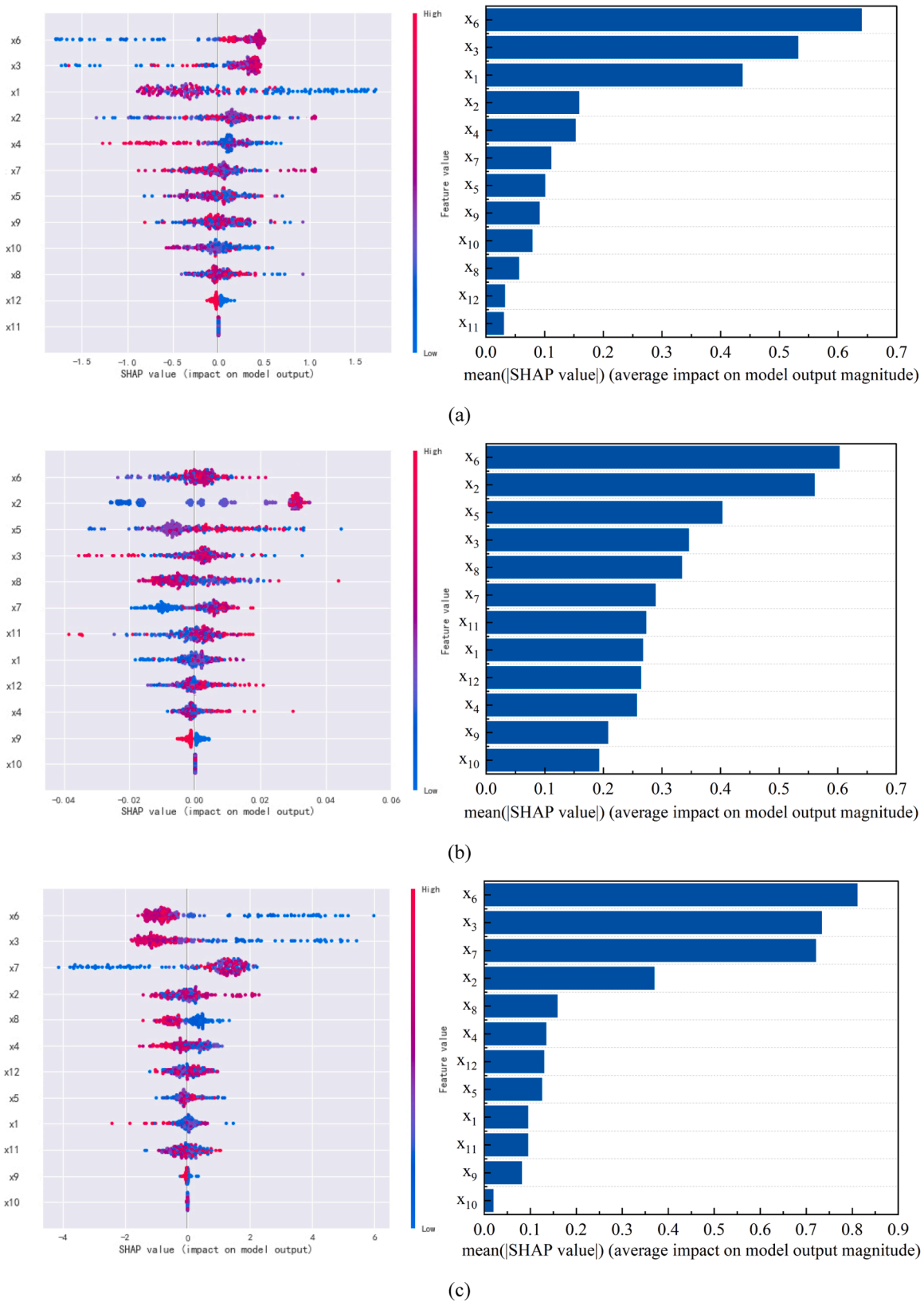


Fig. 10. SHAP analysis results for each input parameter for the control targets: (a) surface settlement; (b) cutter wear; (c) advance speed.

during construction [74]. Increasing the cutterhead speed can improve the advance speed, reduce the disturbance time of the soil layer, and alleviate the surface settlement caused by the compression of the underlying soil layer by the TBM [27]. A proper cutterhead torque can mitigate the disturbance caused by cutterhead rotation to the surrounding soil layer and the circumferential friction between the shield shell and the soil layer, thereby decreasing surface settlement [75]. When the

grouting amount is less than the building gap or the grouting pressure is less than the initial earth pressure, the surface above the shield tail settles. When the grouting amount is greater than the building gap or the grouting pressure is greater than the initial earth pressure, the soil mass will be destroyed by passive compression, and the surface above the shield tail will be uplifted [76,77]. The amount of shield muck affects the chamber pressure and ground behavior during excavation. A small amount of shield muck can lead to greater chamber pressure and ground heave, while a large amount of shield muck can result in low chamber pressure and surface settlement due to excessive ground loss at the excavated surface [78].

(3) The influence of construction parameters (x_1-x_8) on the three targets is greater than that of geometric parameters (x_9-x_{10}) and geological parameters (x_{11}-x_{12}). Fig. 10 illustrates that the effects of the tunnel burial depth, span ratio, internal friction angle and soil cohesion are relatively small because of the relatively concentrated distribution of these factors in the sample dataset. Although geological parameters in actual engineering have a considerable impact on surface settlement, cutter wear and advance speed, the shield tunneling stratum in this study is homogeneous, and the soil around the shield is relatively simple, with very small differences in various properties; thus, the impact on surface settlement, cutter wear and advance speed is relatively small.

3.4. Multiobjective optimization of shield construction parameters based on BO-RF-NSGA-III

In actual engineering, only a few shield construction parameters can be adjusted, and many resources are required to adjust all the parameters, which is not cost effective. Therefore, based on model interpretability analysis and the relevant literature [14,45], the three most influential shield construction parameters, namely, cutterhead speed x_2, gross thrust x_3 and chamber earth pressure x_6, are chosen for adjustment to minimize surface settlement and cutter wear and maximize advance speed. These parameters are explored in seven different scenarios, as presented in Table 6. The other parameters are set according to the average values. In scenarios 1-3, only one shield construction parameter, x_2, x_3 or x_6, is adjusted. In scenarios 4-6, two shield construction parameters are adjusted: x_2 and x_3 in scenario 4, x_2 and x_6 in scenario 5, and x_3 and x_6 in scenario 6. In scenario 7, three shield construction parameters are adjusted.

3.4.1. Determination of the objective function and constraints

Before multiobjective optimization, the relationships between the shield construction parameters and surface settlement, cutter wear, and advance speed can be fitted via RF regression prediction. The resulting regression function can subsequently serve as the fitness function for multiobjective optimization. Taking scenario 7 as an illustration, the objective functions for surface settlement, cutter wear, and advance speed, derived from the RF algorithm, are expressed in Eq. (22).

$$\begin{cases} \min f_1 = \min[RF(x_2, x_3, x_6)] \\ \min f_2 = \min[RF(x_2, x_3, x_6)] \\ \max f_3 = \max[RF(x_2, x_3, x_6)] \end{cases} \tag{22}$$

where f_1, f_2 and f_3 are fitness functions of surface settlement, cutter wear and advance speed, respectively, and x_2, x_3 and x_6 are the cutterhead speed, gross thrust and chamber earth pressure, respectively.

To ensure that the shield construction parameters generated during optimization are practical, it is important to establish reasonable limits for each decision variable based on actual project conditions and relevant specifications. The initial decision variable ranges in this paper are determined based on reasonable parameter ranges for the specific shield type as the main reference, as shown in Table 7.

3.4.2. Obtaining a Pareto front

Prior to the multiobjective optimization of BO-RF-NSGA-III, the model parameters are set. With reference to previous literature [79], the parameters of the NSGA-III are set. The GA target number is set to 3, with a population size of 100. The crossover operator is set to 0.7, the mutation operator is set to 0.01, and the maximum number of iterations is set to 60. Once the parameters are set, the NSGA-III algorithm is executed to obtain the Pareto front solution set.

Figs. 11–17 show the optimization results for scenarios 1-7. With Fig. 11 as an example, Fig. 11(a) depicts the Pareto frontier of the target, and Fig. 11(b) showcases the improvement percentage of the Pareto solution set. A comparison of the improvement degrees of the three targets in the different scenarios is presented in Fig. 18. The optimization results for surface settlement, cutter wear and advance speed in the different scenarios are shown in Table 8.

3.4.3. Analysis of results

According to the multiobjective optimization results, multiple shield construction parameter optimization schemes can be obtained based on NSGA-III. In practice, only one optimized scheme is necessary to provide guidance for field construction. To achieve the best optimization effect, the ideal point method is employed to identify the optimal decision from the Pareto solution set. With scenario 1 as an example, Fig. 11 shows that the ideal point coordinates formed when the surface settlement, cutter wear and advance speed reach the optimal values are E (-2.93, 0.144, 26.54). These coordinates are subsequently applied in Eq. (20) to calculate the distance between each Pareto solution and the ideal point. The selection of the optimal decision in the multiobjective optimization of shield construction parameters is based on the point with the smallest distance from the target, as determined by Eq. (21). The results showed that the surface settlement is 3.03 mm, the cutter wear is 0.166 mm, and the advance speed is 26.54 mm·min^{-1} at this point. The average improvement percentage is 11.74 % when compared to the original data sample average. The optimal decisions under the seven scenarios are shown in Table 8.

(1) This method effectively addresses multiobjective problems by optimizing and controlling surface settlement, cutter wear, and advance speed in the seven scenarios. The method consistently constructs the Pareto front and yields optimal decisions for each scenario, as shown in Figs. 11 to 17. As detailed in Table 8, compared with the average for the original samples, the average improvement percentage in the seven scenarios is 12.56 %. A case study further verifies these multiobjective optimization results. This demonstrates that the solution derived from the BO-RF-NSGA-III framework successfully achieves the triple objective of reducing surface settlement and cutter wear and increasing the advance speed.

Table 6
Parameters adjusted in each scenario.

Scenarios	Adjustment parameter	Scenarios	Adjustment parameter
Scenario 1	x_2	Scenario 5	x_2,x_6
Scenario 2	x_3	Scenario 6	x_3,x_6
Scenario 3	x_6	Scenario 7	x_2,x_3,x_6
Scenario 4	x_2,x_3		

Table 7
Input parameter ranges.

Shield construction parameter	Parameter range
Cutterhead speed (x_2)/(rad·min^{-1})	$1 \le x_3 \le 3$
Gross thrust (x_3)/kN	$7000 \le x_3 \le 16000$
Chamber earth pressure (x_6)/bar	$0 \le x_6 \le 2$

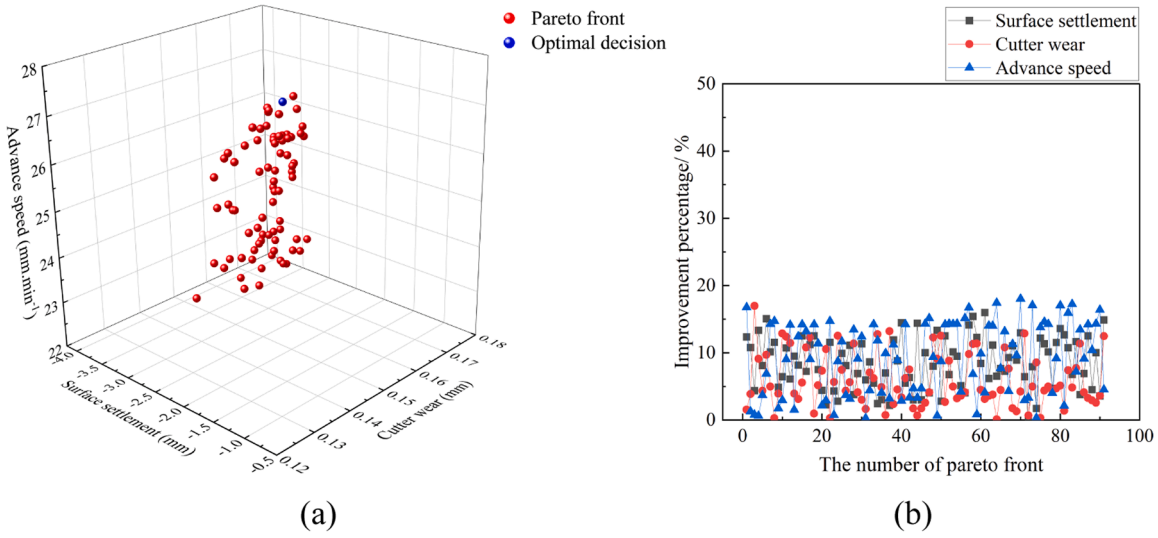


Fig. 11. Optimization results in scenario 1: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

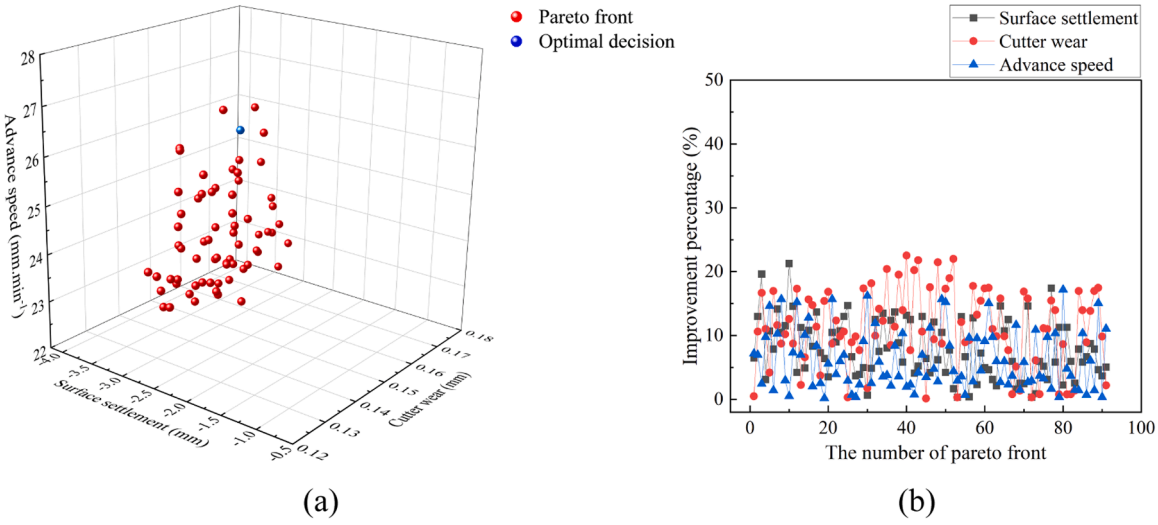


Fig. 12. Optimization results in scenario 2: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

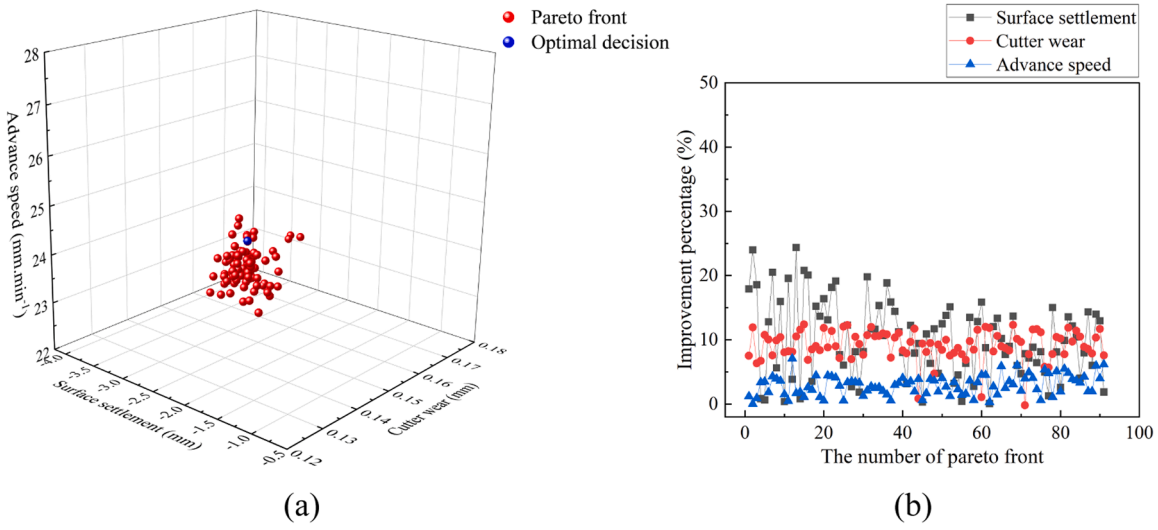


Fig. 13. Optimization results in scenario 3: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

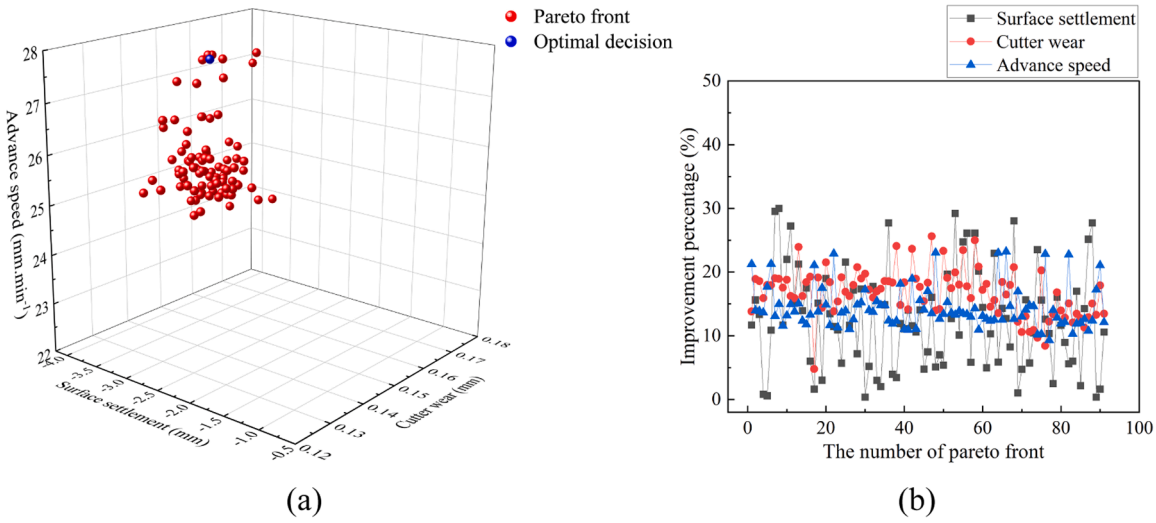


Fig. 14. Optimization results in scenario 4: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

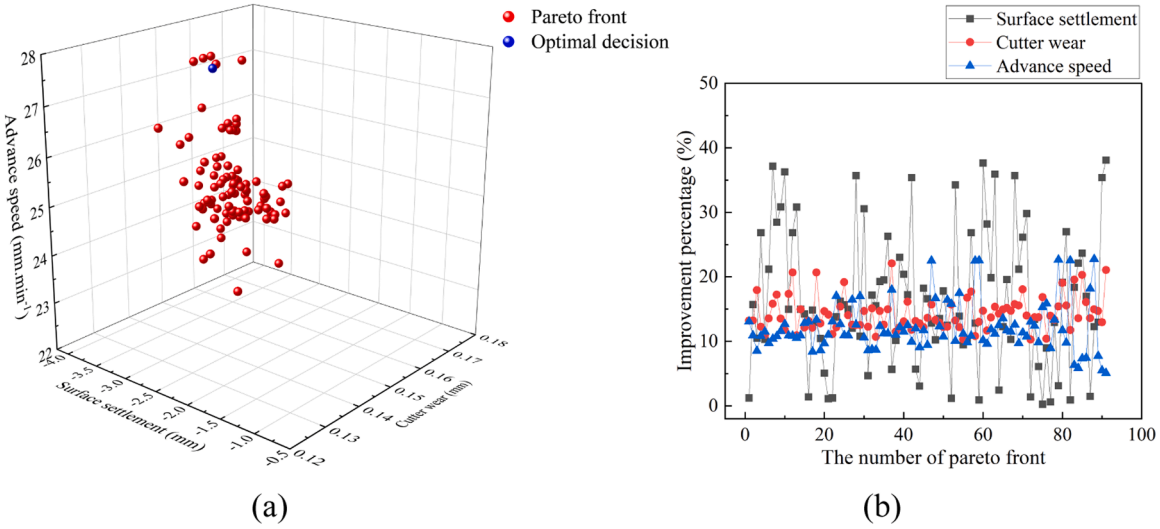


Fig. 15. Optimization results in scenario 5: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

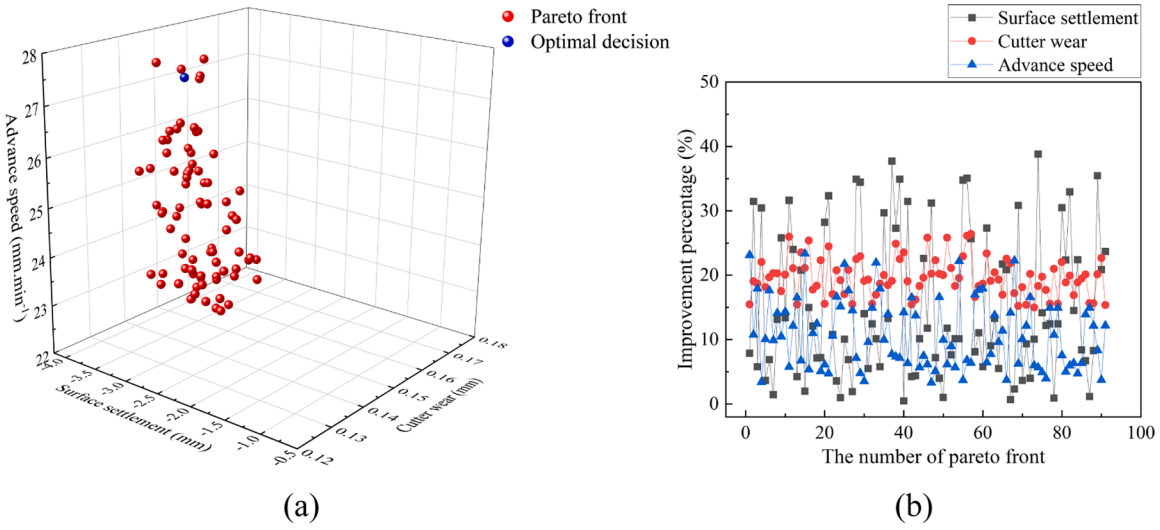


Fig. 16. Optimization results in scenario 6: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

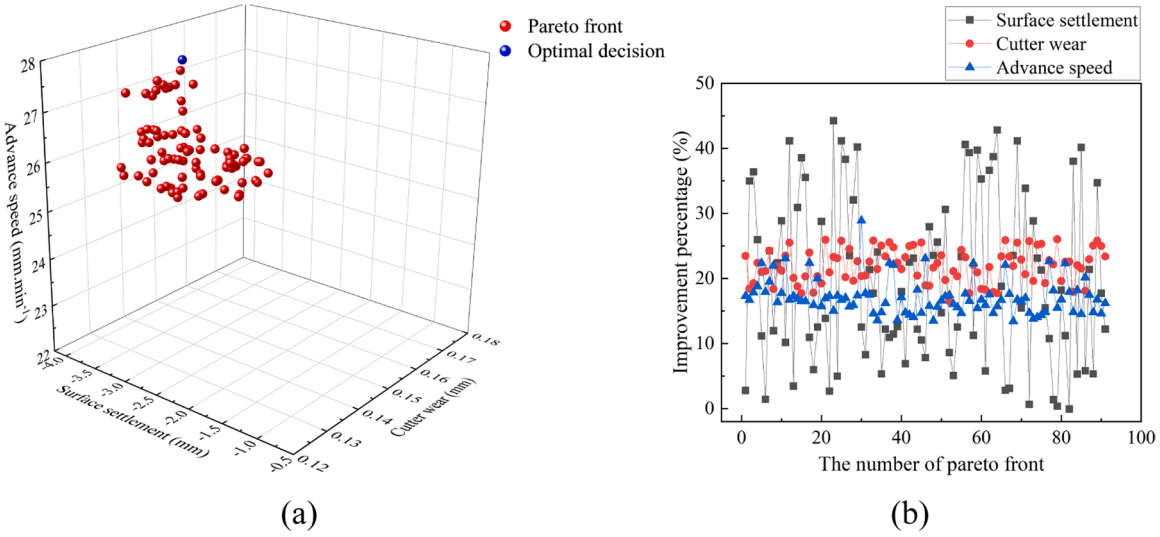


Fig. 17. Optimization results in scenario 7: (a) Pareto frontier; (b) improvement percentage of the Pareto solution set.

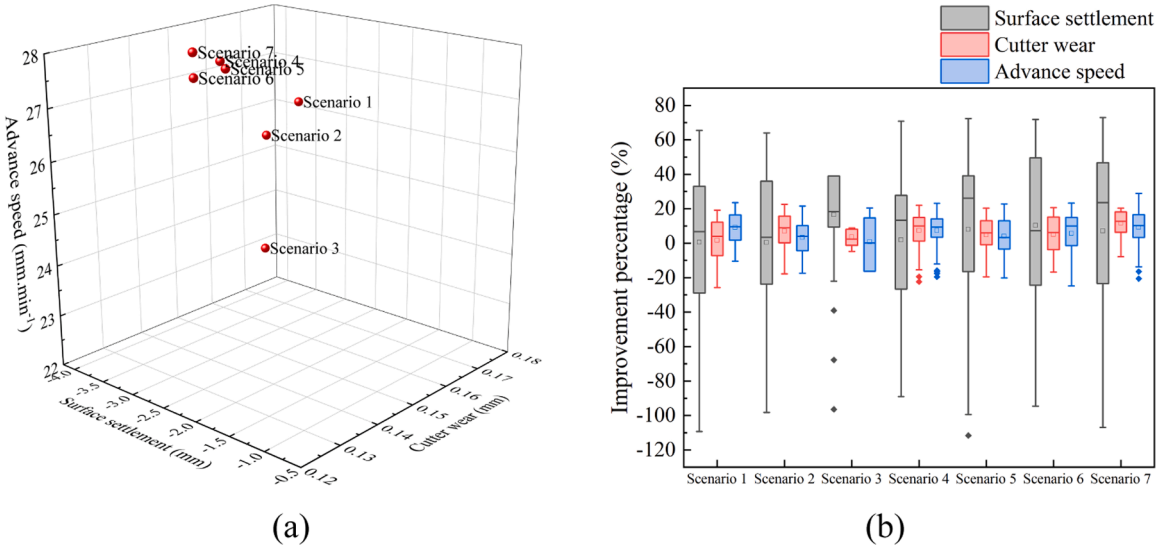


Fig. 18. Comparison of the optimization results: (a) comparison of the optimal decisions in the seven scenarios; (b) comparison of the improvement percentages in the seven scenarios.

Table 8
The degree of improvement of the three objectives in different scenarios.

Optimization objectives		Adjusted parameter								Average
		Original	x_2	x_3	x_6	x_2+x_3	x_2+x_6	x_3+x_6	$x_2+x_3+x_6$	
f_1 /mm	Mean	-3.49	-3.19 (8.48 %)	-3.21 (7.84 %)	-3.13 (10.12 %)	-3.05 (12.58 %)	-2.91 (16.54 %)	-2.96 (15.21 %)	-2.79 (19.89 %)	-3.03 (12.95 %)
f_2 /mm	Mean	0.174	0.163 (5.92 %)	0.154 (11.06 %)	0.158 (9.12 %)	0.144 (16.73 %)	0.149 (14.29 %)	0.139 (19.62 %)	0.135 (21.97 %)	0.149 (14.32 %)
f_3 /(mm·min^{-1})	Mean	22.48	24.56 (9.24 %)	23.89 (6.23 %)	23.15 (2.94 %)	25.73 (14.44 %)	25.22 (12.16 %)	24.90 (10.74 %)	26.35 (17.70 %)	24.83 (10.42 %)
Average improvement percentage		-	7.88 %	8.38 %	7.40 %	14.58 %	14.33 %	15.19 %	19.67 %	12.56 %
f_1 /(mm)	Optimal solution	-3.49	-3.03 (12.97 %)	-3.12 (10.48 %)	-3.00 (13.99 %)	-3.04 (12.71 %)	-3.04 (12.80 %)	-3.13 (10.12 %)	-3.05 (12.54 %)	-3.05 (12.23 %)
f_2 /(mm)	Optimal solution	0.174	0.166 (4.24 %)	0.158 (8.71 %)	0.156 (10.34 %)	0.147 (15.28 %)	0.130 (24.93 %)	0.140 (19.54 %)	0.138 (20.34 %)	0.148 (14.77 %)
f_3 /(mm·min^{-1})	Optimal solution	22.48	26.54 (18.02 %)	26.01 (15.67 %)	23.83 (5.98 %)	27.71 (23.22 %)	27.54 (22.50 %)	27.47 (22.18 %)	27.98 (24.45 %)	26.73 (18.86 %)
Average improvement percentage		-	11.74 %	11.62 %	10.10 %	17.07 %	20.08 %	17.28 %	19.11 %	15.29 %

Note: The values in brackets are the improvement percentages of the optimized values compared with the means of the original data.

(2) As more shield construction parameters are adjusted, the optimization control of surface settlement, cutter wear and driving speed improve. As shown in Table 8, in scenarios 1-3, when only one shield construction parameter is adjusted, the average improvement percentage is 7.89 %. When two shield construction parameters are adjusted in scenarios 4-6, the average improvement percentage is 14.70 %. Adjusting the three shield construction parameters simultaneously in scenario 7 yields the highest average improvement of 19.67 %. Fig. 18(a) shows that the ground settlement and cutter wear of the optimal Pareto solution generated in scenario 7 are lower than those generated in scenarios 1-6, and the tunneling speed of the optimal Pareto solution generated in scenario 7 is greater than that of the optimal Pareto solution generated in scenarios 1-6. The results show that adjusting three shield construction parameters yields the best optimization effect. Therefore, identifying and adjusting additional shield construction parameters can enhance the optimization effect. Guo et al. [35] showed that by considering more parameters and constraints, a multiobjective optimization problem can be further optimized. Zhang et al. [34] showed that optimizing more objectives and new variables will enhance the optimization effect. However, it takes more time to adjust more shield construction parameters.

(3) Conflicts exist among the three objectives, and there is a complex relationship between the shield construction parameters and the optimization objectives. As shown in Fig. 15, the Pareto frontier solution is widely distributed, covering the two extremes of goals f_1, f_2 and f_3. The optimization of one goal needs to be achieved at the expense of the other two goals. Conflicts exist among the three goals, and it is difficult to achieve simultaneous optimization of multiple goals. In addition, different shield construction parameters can improve different optimization objectives to different degrees. For target f_1, the average percentage improvement is 14.37 % for the scenario considering parameter x_2, 13.88 % for the scenario considering parameter x_3, and 15.44 % for the scenario considering parameter x_6. For target f_2, the average percentage improvement is 14.73 % for the scenario considering parameter x_2, 17.35 % for the scenario considering parameter x_3, and 16.25 % for the scenario considering parameter x_6. For target f_3, the average percentage improvement is 13.24 % for the scenario considering parameter x_2, 12.13 % for the scenario considering parameter x_3, and 10.74 % for the scenario considering parameter x_6. By comparison, parameter x_6 has a considerable influence on the optimization improvement of target f_1, parameter x_3 has a considerable influence on the optimization improvement of target f_2, and parameter x_2 has a considerable influence on the optimization improvement of target f_3. Preferentially adjusting the shield construction parameters associated with the corresponding targets can improve the efficiency and effectiveness of modeling, especially if the control requirement for a certain target is high in practical engineering.

4. Discussion

The case study strongly supports the effectiveness of this method for minimizing surface settlement and cutter wear and maximizing advance speed by optimizing and adjusting shield construction parameters. In practice, however, more than three objectives may need to be considered simultaneously. In this section, we introduce a fourth objective to assess the scalability of the proposed method.

The fourth objective of optimization is to minimize the energy consumption of the shield machine drive system. Incorrect estimations of energy consumption may lead to issues such as cutter wear and the fracture of the shield machine, and damage to the electromechanical machine. Consequently, the energy consumption of the shield machine drive system serves as a crucial indicator for assessing the cost and efficiency of shield construction [80]. The same steps as in the method developed in Section 4 are followed for optimization. Fig. 19 shows the percentage improvement in the optimization of the Pareto solution set under the different scenarios. Fig. 20 provides a comparison of the results for different scenarios, and Fig. 21 illustrates a comparison of the optimization results considering three or four objectives. The optimization results of each scenario are presented in Table 9.

(1) This method demonstrates exceptional scalability, efficiently addressing problems involving more than three objectives. The surface settlement, cutter wear, advance speed and energy consumption of the shield machine drive system are well optimized and controlled in the seven scenarios. Fig. 19 shows the percentage improvement in the optimization of the Pareto solution set under the different scenarios. Compared with the average for the original data sample, the average improvement percentage in the seven scenarios is 10.66 %.

(2) When the four-objective optimization is carried out, additional adjustments of the shield construction parameters can improve the degree of optimization of the objective. As shown in Table 9, in scenarios 1-3, when only one shield construction parameter is adjusted, the average improvement percentage is 6.25 %. When two shield construction parameters are adjusted in scenarios 4-6 at the same time, the average improvement percentage is 12.82 %. Adjusting the three shield construction parameters simultaneously in scenario 7 yields the highest average improvement of 17.39 %. The study results demonstrate that optimizing the three shield construction parameters yields the best effect in multiobjective optimization.

(3) The optimization considering the fourth objective highlights conflicts among objectives in the multiobjective optimization problem. As indicated in Table 9, the consideration of the fourth objective leads to a reduction in the optimization improvement for the first three objectives, resulting in decreases of 1.82 %, 1.68 %, and 2.23 %, respectively, on average. There are conflicts among the four objectives, and it is difficult to optimize multiple goals simultaneously. Due to certain conflicts among different goals, the optimization of one factor may occur at the expense of another [70]. This shows that when considering new goals, the relationships among goals may be complex and nonintuitive, and it may be difficult to achieve the optimization and improvement of all goals at the same time. Guo et al. [45] carried out the multiobjective optimization of the tunnel damage problem and verified the conflict among the objectives of the multiobjective optimization problem in complex underground engineering. He et al. [81] also found that there are certain conflicts among the objectives in multiobjective optimization problems, and it is difficult to optimize these objectives simultaneously.

5. Conclusion

To ensure the safety, economy, and efficiency of tunnel construction, minimizing surface settlement and cutter wear while enhancing advance speed is crucial. This paper establishes a multiobjective optimization model for shield construction parameters, leveraging the BO-RF-NSGA-III framework within the context of Guiyang Rail Transit Line 3. An RF model predicts surface settlement, cutter wear, and advance speed, and NSGA-III optimizes the shield construction parameters. The optimal decision is obtained through the ideal point method.

(1) The BO-RF method can reliably predict the surface settlement, cutter wear and advance speed. Its prediction accuracy for the test set surpasses that of the BO-BPNN, BO-SVM, and BO-GBDT models, with greater efficiency than the BO-LSTM model. The R^2 value for surface settlement prediction is 0.930, with an RMSE of 0.172 and MAE of 0.138. For cutter wear, the R^2 is 0.931, with

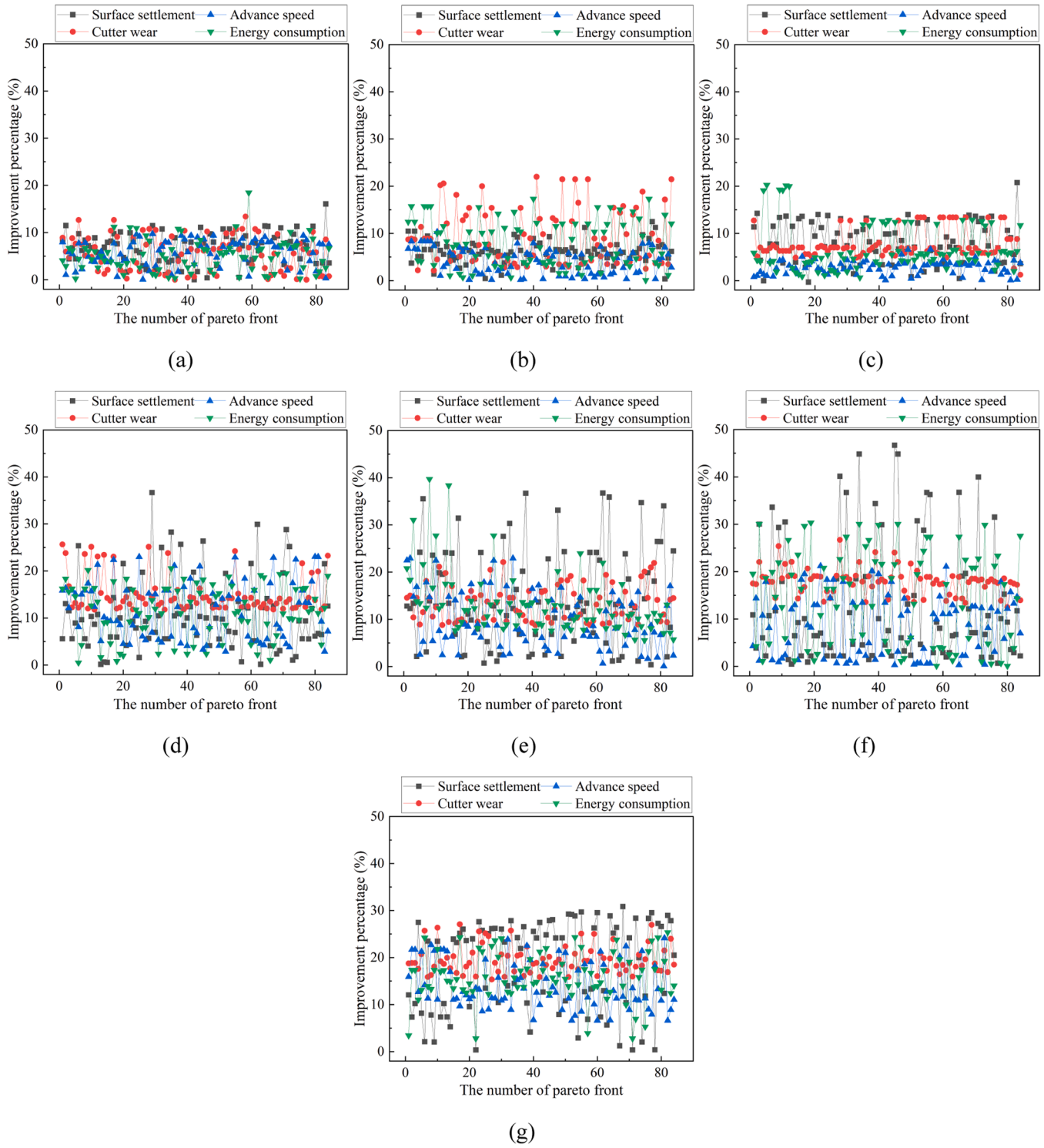


Fig. 19. Improvement percentage of the Pareto solution set under different scenarios: (a) scenario 1; (b) scenario 2; (c) scenario 3; (d) scenario 4; (d) scenario 5; (f) scenario 6; (g) scenario 7.

an RMSE of 0.153 and MAE of 0.114. Additionally, the R^2 for advance speed prediction is 0.938, with an RMSE of 0.138 and MAE of 0.112.

(2) The BO-RF algorithm identifies three key influencing factors, and optimization is performed under seven different scenarios to achieve the multiobjective optimization of surface settlement, cutter wear, and advance speed. Based on the proposed optimization principle, the optimal scheme is determined through calculations by adjusting three key shield construction parameters, namely, the chamber earth pressure, gross thrust and cutterhead torque. Compared with the average of the original sample, the average improvement percentage in the seven scenarios is 12.56 %. By adjusting the three key shield construction parameters at the same time, the average improvement percentage is maximized at 19.67 %.

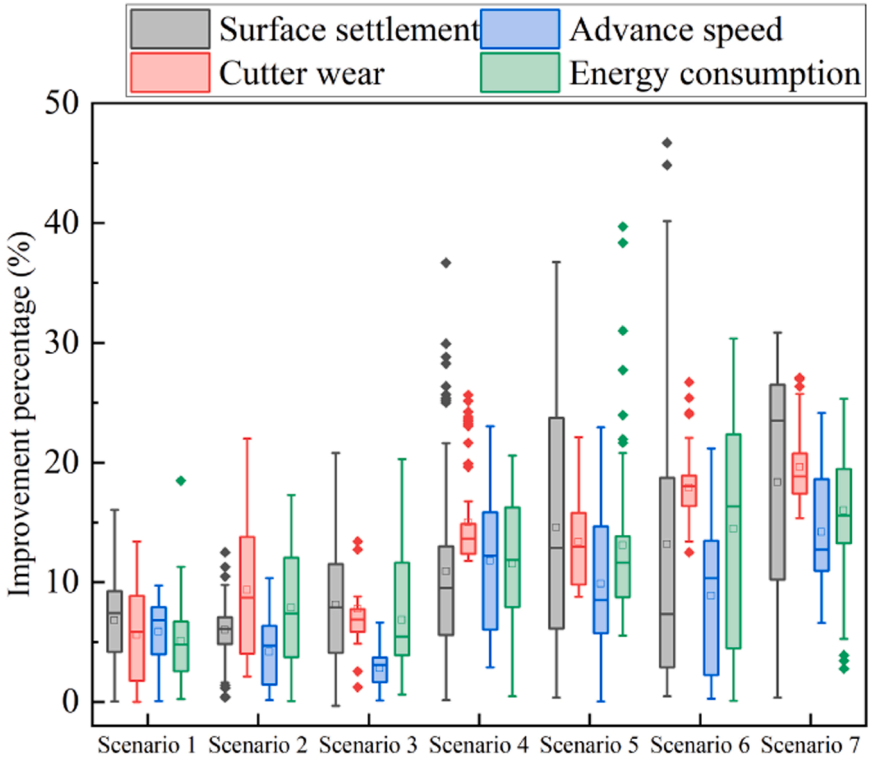


Fig. 20. Comparison of the improvement percentages in scenarios 1–7.

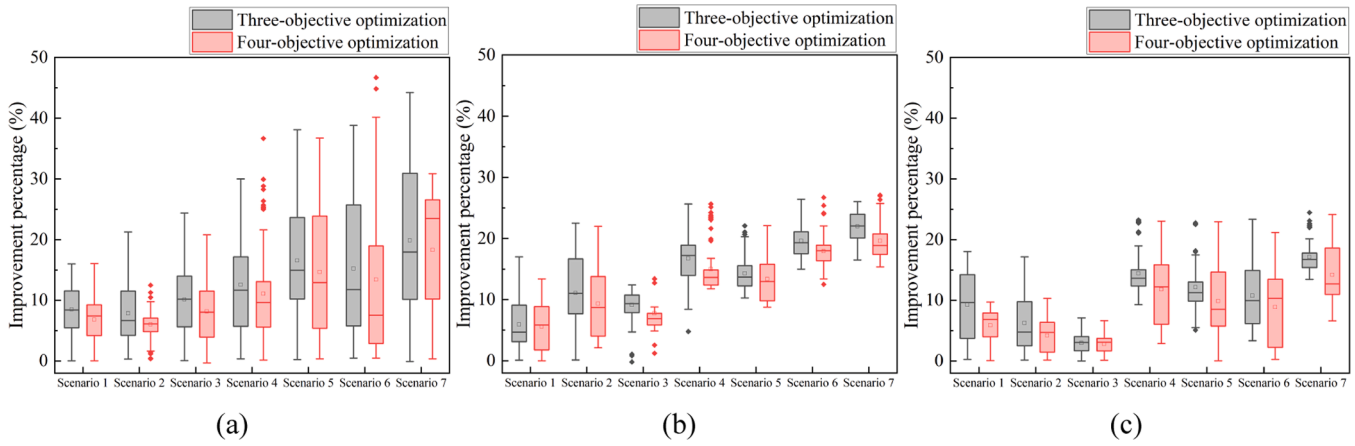


Fig. 21. Comparison of the results of three-objective optimization and four-objective optimization: (a) surface settlement; (b) cutter wear; (c) advance speed.

(3) Considering a fourth objective, the multiobjective optimization model is further improved, but due to the conflicts among the objectives, the optimization of the fourth objective reduces the optimization improvement compared to that in the case with only three objectives; notably, the energy consumption of the shield machine drive system is reduced by 10.70 %, and the surface settlement, cutter wear and advance speed optimization performance is reduced by 1.82 %, 1.46 % and 2.23 %, respectively.

This study has several potential limitations. First, although the shield parameters are adjusted and optimized to optimize the surface settlement, cutter wear and advance speed, the sample data used are only from a single engineering case. Future research should be based on additional engineering samples to augment the sample size and extend the application scope of the proposed method. Second, because the synchronous grouting hole is located behind the excavation face, the grouting parameters (grouting pressure and grouting amount) have a certain lag phenomenon compared to other inputs, so special treatment is needed. Additionally, different engineering projects have different objective requirements. When prioritizing a particular objective, it becomes essential to validate the effectiveness of the proposed method in the specific context of a project. Finally, Guizhou Province, where the metro case project is located, is the core area of contiguous karst landforms in southwestern China, and the concealed nature and uncertainty of various karst caves contained in karst landforms are major threats to tunnel construction.

Based on this study, the construction party of the China Railway Tunnel set up a shield tunnel test section, and the Line 3 shield tunnel was finished one year ahead of the original plan. The geological conditions of China and other parts of the world are different. According to the specific geological conditions, future research can further improve the multiobjective optimization model of shield parameters. According

Table 9
The degree of improvement of the four objectives in different scenarios.

Optimization objectives		Adjusted parameter								Average
		Original	x_2	x_3	x_6	x_2+x_3	x_2+x_6	x_3+x_6	$x_2+x_3+x_6$	
f_1 /mm	Mean	-3.49	-3.25 (6.81 %)	-3.28 (6.01 %)	-3.20 (8.11 %)	-3.11 (10.91 %)	-2.98 (14.59 %)	-3.03 (13.16 %)	-2.85 (18.35 %)	-3.10 (11.13 %)
f_2 /mm	Mean	0.174	0.164 (5.57 %)	0.157 (9.33 %)	0.160 (7.74 %)	0.147 (15.00 %)	0.150 (13.36 %)	0.142 (17.89 %)	0.139 (19.61 %)	0.151 (12.64 %)
f_3 /(mm·min^{-1})	Mean	22.48	23.80 (5.85 %)	23.43 (4.19 %)	23.11 (2.81 %)	25.13 (11.78 %)	24.70 (9.86 %)	24.48 (8.86 %)	25.68 (14.19 %)	24.33 (8.19 %)
f_4 /(kW·h)	Mean	737.75	700.15 (5.10 %)	679.67 (7.87 %)	687.26 (6.84 %)	652.74 (11.52 %)	641.21 (13.09 %)	631.18 (14.44 %)	619.70 (16.00 %)	658.85 (10.70 %)
Average improvement percentage		-	6.08 %	6.51 %	6.15 %	12.56 %	12.60 %	13.30 %	17.39 %	10.66 %
f_1 /mm	Optimal solution	-3.49	-3.36 (3.58 %)	-3.18 (8.90 %)	-3.27 (6.34 %)	-3.34 (4.02 %)	-3.00 (13.91 %)	-3.41 (2.21 %)	-2.48 (28.97 %)	-3.15 (9.70 %)
f_2 /mm	Optimal solution	0.174	0.162 (6.85 %)	0.164 (5.29 %)	0.163 (6.32 %)	0.153 (12.01 %)	0.148 (14.55 %)	0.142 (18.39 %)	0.144 (16.89 %)	0.154 (11.47 %)
f_3 /(mm·min^{-1})	Optimal solution	22.48	22.64 (0.70 %)	24.27 (7.95 %)	22.72 (1.03 %)	26.39 (17.36 %)	25.78 (14.67 %)	24.33 (8.21 %)	23.97 (6.61 %)	24.30 (8.07 %)
f_4 /(kW·h)	Optimal solution	737.75	601.44 (18.48 %)	610.21 (17.29 %)	588.16 (20.28 %)	588.97 (20.17 %)	444.89 (39.70 %)	513.77 (30.36 %)	550.87 (25.33 %)	556.90 (24.51 %)
Average improvement percentage		-	3.71 %	7.38 %	4.56 %	11.13 %	14.38 %	9.60 %	17.49 %	9.75 %

to the characteristics of specific engineering projects, different objectives should be assigned corresponding weights to optimize and adjust the shield construction parameters. Additionally, the possible negative influence of the hysteresis effect of grouting parameters can be reduced or compensated for to further improve the efficiency and safety of tunnel construction processes.

CRediT authorship contribution statement

Hongyu Chen: Writing – review & editing, Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Jun Liu:** Writing – review & editing, Writing – original draft, Software, Methodology, Data curation, Conceptualization. **Geoffrey Qiping Shen:** Writing – review & editing. **Luis Martínez:** Writing – review & editing. **Muhammet Deveci:** Writing – review & editing, Resources. **Zhen-Song Chen:** Writing – review & editing, Resources, Funding acquisition. **Yang Liu:** Writing – review & editing, Resources, Funding acquisition.

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.

Data availability

The data that has been used is confidential.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 72171182 and 72031009), the Natural Science Foundation of Hubei Province (Grant No. 2023AFC026) and the Social Science Foundation of Hubei Province (Grant No. 22ZD004).

References

[1] G. Shi, C. Qin, J. Tao, C. Liu, A VMD-EWT-LSTM-based multi-step prediction approach for shield tunneling machine cutterhead torque, Knowl. Based Syst. 228 (2021) 107213.
[2] H.Y. Chen, S. Yang, Z.B. Feng, Y. Liu, Y.W. Qin, Safety evaluation of buildings adjacent to shield construction in karst areas: an improved extension cloud approach, Eng. Appl. Artif. Intel. 124 (2023) 106386.
[3] A. Mahmoodzadeh, H.R. Nejati, M. Mohammadi, Optimized machine learning modelling for predicting the construction cost and duration of tunnelling projects, Autom. Constr. 139 (2022) 104305.
[4] S. Leng, J.R. Lin, Z.Z. Hu, X. Shen, A hybrid data mining method for tunnel engineering based on real-time monitoring data from tunnel boring machines, IEEE Access 8 (2020) 90430–90449.
[5] R.P. Chen, P. Zhang, X. Kang, Z.Q. Zhong, Y. Liu, H.N. Wu, Prediction of maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods, Soils Found. 59 (2) (2019) 284–295.
[6] Y. Liu, H.Y. Chen, L.M. Zhang, X.J. Wang, Risk prediction and diagnosis of water seepage in operational shield tunnels based on random forest, J. Civ. Eng. Manag. 27 (2021) 539–552.
[7] M.W. Gui, S.L. Chen, Estimation of transverse ground surface settlement induced by DOT shield tunneling, Tunn. Undergr. Space Technol. 33 (2013) 119–130.
[8] L. Wang, Y. Kang, Z. Cai, Q. Zhang, Y. Zhao, H. Zhao, P. Su, The energy method to predict disc cutter wear extent for hard rock TBMs, Tunn. Undergr. Space Technol. 28 (2012) 183–191.
[9] X. Hu, C. He, G. Walton, Y. Fang, G. Dai, Laboratory model test of EPB shield tunneling in a cobble-rich soil, J. Geotechn. Geoenviron. Eng. 146 (10) (2020) 04020112.
[10] Z. Sun, H. Zhao, K. Hong, K. Chen, J. Zhou, F. Li, B. Zhang, F. Song, Y. Yang, R. He, A practical TBM cutter wear prediction model for disc cutter life and rock wear ability, Tunn. Undergr. Space Technol. 85 (2019) 92–99.
[11] M. Melis, L. Medina, J.M. Rodríguez, Prediction and analysis of subsidence induced by shield tunnelling in the Madrid Metro extension, Can. Geotechn. J. 39 (6) (2002) 1273–1287.
[12] Y. Fang, Z. Yao, W. Xu, Q. Tian, C. He, S. Liu, The performance of TBM disc cutter in soft strata: A numerical simulation using the three-dimensional RBD-DEM coupled method, Eng. Failure Anal. 119 (2021) 104996.
[13] J. Fu, J. Yang, H. Klapperich, S. Wang, Analytical prediction of ground movements due to a nonuniform deforming tunnel, Int. J. Geomech. 16 (4) (2016) 04015089.
[14] X. Ling, X. Kong, L. Tang, Y. Zhao, W. Tang, Y. Zhang, Predicting earth pressure balance (EPB) shield tunneling-induced ground settlement in compound strata using random forest, Transp. Geotech. 35 (2022) 100771.
[15] X.G. Wu, Z.B. Feng, Y. Liu, Y.W. Qin, T.Y. Yang, J.C. Duan, Enhanced safety prediction of vault settlement in urban tunnels using the pair-copula and Bayesian network, Appl. Soft. Comput. 132 (2023) 109711.
[16] S. Yüksel, S. Eti, H. Dinçer, Y. Gökalp, Comprehensive risk analysis and decision-making model for hydroelectricity energy investments, J. Soft Comput. Decis. Anal. 2 (1) (2024) 28–38.
[17] Y. Gökalp, H. Dinçer, S. Eti, S. Yüksel, Generating a novel artificial intelligence-based decision-making model for determining priority strategies for improving community health, J. Intell. Stud. Bus. 2 (1) (2024) 1–13.
[18] H. Huang, J. Chang, D. Zhang, J. Zhang, H. Wu, G. Li, Machine learning-based automatic control of tunneling posture of shield machine, J. Rock Mech. Geotech. Eng. 14 (4) (2022) 1153–1164.
[19] S. Seker, M.T. Ergün, Investigation the effect of Covıd-19 pandemic in the sales for online education using machine learning methods, J. Soft Comput. Decis. Anal. 1 (1) (2023) 273–282.
[20] K. Wada, H. Sugiyama, K. Nozawa, M. Honda, S. Yamamoto, Guidance System for directional control in shield tunneling using machine learning techniques, in: Proceedings of the 18th International Conference on Computing in Civil and Building Engineering: ICCCBE 2020, Springer, 2021, pp. 73–88.
[21] X. Yan, D. She, Y. Xu, M. Jia, Deep regularized variational autoencoder for intelligent fault diagnosis of rotor–bearing system within entire life-cycle process, Knowl. Based Syst. 226 (2021) 107142.

[22] Y. Liu, H.Y. Chen, L.M. Zhang, X.G. Wu, X.J. Wang, Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: a case study in China, J. Clean. Prod. 272 (2020) 122542.
[23] M. Ma, G. Zhao, B. He, Q. Li, H. Dong, S. Wang, Z.J. Wang, XGBoost-based method for flash flood risk assessment, 598 (2021) 126382.
[24] L. Breiman, Random forests. 2001; 45 (1): 5–32.
[25] V. Kohestani, M. Bazarganlari, J. Asgari Marnani, Prediction of maximum surface settlement caused by earth pressure balance shield tunneling using random forest, J. AI Data Min. 5 (1) (2017) 127–135.
[26] R. Chen, P. Zhang, H. Wu, Z. Wang, Z. Zhong, Prediction of shield tunneling-induced ground settlement using machine learning techniques, Front. Struct. Civ. Eng. 13 (6) (2019) 1363–1378.
[27] Z. Huang, H. Zhang, Z. Long, W. Qiu, G. Meng, L. Zhu, Field test optimization of shield tunnelling parameters undercrossing an existing high-speed railway tunnel: a case study, Geotech. Geol. Eng. 39 (2) (2021) 1381–1398.
[28] L. Wang, H. Li, X. Zhao, Q. Zhang, Development of a prediction model for the wear evolution of disc cutters on rock TBM cutterhead, Tunn. Undergr. Space Technol. 67 (2017) 147–157.
[29] M. Alber, Advance rates of hard rock TBMs and their effects on project economics, Tunn. Undergr. Space Technol. 15 (1) (2000) 55–64.
[30] H. Li, Q. Zhang, Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II, IEEE Trans. Evol. Comput. 13 (2) (2008) 284–302.
[31] B. Liu, Y. Wang, G. Zhao, B. Yang, R. Wang, D. Huang, B. Xiang, Intelligent decision method for main control parameters of tunnel boring machine based on multi-objective optimization of excavation efficiency and cost, Tunn. Undergr. Space Technol. 116 (2021) 104054.
[32] K. Guo, L. Zhang, Multi-objective optimization in tunnel line alignment under uncertainty, Autom. Constr. 122 (2021) 103504.
[33] Z. Fan, Z. Zheng, B. Xu, W. Li, Y. Zhang, Z. Hao, Performance optimization of hard rock tunnel boring machine using multi-objective evolutionary algorithm, Comput. Ind. Eng. 169 (2022) 108251.
[34] L. Zhang, P. Lin, Multi-objective optimization for limiting tunnel-induced damages considering uncertainties, Reliab. Eng. Syst. Saf. 216 (2021) 107945.
[35] K. Guo, L. Zhang, Adaptive multi-objective optimization for emergency evacuation at metro stations, Reliab. Eng. Syst. Saf. 219 (2022) 108210.
[36] R. Arabahmadi, M. Mohammadi, M. Samizadeh, M. Rabbani, K. Gharibi, Facility location optimization for technical inspection centers using multi-objective mathematical modeling considering uncertainty, J. Soft Comput. Decis. Anal. 1 (1) (2023) 181–208.
[37] Y. Liu, T.J. Li, W.S. Xu, Q. Wang, H. Huang, B.J. He, Building information modelling-enabled multi-objective optimization for energy consumption parametric analysis in green buildings design using hybrid machine learning algorithms, Energy Build. 300 (2023) 113665.
[38] G. Wang, X. Li, L. Gao, P. Li, Energy-efficient distributed heterogeneous welding flow shop scheduling problem using a modified MOEA/D, Swarm Evol. Comput. 62 (2021) 100858.
[39] K. Zhong, G. Zhou, W. Deng, Y. Zhou, Q. Luo, MOMPA: multi-objective marine predator algorithm, Comput. Methods Appl. Mech. Eng. 385 (2021) 114029.
[40] S. Divall, R.J. Goodey, Twin-tunnelling-induced ground movements in clay, Proc. Inst. Civ. Eng. Geotech. Eng. 168 (3) (2015) 247–256.
[41] L. Zhang, X. Wu, W. Ji, S.M. AbouRizk, Intelligent approach to estimation of tunnel-induced ground settlement using wavelet packet and support vector machines, J. Comput. Civ. Eng. 31 (2) (2017) 04016053.
[42] Q. Zhang, X. Feng, X. Chen, K. Lu, Mix design for recycled aggregate pervious concrete based on response surface methodology, Constr. Build. Mater. 259 (2020) 119776.
[43] H. Chen, T. Deng, T. Du, B. Chen, M.J. Skibniewski, L. Zhang, An RF and LSSVM–NSGA-II method for the multi-objective optimization of high-performance concrete durability, Cem. Concr. Compos. 129 (2022) 104446.
[44] H. Chen, X. Li, Z. Feng, L. Wang, Y. Qin, M.J. Skibniewski, Z.S. Chen, Y. Liu, Shield attitude prediction based on Bayesian-LGBM machine learning, Inf. Sci. 632 (2023) 105–129.
[45] K. Guo, L. Zhang, Data-driven optimization for mitigating tunnel-induced damages, Appl. Soft Comput. 115 (2022) 108128.
[46] C. Fu, P. Wang, L. Zhao, X. Wang, A distance correlation-based Kriging modeling method for high-dimensional problems, Knowl. Based Syst. 206 (2020) 106356.
[47] D.J. Ren, S.L. Shen, A. Arulrajah, W.C. Cheng, Prediction model of TBM disc cutter wear during tunnelling in heterogeneous ground, Rock Mech. Rock Eng. 51 (2018) 3599–3611.
[48] Y. Liu, H. Chen, L. Zhang, Z. Feng, Enhancing building energy efficiency using a random forest model: a hybrid prediction approach, Energy. Rep. 7 (2021) 5003–5012.
[49] Y. Liu, Y. Cao, L. Wang, Z.-S. Chen, Y. Qin, Prediction of the durability of high-performance concrete using an integrated RF-LSSVM model, Constr. Build. Mater. 356 (2022) 129232.
[50] T.T. Joy, S. Rana, S. Gupta, S. Venkatesh, Fast hyperparameter tuning using Bayesian optimization with directional derivatives, Knowl. Based Syst. 205 (2020) 106247.
[51] Y. Liu, X.J. Wang, Z.S. Chen, Y. Zhang, S. Zhao, M. Devici, L. Jin, M.J. Skibniewski, Evaluating digital health services quality via social media, IEEE Trans. Eng. Manag. (2023) 1–13.
[52] L. Yang, Y. Ao, J. Ke, Y. Lu, Y. Liang, To walk or not to walk? Examining non-linear effects of streetscape greenery on walking propensity of older adults, J. Transp. Geogr. 94 (2021) 103099.
[53] H. Chen, Y. Cao, Y. Liu, Y. Qin, L. Xia, Enhancing the durability of concrete in severely cold regions: mix proportion optimization based on machine learning, Construct. Build. Mater. 371 (2023) 130644.
[54] L. Yang, H. Yang, B. Yu, Y. Lu, J. Cui, D. Lin, Exploring non-linear and synergistic effects of green spaces on active travel using crowdsourced data and interpretable machine learning, Travel Behav. Soc. 34 (2024) 100673.
[55] Y.C. Cao, W.W. Yu, W. Ren, G.R. Chen, An overview of recent progress in the study of distributed multi-agent coordination, IEEE Trans. Ind. Inf. 9 (1) (2013) 427–438.
[56] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints, IEEE Trans. Evol. Comput. 18 (4) (2013) 577–601.
[57] Y. Liu, H. Qin, Z. Zhang, L. Yao, C. Wang, L. Mo, S. Ouyang, J. Li, A region search evolutionary algorithm for many-objective optimization, Inf. Sci. 488 (2019) 19–40.
[58] H.Y. Chen, J.Y. Wang, Z.B. Feng, Y. Liu, W. Xu, Y.W. Qin, Research on the risk evaluation of urban wastewater treatment projects based on an improved fuzzy cognitive map, Sustain Cities Soc. 98 (2023) 104796.
[59] Q. Fang, J.m. Du, J.y. Li, D.l. Zhang, L.q. Cao, Settlement characteristics of large-diameter shield excavation below existing subway in close vicinity, J. Cent. South Univ. 28 (3) (2021) 882–897.
[60] D. Kim1b, K. Pham1a, S. Park1b, J.Y. Oh2a, H. Choi, Determination of effective parameters on surface settlement during shield TBM, (2020).
[61] H. Yu, X. Zhou, X. Zhang, M. Mooney, Enhancing earth pressure balance tunnel boring machine performance with support vector regression and particle swarm optimization, Autom. Constr. 142 (2022) 104457.
[62] X. Liu, S. Xu, Y. Huang, Optimal control for earth pressure balance of shield machine based on action-dependent heuristic dynamic programming, ISA Trans. 94 (2019) 28–35.
[63] D. Jin, Y. Guo, X. Li, Y. Yang, Y. Fang, Numerical study on the muck flow behavior in the screw conveyor during EPB shield tunneling, Tunn. Undergr. Space Technol. 134 (2023) 105017.
[64] Y. Wei, Y. Yang, M. Tao, D. Wang, Y. Jie, Earth pressure balance shield tunneling in sandy gravel deposits: a case study of application of soil conditioning, Bull. Eng. Geol. Environ. 79 (2020) 5013–5030.
[65] C. Heng, S. Sun, Z. Zhou, J. Zhang, Prediction of surface settlement with ultra-shallow-burial and large rectangular cross-section urban underpass, KSCE J. Civ. Eng. 23 (11) (2019) 4641–4650.
[66] X. Liu, D. Liu, F. Xiong, Y. Han, R. Liu, Q. Meng, Z. Zhong, Q. Chen, C. Weng, W. Liu, Experimental and numerical study on the stability of slurry shield tunneling in circular-gravel layer with different cover-span ratios, Geomech. Eng. 28 (3) (2022) 265–281.
[67] V. Ghiasi, M. Koushki, Numerical and artificial neural network analyses of ground surface settlement of tunnel in saturated soil, SN Appl. Sci. 2 (5) (2020) 1–14.
[68] X.G. Wu, X.Y. Li, Y.W. Qin, W. Xu, Y. Liu, Intelligent multiobjective optimization design for NZEBs in China: four climatic regions, Appl. Energ. 339 (2023) 120934.
[69] Y. Cheng, W.H. Zhou, T. Xu, Tunneling-induced settlement prediction using the hybrid feature selection method for feature optimization, Transp. Geotech. 36 (2022) 100808.
[70] D. Wu, C. Jennings, J. Terpenny, R.X. Gao, S. Kumara, A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests, J. Manuf. Sci. Eng. 139 (7) (2017).
[71] J. Zhou, E. Li, F. Gong, M. Wang, Q. Qiao, Development of random forests and Cubist models for predicting TBM penetration rate in hard rock condition, in: Proceedings of the 15th Annual Meeting of Chinese rock mechanics and Engineering (China Rock 2018), Beijing, China, 2018, p. 10.
[72] L. Tang, S. Na, Comparison of machine learning methods for ground settlement prediction with different tunneling datasets, J. Rock Mech. Geotech. Eng. 13 (6) (2021) 1274–1289.
[73] K. Zhang, H.M. Lyu, S.L. Shen, A. Zhou, Z.Y. Yin, Evolutionary hybrid neural network approach to predict shield tunneling-induced ground settlements, Tunn. Undergr. Space Technol. 106 (2020) 103594.
[74] Y. Liang, X. Chen, J. Yang, J. Zhang, L. Huang, Analysis of ground collapse caused by shield tunnelling and the evaluation of the reinforcement effect on a sand stratum, Eng. Failure Anal. 115 (2020) 104616.
[75] X.P. Zhou, S.F. Zhai, Estimation of the cutterhead torque for earth pressure balance TBM under mixed-face conditions, Tunn. Undergr. Space Technol. 74 (2018) 217–229.
[76] C. Li, Simplified algorithm for grouting pressure and grouting quantity in shield construction, Int. J. Civ. Eng. 18 (4) (2020) 419–428.
[77] L. Cao, D. Zhang, Q. Fang, L. Yu, Movements of ground and existing structures induced by slurry pressure-balance tunnel boring machine (SPB TBM) tunnelling in clay, Tunn. Undergr. Space Technol. 97 (2020) 103278.
[78] L. Wang, H.Y. Chen, Y. Liu, H. Li, W.J. Zhang, Application of copula-based Bayesian network method to water leakage risk analysis in cross river tunnel of Wuhan Rail Transit Line 3, Adv. Eng. Inform. 57 (2023) 102056.
[79] X. Wu, L. Wang, B. Chen, Z. Feng, Y. Qin, Q. Liu, Y. Liu, Multi-objective optimization of shield construction parameters based on random forests and NSGA-II, Adv. Eng. Inf. 54 (2022) 101751.
[80] K. Elbaz, T. Yan, A. Zhou, S.L. Shen, Deep learning analysis for energy consumption of shield tunneling machine drive system, Tunn. Undergr. Space Technol. 123 (2022) 104405.
[81] R. He, L. Zhang, Multi-objective optimization for cost-effective aseismic design of submerged floating tunnels considering weighted preferences, Ocean Eng. 250 (2022) 110976.