

Lulu WANG, Penghui LIN, Yongsheng LI, Hui LUO, Limao ZHANG

# Physics-informed online deep learning for advanced control of shield tail clearance in tunnel construction

© Higher Education Press 2025

**Abstract** To more accurately estimate and control the magnitude of the shield tail clearance, a hybrid deep learning model with the integration of an online physics-informed deep neural network (online PDNN) and non-dominated sorting genetic algorithm-II (NSGA-II) is developed. The online PDNN has evolved from a deep learning framework constrained by the underlying physical mechanism of shield tail clearance measurements. The algorithm is used to forecast the shield tail clearance in tunnel boring machines (TBMs). The NSGA-II is employed to conduct the multi-objective optimization (MOO) process for shield tail clearance. The proposed method is validated in a tunnel case in China. Experimental results reveal that: (1) In comparison with some state-of-the-art algorithms, the online PDNN model demonstrates superior capability in predicting shield tail clearance above, upper-left, and upper-right, with  $R^2$  scores of 0.93, 0.90, and 0.90, respectively; (2) The MOO achieves a comprehensive optimal solution, with the overall improvement percentage of shield tail clearance reaching 30.87% and a hypervolume of 32 under the 20% constraint condition, which surpasses the average performance of other MOO frameworks by 23 and 5.48%, respectively. The novelty of this research lies in coupling the constructed physical constraints and the online update mechanism into a causal analysis-oriented data-driven model, which not only enhances the model's performance and interpretability but also realizes the

control for the shield tail clearance by the integration of NSGA-II.

**Keywords** physics-informed neural network, online update mechanism, multi-objective optimization, NSGA-II, shield tail clearance

## 1 Introduction

Nowadays, rapid urbanization and population growth make the problem of traffic congestion more and more serious. To alleviate this situation, large-scale urban rail transit construction has been accelerating. A tunnel boring machine (TBM), as a piece of advanced equipment for tunneling, has been widely used in metro tunnel excavation. Shield tail clearance, referring to the clearance between the inner edge of the shield shell and the outer edge of the lining, plays a crucial role in the rational selection of the lining segment, the attitude correction of the shield machine, and the control of tunneling orientation (Liu et al., 2021). However, shield tail clearance estimation and control face various challenges under different tunneling conditions. In soft-hard uneven composite strata, the unbalanced distribution of thrust on the tunnel face can cause shield attitude deviations and over-excavation, directly affecting the clearance variation (Feng et al., 2022; Tang et al., 2022). Furthermore, silty fine sand and sandy clay interlayers tend to accumulate around the shield, forming a hard-plastic shell as water drains. This increases friction resistance, leading to uneven shield tail deformation and impacting clearance size (Han et al., 2021). These issues are particularly pronounced in shield TBMs, which have a larger contact area with the surrounding rock compared to gripper TBMs (Xu et al., 2021). The practice has shown that shield tail clearances that are either too narrow or too large can compromise construction safety (Liu and Bao, 2021; Lou et al., 2023). Therefore, it is necessary to study how to precisely measure and control the shield tail

Received Aug. 12, 2024; revised Jan. 20, 2025; accepted Mar. 8, 2025

Lulu WANG, Yongsheng LI, Hui LUO, Limao ZHANG (✉)  
School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China  
E-mail: limao\_zhang@hotmail.com

Penghui LIN  
School of Civil and Environmental Engineering, Nanyang Technological University, Singapore 639798, Singapore

This work is supported in part by the National Natural Science Foundation of China (Grand No. 72271101) and the Key Laboratory of Water Grid Project and Regulation of Ministry of Water Resources, China (No. QTKS0034W25018).

clearance to ensure the high efficiency and reliability of shield tunnel construction.

The main methods for shield tail clearance measurement include theoretical analysis, manual measurement, contact measurement, and non-contact measurement. The theory of minimum shield tail clearance calculation (Yu, 2022) is one of the effective methods for estimating shield tail clearance. However, due to its simplification of engineering reality and inability to pinpoint the location of the minimal shield tail clearance, this method is not widely used in engineering applications. At present, the most common way to evaluate shield tail clearance is to adopt a manual measurement technique that is user-friendly and yields intuitive results. Nevertheless, manual measurement has many drawbacks, such as human error proneness, low levels of automation, poor real-time performance, and safety hazards (He et al., 2021). The contact measurement method realizes the measurement of shield tail clearance by installing a mechanical device at the shield tail. In fact, the complexity of the shield construction environment makes it inconvenient to install and replace the measuring equipment, thereby increasing the likelihood of damage to the tube segment (Yang, 2021). Compared with the mentioned three measurement ways, the non-contact measurement method is also popular in clearance measurement.

Non-contact measurement mainly utilizes ultrasonic, laser, vision, and imaging technologies to measure shield tail clearance. Specifically, the ultrasonic-based measurement method utilizes ultrasonic sensors embedded inside the shield shell to send and receive ultrasonic signals uniformly distributed on the inner wall of the shield shell to accomplish the clearance measurement. However, the effectiveness of this method's application is influenced by the thickness of the shield shell, the cleanliness of the environment, and the performance of the ultrasonic rangefinder (VMT, 2021). The laser-based measurement technology uses a laser range finder to scan the inner wall of the shield shell and the outer surface of the lining to locate their edges. Subsequently, shield tail clearance can be calculated indirectly by combining the scanning angles. Yet, its main drawback is that the accumulation of errors could affect the accuracy and stability of data (Sun and Zhuang, 2016). Currently, shield tail clearance measurement based on machine vision is a research direction that has attracted much attention. This method focuses on utilizing digital image processing techniques to calibrate the captured shield tail clearance images, extract the edges of the shield shell inwall and segment outwall, and calculate the clearance by considering the spatial structure of the shield tail (Chen et al., 2021). Although machine vision-based (MVB) systems perform well in laboratory tests, they are prone to interference from stray light, dust, and muck in the underground working environment. This vulnerability leads to unsatisfactory measurement accuracy in practical applications (Chen et al., 2020).

With the development of information technology, the vast amount of data collected by sensors has facilitated the application of machine learning (ML) and deep learning (DL) technologies in the field of civil engineering. At present, these methods are mainly applied in posture recognition (Chen et al., 2025), construction management (Zhou et al., 2018), and risk analysis (Zhang et al., 2024b), but they are less applied in shield tail clearance measurement. Furthermore, similar to the trajectory planning problem proposed by Chai et al. (2020), the shield tail clearance control task can also be regarded as a multi-objective optimization (MOO) problem. The shield tail clearances in different orientations can be viewed as different optimization objectives, which can then be minimized by appropriately adjusting the TBM's key operating parameters. This approach yields a reasonable adjustment range for these parameters to keep the clearance within specified limits, thereby ensuring the smooth progress of the excavation process. Statistically, the non-dominant sorting genetic algorithm-II (NSGA-II) is one of the most popular methods for solving MOO problems (Verma et al., 2021). For example, Chai et al. (2017) efficiently solved the multi-objective trajectory optimization problem for spacecraft using an improved NSGA-II method. Zhang and Lin (2021) implemented a hybrid approach that integrates extreme gradient boosting (Xgboost) and NSGA-II to mitigate tunneling-induced limit support pressure and ground surface deformation. However, to the best of our knowledge, no researchers have attempted to apply NSGA-II to solve the control problems related to shield tail clearance.

Overall, previous studies have significantly contributed to the advancement of shield tail clearance measurement techniques. However, due to the time-varying nature of shield tail clearance, existing methods still have some deficiencies, such as their inability to provide valuable evidence for controlling shield tail clearance. To address these challenges, exhaustive efforts herein develop an online physics-informed DL model to support the automatic measurement and control of shield tail clearance. There are three key research questions to be addressed: (1) How to construct the physical equations for calculating shield tail clearance at arbitrary positions and utilize them as physical constraints for training deep neural network (DNN) models; (2) how to facilitate the online training process of physics-informed DL models to realize dynamic prediction of shield tail clearance; (3) how to implement the MOO of shield tail clearance to provide a certain reference for on-site operators to control the shield tail clearance. The innovations of this research can be summarized in two key aspects: (a) Integrating the derived physical equations for clearance measurement and the online updating mechanism into a DNN framework to generate the online physics-informed DNN (online PDNN) that can accurately predict the shield tail clearance, and (b) incorporating the developed online PDNN

model with the NSGA-II algorithm to realize intelligent control of the shield tail clearance at arbitrary positions.

The remaining parts of this paper are organized as follows. Section 2 reviews existing research on the application of ML, physics-informed neural networks (PINNs), and online learning methods to tunnel problems. Section 3 introduces the details of our method, including the physical mechanism of shield tail clearance measurement, the framework of the online PDNN, and the MOO for shield tail clearance control. Section 4 applies the developed online PDNN approach to the actual construction project of Wuhan Metro Line 12, Guobo Center South Station–Lingwucun Station, aiming to verify its feasibility and effectiveness. Section 5 discusses the effect of using different algorithms in estimating and controlling shield tail clearance to highlight the advantages of the proposed approach. Section 6 summarizes the findings and presents future work.

---

## 2 Related studies

With the progress of modern information technology, the tunnel construction industry undergoing an important stage of digital transformation, where an immersing interest is to apply ML technology to improve the automation level of shield tunneling. For one thing, the growing abundance of monitoring data related to TBM can create a big-data environment suitable for training ML models (Yin et al., 2022). For another, ML has powerful learning capabilities to construct potentially hidden relationships among multidimensional data, laying a critical foundation for the autonomous operation of control systems (Chai et al., 2024). For example, Afradi et al. (2020) accurately predicted the TBM penetration rate using a support vector machine (SVM). Mahmoodzadeh et al. (2021) found that Gaussian process regression (GPR) is more suitable for predicting the disc cutter's life of TBM than support vector regression (SVR), decision trees (DT), and K-nearest neighbors (KNN) algorithms. Although these methods have been proven effective, the shallow learning mechanisms inherent in ML often lead to insufficient accuracy when capturing more complex nonlinear relationships between variables.

To break through the inherent limitations of ML, DL has been developed as an important branch of ML. The distinguishing advantage of DL is its ability to automatically carry out feature engineering, learn the implicit relationships between input features and output targets, and provide accurate prediction results. For example, Feng et al. (2021) predicted the field penetration index of three rock types with an average relative error of less than 0.15 by deploying a deep belief network (DBN). Zhang et al. (2024a) developed a hybrid DL model that integrates graph convolutional network (GCN) and long short-term memory (LSTM) networks, which is capable of predicting

TBM penetration rate, over-excavation rate, energy consumption, and tool wear with a coefficient of determination ( $R^2$ ) score of no less than 0.83. In summary, the results of these existing studies confirm the effectiveness of data-driven approaches in addressing tunneling problems. However, the application of ML/DL strategies in tunneling projects still faces several challenges (Phoon and Zhang, 2023): (1) Shield machine tunneling reflects interactions between the ground and the machine, constrained by specific physical mechanisms, but purely data-driven ML/DL models often fail to offer interpretable insights into these mechanisms. (2) One of the prerequisites for ML/DL methods to describe complex interactions is the need for a large amount of training data, which is often very limited for many engineering projects.

To address these fundamental challenges, PINN was proposed in 2019 (Raissi et al., 2019). Since its inception, the PINN has gained more attention in the tunneling engineering domain. For example, Zhang et al. (2023) and Liu et al. (2023) performed PINN models for estimating ground deformation caused by tunnel excavation. Shaban et al. (2023) used a PDNN to model the diffusion of chloride ions and predict the distribution of chloride concentrations in concrete. Lin et al. (2023) used a physics-informed DL approach to reliably predict the soil chamber pressure of TBM with an  $R^2$  score of 0.96. Although most conventional PINNs yield acceptable results, the prerequisite is that the entire training data set needs to be provided before the learning task, which is detrimental to the shield tail clearance measurement task. As a possible solution, an online learning manner can therefore be taken into account to address the aforementioned limitations, which can potentially be more practical and robust in practice compared to offline PINN.

Online learning methods can adaptively update the model with the incoming stream of observations, enabling the model to adapt more efficiently to the dynamic excavation process. Recently, some research on online ML/DL modeling in the field of tunnel engineering has been reported. Pan et al. (2022) developed an online version of the att-GCN model to predict the penetration and energy consumption of the shield machine. The results indicated that the online learning mechanism can improve the prediction accuracy of the model. Zhang et al. (2020a) integrated a reinforcement learning (RL) based optimizer with the extreme learning machine (ELM) model to enhance the forecasting of tunneling-induced ground response. The findings demonstrate that the optimized model achieves improved accuracy while reducing computational costs. The key to developing online ML/DL models is the determination of hyperparameters. The common methods for determining hyperparameters include grid search, random search, and Bayesian Optimization (BO). Compared with the other two methods, the BO selects parameter configurations for the next iteration by considering the results of previous

iterations, which confers a notable edge in tackling complex global optimization problems (Shin et al., 2020). Consequently, the BO algorithm is used in this study to automatically search for optimal hyperparameter combinations for online PDNN models.

To address existing challenges and fill research gaps, herein, a hybrid approach with the integration of online PDNN and NSGA-II is proposed to conduct the prediction and optimization of shield tail clearance. As expected, the developed online PDNN model will be more practical and reliable. To be more specific, the established physical equations are incorporated into the loss function of the DL model, which contributes to reducing over-reliance on training data and improving the prediction performance and interpretability of the model (Rao et al., 2020). The developed physics-informed DL model is trained online to accurately capture and represent the dynamic changes in shield tail clearance during tunnel excavation. Combining online PDNN with NSGA-II for performing MOO, the optimal operational decision obtained could serve as a reference for operators to control shield tail clearance, ensuring the safety and efficiency of the excavation process.

### 3 Methodology

In this paper, a physics-informed online DL network is developed to estimate and control the shield tail clearance during tunnel excavation. In this section, the physical laws relevant to computing shield tail clearance are principally introduced. The structure of the PDNN algorithm after integrating physical laws and the online update mechanism is then illustrated. Methods for evaluating and explaining the model are provided. At the end of this section, the MOO method for shield tail clearance control is described. An overall framework illustrating the proposed approach is shown in Fig. 1.

#### 3.1 Physical mechanism in shield tail clearance calculation

Currently, the minimum shield tail clearance calculation theory serves as an important tool for analyzing shield tail clearance. This method primarily estimates the minimum clearance value based on the geometric relationship between the clearance and factors such as the radius of the tunnel curve, the pitch angle, and the length of the shield tail covering segments (Yu, 2022). By subsequently comparing the results with existing minimum shield tail clearance standards, it assesses whether the current clearance is too small. However, this method requires construction personnel to accurately identify the location of the minimum shield tail clearance, which can be challenging. In practice, field workers typically monitor shield tail clearance at eight specific locations: above,

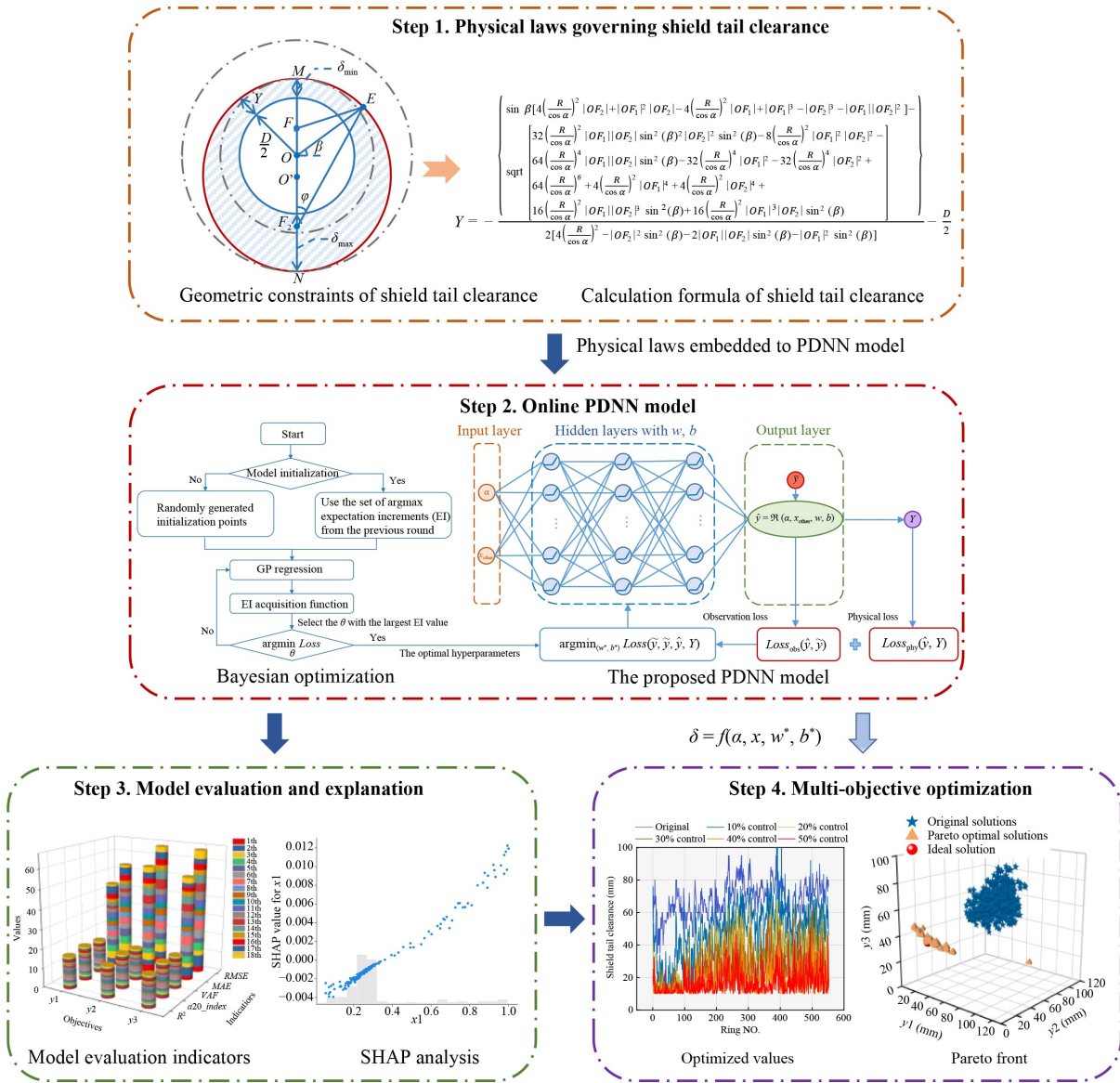
below, left, right, upper left, lower left, upper right, and lower right. Under these conditions, the minimum shield tail clearance theory becomes less applicable. Additionally, another limitation of this method is its inability to provide effective guidance for controlling shield tail clearance. To enhance its applicability in engineering scenarios, an improved method for calculating the shield tail clearance at different positions is proposed.

Assuming the assembly of the lining rings is in an ideal state, there are six models for the calculation of the minimum shield tail clearance (as shown in Figs. 2(a)–2(f)). In the first model, the TBM excavates along a straight line, and the central axes of the TBM, the lining ring, and the excavation route coincide (as exhibited in Fig. 3). The size of the shield tail clearance  $Y$  at different orientations is equal to the initial clearance  $\delta_0$  mm (as presented in Fig. 2(a)). In the second one, the variation of shield tail clearance generated by shield pitch angle  $\alpha$  ( $^\circ$ ) is  $\delta_1$  mm, where when the TBM is up tilted,  $\alpha > 0$ ; otherwise,  $\alpha < 0$ . Due to the influence of the shield pitch angle, the shield tail clearance values of the relative positions are not equal, in which the size of the larger shield tail clearance is  $\delta_{\max}$  mm and the smaller is  $\delta_{\min}$  mm (as depicted in Fig. 2(b)). In the third condition, since the TBM tunnels along a route with a deflection angle  $\eta$  ( $^\circ$ ) and a tunnel curve radius  $R'$  m, the shield tail clearance at relative positions differs by  $\delta_2$  mm (as displayed in Fig. 2(c)). In the fourth to sixth cases, the effects of the pitch angle and the radius of the tunnel curve on the shield tail clearance are considered at the same time (as shown in Figs. 2(d)–2(f)). Among them, Fig. 2(d) gives the case when  $\delta_1 < \delta_2$ , where the minimum shield tail clearance appears directly above, and the maximum shield tail clearance occurs directly below. Whereas when  $\delta_1 > \delta_2$ , the situation is the opposite, as indicated in Fig. 2(e). That is to say, the minimum shield tail clearance appears directly below, and the maximum shield tail clearance appears directly above. From the geometries of Fig. 2,  $\delta_1$ ,  $\delta_2$ ,  $\delta_{\min}$ , and  $\delta_{\max}$  can be estimated by Eqs. (1)–(4), respectively.

$$\delta_1 = l \tan \alpha, \quad (1)$$

$$\delta_2 = \left( R' - \frac{D}{2} \right) \left( 1 - \sqrt{1 - \left( \frac{l}{R' - D/2} \right)^2} \right), \quad (2)$$

$$\delta_{\min} = \begin{cases} \delta_0 - \frac{1}{2} \delta_1 & (\alpha \neq 0, \eta = 0) \\ \delta_0 - \frac{1}{2} \delta_2 & (\alpha = 0, \eta \neq 0) \\ \delta_0 + \frac{1}{2} \delta_1 - \frac{1}{2} \delta_2 & (\alpha > 0, \eta \neq 0, \text{ and } \delta_1 < \delta_2) \\ \delta_0 - \frac{1}{2} \delta_1 + \frac{1}{2} \delta_2 & (\alpha > 0, \eta \neq 0, \text{ and } \delta_1 > \delta_2) \\ \delta_0 - \frac{1}{2} \delta_1 - \frac{1}{2} \delta_2 & (\alpha < 0, \eta \neq 0) \end{cases}, \quad (3)$$



**Fig. 1** Overall flowchart of the developed online PDNN approach.

$$\delta_{\max} = 2\delta_0 - \delta_{\min}, \quad (4)$$

where  $l$  refers to the length of the segment covered by the shield tail, which is assumed to be equal to the width of the lining ring in this study;  $D$  represents the outer diameter of the lining ring.

Taking the subgraph Fig. 2(f) as an example, the basic principle of shield tail clearance calculation is shown in Fig. 4. The three concentric circles centered on  $O$  represent the lining ring with radius  $D/2$ , and the two auxiliary circles with radii  $(D/2 + \delta_{\min})$  and  $(D/2 + \delta_{\max})$ , respectively. The ellipse centered at  $O'$  is a 1-1 section of the shield shell.  $F_1$  and  $F_2$  are the two foci of the ellipse. The long and short axes of the ellipse are  $2R/\cos\alpha$  and  $2R$ , respectively, where  $R$  is the inner diameter of the shield shell. According to the definition of the ellipse and the cosine theorem, we can obtain the Eqs. (5)–(8), which

implicitly represents the geometric relationships between the shield tail clearance, pitch angle, radius of the tunnel curve, and horizontal angle ( $\beta$ ).

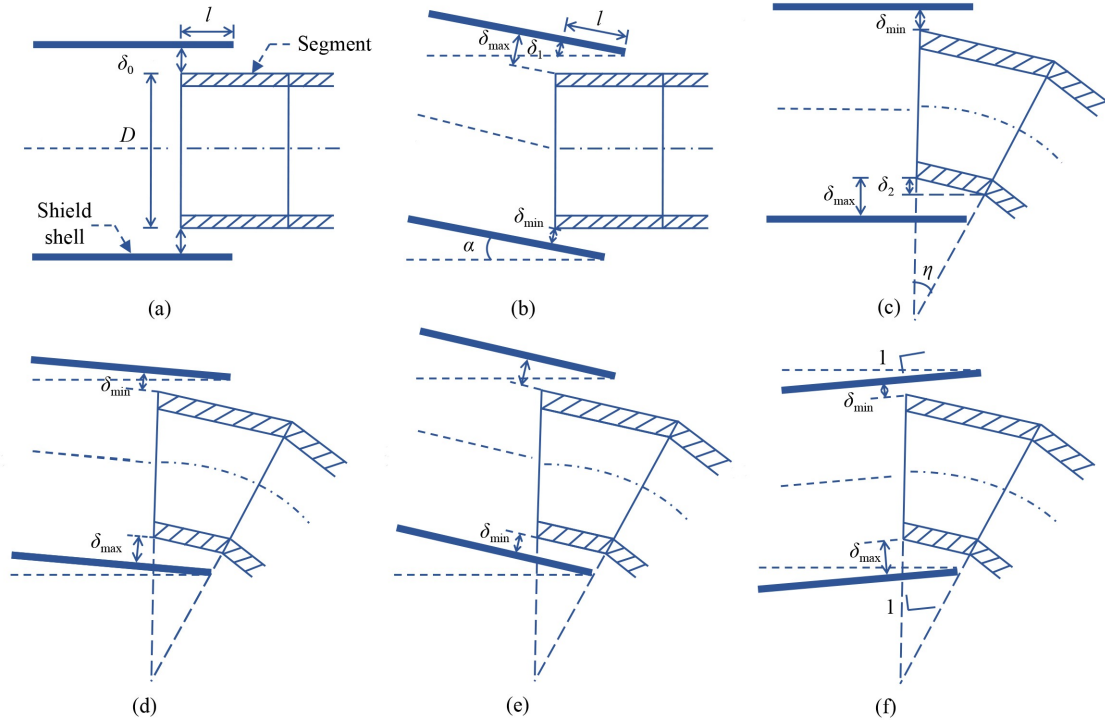
$$|EF_1| + |EF_2| = \frac{2R}{\cos\alpha}, \quad (5)$$

$$|EF_1|^2 = (|OF_1| + |OF_2|)^2 + |EF_2|^2 - 2(|OF_1| + |OF_2|)|EF_2|\cos\phi, \quad (6)$$

$$|EF_1|^2 = |OF_1|^2 + (R + Y)^2 - 2|OF_1|(R + Y)\sin\beta, \quad (7)$$

$$\left(\frac{D}{2} + Y\right)^2 = |OF_2|^2 + |EF_2|^2 - 2|OF_2||EF_2|\cos\phi, \quad (8)$$

where  $|EF_1|$  and  $|EF_2|$  are the distances from the measuring



**Fig. 2** Calculation models of shield tail clearance: (a) TBM excavates along the straight line, (b) TBM excavates in a pitch attitude, (c) TBM excavates along the curve, (d) TBM is up tilted and excavates along the curve ( $\delta_1 < \delta_2$ ), (e) TBM is up tilted and excavates along the curve ( $\delta_1 > \delta_2$ ), and (f) TBM is down tilted and excavates along the curve.

point to the focal points  $F_1$  and  $F_2$ , respectively, and  $\varphi$  is the angle between  $EF_2$  and  $F_1F_2$ .

After eliminating the three unknown variables  $|EF_1|$ ,

$|EF_2|$ , and  $\varphi$  by combining Eqs. (5)–(8), Eq. (9) for the calculation of the shield tail clearance at any position can be derived.

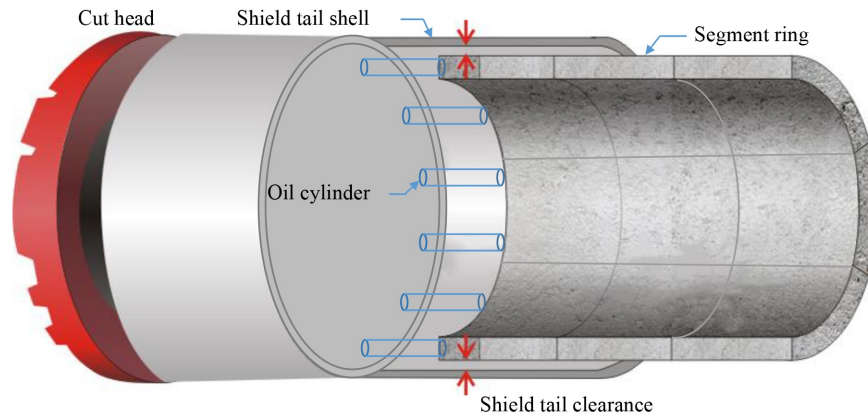
$$Y = \frac{\left\{ \begin{array}{l} \sin\beta \left[ 4\left(\frac{R}{\cos\alpha}\right)^2 |OF_2| + |OF_1|^2 |OF_2| - 4\left(\frac{R}{\cos\alpha}\right)^2 |OF_1| + |OF_1|^3 - |OF_2|^3 - |OF_1| |OF_2|^2 \right] - \\ \text{sqrt} \left[ \begin{array}{l} 32\left(\frac{R}{\cos\alpha}\right)^2 |OF_1| |OF_2| \sin^2(\beta) |OF_2|^2 \sin^2(\beta) - 8\left(\frac{R}{\cos\alpha}\right)^2 |OF_1|^2 |OF_2|^2 - \\ 64\left(\frac{R}{\cos\alpha}\right)^4 |OF_1| |OF_2| \sin^2(\beta) - 32\left(\frac{R}{\cos\alpha}\right)^4 |OF_1|^2 - 32\left(\frac{R}{\cos\alpha}\right)^4 |OF_2|^2 + \\ 64\left(\frac{R}{\cos\alpha}\right)^6 + 4\left(\frac{R}{\cos\alpha}\right)^2 |OF_1|^4 + 4\left(\frac{R}{\cos\alpha}\right)^2 |OF_2|^4 + \\ 16\left(\frac{R}{\cos\alpha}\right)^2 |OF_1| |OF_2|^3 \sin^2(\beta) + 16\left(\frac{R}{\cos\alpha}\right)^2 |OF_1|^3 |OF_2| \sin^2(\beta) \end{array} \right] \end{array} \right\}}{2 \left[ 4\left(\frac{R}{\cos\alpha}\right)^2 - |OF_2|^2 \sin^2(\beta) - 2|OF_1| |OF_2| \sin^2(\beta) - |OF_1|^2 \sin^2(\beta) \right]} - \frac{D}{2}, \quad (9)$$

$$|OF_1| = -\frac{R}{\cos\alpha} + R \tan\alpha + \frac{D}{2} + \delta_{\min}, \quad (10)$$

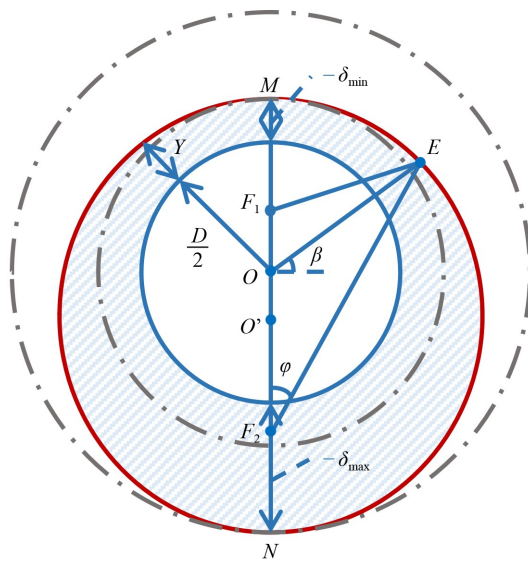
$$|OF_2| = \frac{R}{\cos\alpha} + R \tan\alpha - \frac{D}{2} - \delta_{\min}, \quad (11)$$

where  $\beta$  indicates the position at which the shield tail clearance needs to be calculated;  $|OF_1|$  and  $|OF_2|$  can be calculated by Eqs. (10) and (11), respectively. Combining

Eqs. (1)–(4) and (9)–(11), it can be concluded that when the TBM model and segment geometry are fixed, the tunnel curve radius, pitch angle, and length of the shield tail covering segments are the key factors affecting the shield tail clearance. In particular, while the pitch angle changes only slightly during excavation, it significantly impacts the shield tail clearance. From Eq. (1), it can be seen that the greater the pitch angle, the greater the variation



**Fig. 3** Schematic diagram of the desirable shield attitude.



**Fig. 4** Schematic diagram of shield tail clearance calculation.

in shield tail clearance. However, the total clearance at relative positions is a fixed value. Therefore, it is crucial to strictly control the pitch angle of the TBM to prevent excessive pitch angles from causing shield tail clearance changes that exceed the available clearance value, which could lead to issues such as TBM jamming or difficulties in segment assembly.

Equation (9) realizes the estimation of clearance by revealing the influence laws of pitch angle and tunnel curve radius on the shield tail clearance. However, the derivation process of the physical equations doesn't consider the influence of other factors on the shield tail clearance, such as cylinder stroke difference and shield tail deviation. As a result, the theoretical analysis methods often fail to accurately calculate the clearance size. To overcome these shortcomings, this study aims to develop an online PDNN model to better forecast and control shield tail clearance.

### 3.2 Online PDNN for shield tail clearance estimation

#### 3.2.1 The architecture of PDNN model

Since the neural network was proposed, some researchers in the civil engineering field have started to explore the potential application of DNN in tunnel-related challenges (Gao et al., 2019; Lin et al., 2023; Zhang et al., 2020b). However, purely data-driven DNN models have the problems of over-reliance on data and a lack of reasonable explanation of the internal mechanism of the physical system (Chakraborty, 2021). A viable strategy for tackling these issues is to incorporate the physical laws into the framework of DNN to derive a novel PDNN algorithm. Specifically, one of the mainstream design ideas of the PDNN algorithm is to integrate the physical mechanisms into the model constraints. Figure 5 illustrates a general framework of physics-informed DNN.

As shown in Fig. 5, the general structure of a DNN consists of three main components, namely, an input layer that receives feature variables, multiple hidden layers for extracting and learning high-level features through nonlinear transformations, and an output layer for producing prediction results. Generally, the feature variables and predicted values can be represented by the vectors  $x = \{x_1, x_2, \dots, x_m\}$  and  $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_r\}$ , where  $m$  represents the number of feature variables and  $r$  represents the number of the output targets. These layers comprising a different number of neurons are fully linked together through weights and biases, and the information is transformed between neurons through a nonlinear activation function. Mathematically, the relationship can be represented as:

$$o_q^s = \sigma(x_q^s) = \sigma(\sum_q w_q^{s-1} o_q^{s-1} + b^{s-1}), \quad (12)$$

where  $x_q^s$  and  $o_q^s$  denote the input and output values of the  $q$ th neuron in the  $s$ th layer, respectively;  $w_q^{s-1}$  represents the weight of the  $q$ th neuron in the  $(s-1)$ th layer;  $b^{s-1}$

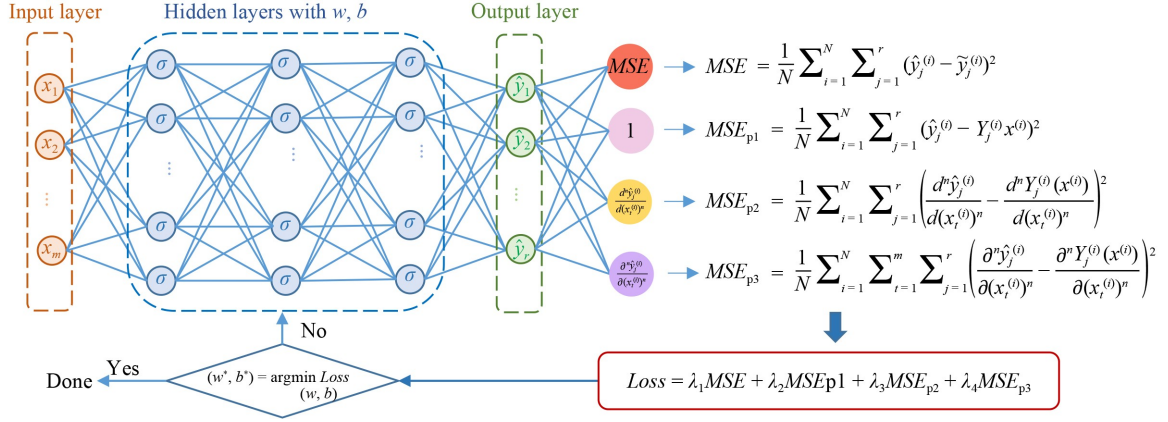


Fig. 5 The general framework of a physics-constrained DNN model.

represents the bias of the  $(s-1)$ th layer;  $\sigma$  is an activation function that acts as a nonlinear mapping. Generally, the values of the hyperparameters  $w_q^{s-1}$ ,  $b^{s-1}$  are estimated automatically by minimizing the mean square error ( $MSE$ ) in the process of training.

$$MSE = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^r (\hat{y}_j^{(i)} - \tilde{y}_j^{(i)})^2, \quad (13)$$

where  $\hat{y}_j^{(i)}$  and  $\tilde{y}_j^{(i)}$  are the predicted value and the real value for the  $j$ th output target of the  $i$ th sample, respectively;  $N$  is the number of collection samples. Compared with the *Sigmoid* and *Tanh* functions, the activation function *ReLU* is not only easy to converge but also solves the gradient vanishing, so it is employed to activate the neurons in this study.

$$\sigma(x_q^s) = ReLU(x_q^s) = \begin{cases} x_q^s, & \text{if } x_q^s \geq 0 \\ 0, & \text{if } x_q^s < 0 \end{cases}, \quad (14)$$

where the final output is determined by the input features, as well as the weights and biases associated with each neuron in the network.

The fundamental idea of the PDNN model is to integrate the physical laws governing the problem into the loss function, thereby constructing a DNN model that incorporates physical constraints. The mathematical equations of fundamental physical laws can be categorized into dependent equations, ordinary differential equations (ODEs), and partial differential equations (PDEs) (Roy et al., 2023), with their physical interpretations varying across different engineering problems. Typically, these equations are used to represent the initial conditions, boundary conditions, and control equations that the problem must satisfy. For instance, Chai et al. (2020), in the task of optimizing the trajectory of a hypersonic vehicle, employed dependent equations and ODEs to describe the boundary conditions of the hypersonic vehicle system and the control equations based on flight kinematics, respectively. Due to the differing mathematical represen-

tations of physical laws, the corresponding physical loss functions are slightly different. As shown in Eqs. (15)–(17),  $MSE_{p1}$ ,  $MSE_{p2}$ , and  $MSE_{p3}$  represent the physical losses associated with the dependent equations, ODEs, and PDEs, respectively.

$$MSE_{p1} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^r (\hat{y}_j^{(i)} - Y_j^{(i)}(x^{(i)}))^2, \quad (15)$$

$$MSE_{p2} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^r \left( \frac{d^n \hat{y}_j^{(i)}}{d(x^{(i)})^n} - \frac{d^n Y_j^{(i)}(x^{(i)})}{d(x^{(i)})^n} \right)^2, \quad (16)$$

$$MSE_{p3} = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^m \sum_{j=1}^r \left( \frac{\partial^n \hat{y}_j^{(i)}}{\partial (x_t^{(i)})^n} - \frac{\partial^n Y_j^{(i)}(x^{(i)})}{\partial (x_t^{(i)})^n} \right)^2, \quad (17)$$

where  $x^{(i)}$  is the feature vector of the  $i$ th sample;  $x_t^{(i)}$  is the  $t$ th feature of the  $i$ th sample;  $\hat{y}_j^{(i)}$  and  $Y_j^{(i)}(x^{(i)})$  are the predicted and theoretical values of the  $j$ th the output for the  $i$ th sample, respectively;  $\frac{d^n \hat{y}_j^{(i)}}{d(x^{(i)})^n}$  and  $\frac{d^n Y_j^{(i)}(x^{(i)})}{d(x^{(i)})^n}$  denote the ordinary differentials of the  $\hat{y}_j^{(i)}$  and  $Y_j^{(i)}(x^{(i)})$  with respect to the  $x_t^{(i)}$ , respectively;  $\frac{\partial^n \hat{y}_j^{(i)}}{\partial (x_t^{(i)})^n}$  and  $\frac{\partial^n Y_j^{(i)}(x^{(i)})}{\partial (x_t^{(i)})^n}$  are the partial differentials of the  $\hat{y}_j^{(i)}$  and  $Y_j^{(i)}(x^{(i)})$  with respect to the  $x_t^{(i)}$ , respectively. The values of  $\frac{d^n \hat{y}_j^{(i)}}{d(x^{(i)})^n}$  and  $\frac{\partial^n \hat{y}_j^{(i)}}{\partial (x_t^{(i)})^n}$  can be obtained using an automatic differentiation technique, while  $\frac{d^n Y_j^{(i)}(x^{(i)})}{d(x^{(i)})^n}$  and  $\frac{\partial^n Y_j^{(i)}(x^{(i)})}{\partial (x_t^{(i)})^n}$  need to be derived by differentiating the physical equations. In summary, the loss function of the PDNN model is expressed as the sum of Eq. (13), and Eqs. (15)–(17):

$$Loss = \lambda_1 MSE + \lambda_2 MSE_{p1} + \lambda_3 MSE_{p2} + \lambda_4 MSE_{p3}, \quad (18)$$

where  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  represent the weight of the  $MSE$ ,  $MSE_{p1}$ ,  $MSE_{p2}$ , and  $MSE_{p3}$ , respectively. The sum of  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  is 1.

The application of physical constraint equations not only helps to reduce the size of the input data set but also enables the output target to satisfy the corresponding physical properties. The PDNN model framework for shield tail clearance measurement after coupling the physical mechanism is shown in Fig. 6. The loss function of the developed PDNN model is formulated as:

$$\begin{aligned} Loss(\hat{y}, \tilde{y}, \hat{y}, Y) &= \lambda_1 Loss_{obs}(\hat{y}, \tilde{y}) + \lambda_2 Loss_{phy}(\hat{y}, Y) \\ &= \lambda_1 \sum_{i=1}^N \sum_{j=1}^r (\hat{y}_j^{(i)} - \tilde{y}_j^{(i)})^2 \\ &\quad + \lambda_2 \sum_{i=1}^N \sum_{j=1}^r (\hat{y}_j^{(i)} - Y_j^{(i)})^2, \end{aligned} \quad (19)$$

where  $\hat{y}$ ,  $\tilde{y}$ , and  $Y$  represent the sets formed by the predicted values, observed values, and theoretical values of the shield tail clearance from all the training samples, respectively; the first part  $\sum_{i=1}^N \sum_{j=1}^r (\hat{y}_j^{(i)} - \tilde{y}_j^{(i)})^2$  represents the data loss  $Loss_{obs}(\hat{y}, \tilde{y})$ , calculated from the predicted value and observed value of shield tail clearance, while the second part  $\sum_{i=1}^N \sum_{j=1}^r (\hat{y}_j^{(i)} - Y_j^{(i)})^2$  represents the physical loss  $Loss_{phy}(\hat{y}, Y)$ , which is derived from the predicted value of shield tail clearance and the theoretical value calculated based on Eq. (9).

Drawing on the research of Chai et al. (2021), which applied physical constraints to optimize overtaking strategies in maneuver planning, the proposed modeling approach that combines DNN with physical laws is expected to improve the model's causal interpretability. Specifically, the model parameters, such as weights and biases, trained using this carefully designed loss function can minimize not only the data-driven component of the loss but also the boundary residual constraints represented by the dependent equations. This helps the model's predicted solutions to approximate the theoretical solution

more closely, resulting in improved alignment between the predicted data set and the observed data set.

The main steps involved in the proposed PDNN are illustrated in Algorithm 1. For training the PDNN, the input variables include shield pitch angle and several other important parameters  $x_{other}$ . The output variables are the magnitude of the shield tail clearance at different positions. Details on the strategy of hyperparameter tuning for model training as well as input and output variables are described in Subsections 3.2.2 and 4.2, respectively. Herein, the Adam optimizer and Xavier initialization are also applied in the training process of PDNN models.

### 3.2.2 Online update mechanism

Although DNN models embedded with physical laws reduce the dependence on data set size and more realistically reflect the engineering mechanisms, offline PDNN models may no longer be valid as the data set is continuously updated. Consequently, we concentrate on developing an online PDNN model in this study, which is significantly different from the conventional PDNN model in the manner of data provided and the automatic updating of the model parameters. Specifically, the data are provided sequentially for the developed online PDNN model, and the hyperparameters are dynamically adjusted as the data set is updated. In contrast, the entire data set is supplied in batches for the conventional offline PDNN model. Once the model has been trained, the parameters of the neural network are frozen, which means that they are not updated as the input data set changes. Figure 7 shows the development process of the online PDNN model.

In the online prediction concept, suppose that the initial data set contains  $N_0$  samples. Then, we choose  $N_1$  of  $N_0$  samples as the training data to train the model, and the

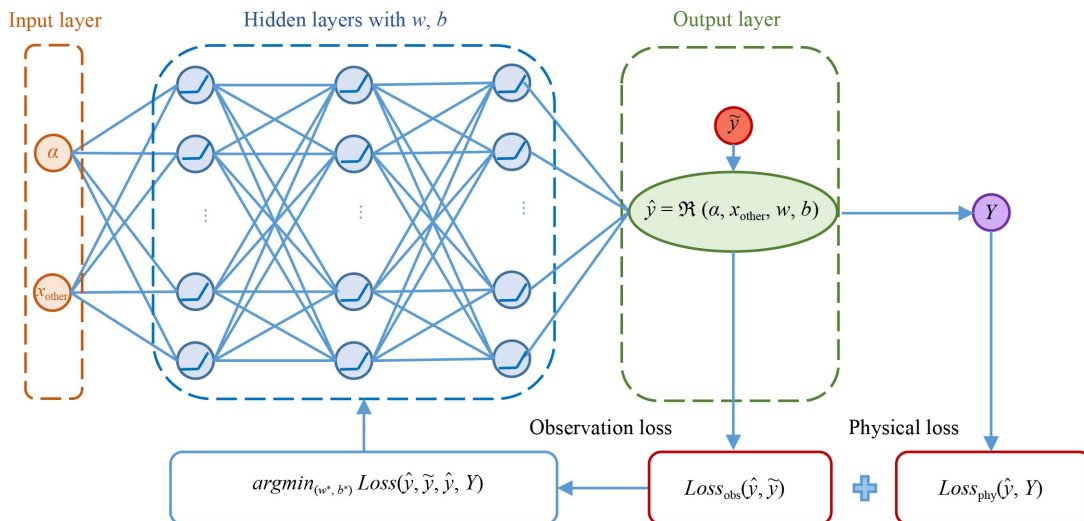
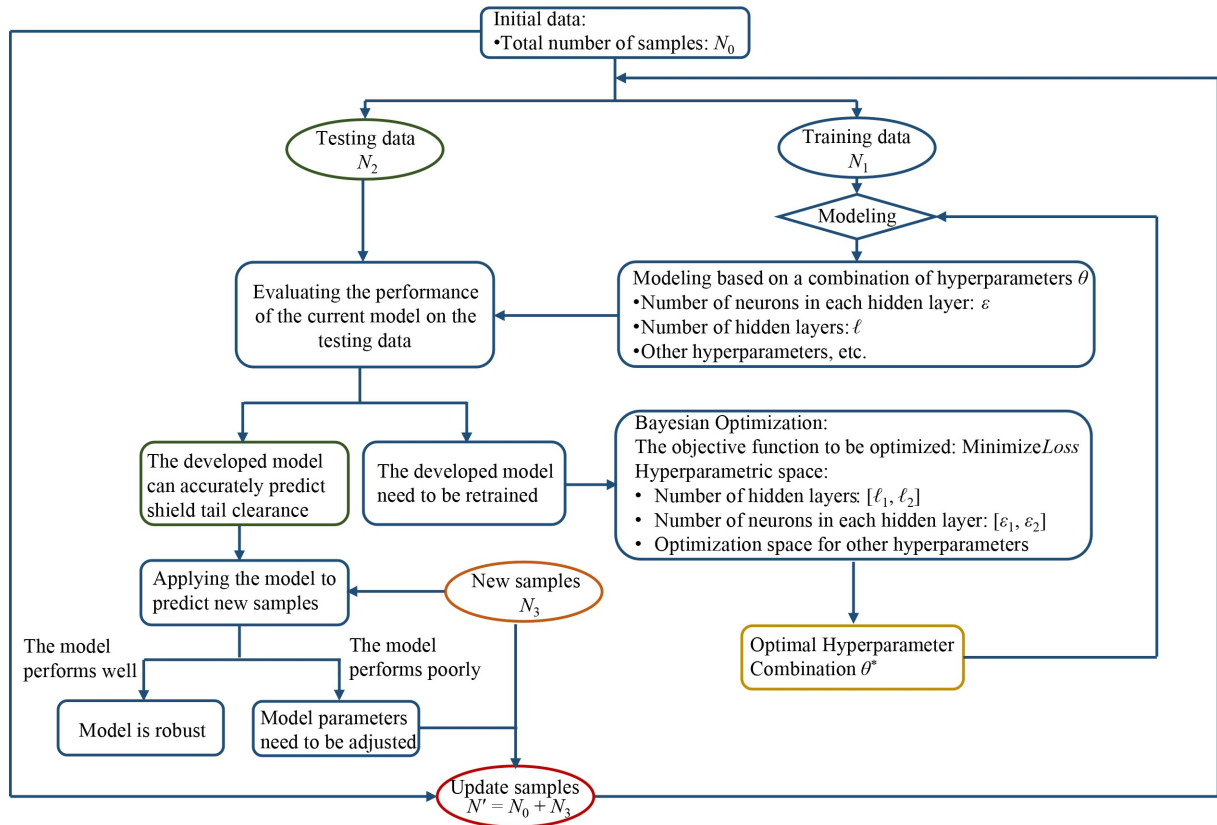


Fig. 6 The framework of the PDNN algorithm for shield tail clearance measurement.

**Algorithm 1** Pseudocode for physics-informed deep neural network algorithm

- 1 Provide training data including pitch angle  $\alpha$ , other input variables  $x_{\text{other}}$ , and the size of the shield tail clearance in a certain direction  $y$
- 2 Specify the structure of the DNN, the number of hidden layers  $\ell$ , and the number of neurons in each hidden layer  $\varepsilon$
- 3 Combine the physical laws of shield tail clearance to formulate a physics-informed loss function
- 4 Express the unknown response  $y = \mathfrak{R}(\alpha, x_{\text{other}}, w, b)$  using the PDNN
- 5 Minimize the physics-informed loss function to find the optimal hyperparameters of the model,  $w^*, b^*$
- 6 Generate the PDNN model with optimal  $w^*, b^*$

**Fig. 7** The development process of an online PDNN model.

remaining  $N_2$  samples as the testing data to validate the model. If the loss function can be minimized on the currently set hyperparameter combination  $\theta$ , it indicates that the current hyperparameter combination  $\theta$  is optimal. Otherwise, hyperparameter optimization algorithms need to be used to find the optimal combination of hyperparameters  $\theta^*$ . To the best of our knowledge, Bayesian Optimization (BO) with Gaussian process priors has good practices in determining the optimal hyperparameters for DL algorithms. Therefore, the BO algorithm is integrated into the proposed online PDNN model aiming to automatically optimize the hyperparameters of the neural

network. A pseudocode for optimizing the hyperparameters based on BO is shown in [Algorithm 2](#). Here are two key aspects to highlight. One is the use of Gaussian processes to construct a probabilistic agent model; the other is the use of expectation increment as an acquisition function to select the next sampling location.

All computations in the DL method can be viewed as nonlinear operators performed by neural networks, where the tunable parameters of the model are continuously computed and updated by minimizing the loss function. Therefore, when new samples  $N_3$  arrive, we need to ascertain whether it performs well with the current

**Algorithm 2** Pseudocode for Bayesian optimization incorporated in online PDNN

**Input:** A PDNN model  $\mathfrak{R}$  established based on the current  $N_0$  (or updated  $N'$ ) samples

Acquisition function  $\mathcal{J}(\theta)$

Probabilistic agent model  $\mathcal{M}$

Tunable hyperparameters space  $\Theta$

$G \leftarrow$  Initialize a set of experimental points using  $\Theta$  and  $\mathfrak{R}$

Maximum number of iterations  $V$

**For**  $v = 1$  to  $V$

$p(\text{Loss}(\hat{y}, \tilde{y}, \hat{y}, Y) | \theta, G) \leftarrow$  Fit agent model ( $\mathcal{M}, G$ )

Select  $\theta' \leftarrow \operatorname{argmax}_{\theta \in \Theta} \mathcal{J}(\theta; p(\text{Loss}(\hat{y}, \tilde{y}, \hat{y}, Y) | \theta, G))$

Evaluate  $\text{Loss}(\hat{y}, \tilde{y}, \hat{y}, Y) \leftarrow \mathfrak{R}(\theta')$

Update  $G \leftarrow G \cup (\theta', \text{Loss}(\hat{y}, \tilde{y}, \hat{y}, Y))$

Check for exit criteria

**end for**

Return  $\theta^*$  with minimum  $\text{Loss}(\hat{y}, \tilde{y}, \hat{y}, Y)$  in  $\Theta$

Update the PDNN model  $\mathfrak{R}$  with  $\theta^*$

model. If so, it indicates that the model has excellent robustness, and we don't need to implement any adjustments to it. Otherwise, we need to recalculate the optimal hyperparameters with the updated data set (containing  $N'$  samples).

### 3.3 Model evaluation and interpretation

Five statistical metrics are employed to assess the performance of the developed online PDNN model, including the  $R^2$ , the variance account for ( $VAF$ ), the  $a20\_index$ , mean absolute error ( $MAE$ ), and root mean square error ( $RMSE$ ), whose calculation formulas are given in Eqs. (20)–(24). These assessment indicators are widely used in the evaluation of regression models. That is to say, the reliability of model quality assessment can be enhanced by utilizing various calculations to assess the differences between predicted and measured values.  $R^2$  reflects the extent to which independent variables contribute to explaining changes in dependent variables.  $a20\_index$  as a criterion of engineering problems quantitatively evaluates the reliability of the regression model with a 20% tolerance (Yavas et al., 2023).  $RMSE$  and  $MAE$  are used to measure the deviation between predicted and measured values. Typically, as the values of  $R^2$ ,  $VAF$ , and  $a20\_index$  get closer to 1, and the values of  $MAE$  and  $RMSE$  get closer to 0, the higher the predictive accuracy of the model.

$$R_j^2 = 1 - \frac{\sum_{i=1}^N (\hat{y}_j^{(i)} - \tilde{y}_j^{(i)})^2}{\sum_{i=1}^N (\tilde{y}_j - \tilde{y}_j^{(i)})^2}, \quad (20)$$

$$VAF_j = 1 - \frac{\sum_{i=1}^N (\hat{y}_j^{(i)} - \tilde{y}_j^{(i)})^2}{\sum_{i=1}^N (\tilde{y}_j - \tilde{y}_j^{(i)})^2}, \quad (21)$$

$$a20\_index_j = \frac{m_j^{20}}{N}, \quad (22)$$

$$MAE_j = \frac{1}{N} \sum_{i=1}^N |\hat{y}_j^{(i)} - \tilde{y}_j^{(i)}|, \quad (23)$$

$$RMSE_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_j^{(i)} - \tilde{y}_j^{(i)})^2}, \quad (24)$$

where  $\tilde{y}_j$  represents the mean of all samples for the  $j$ -th target;  $m_j^{20}$  refers to the number of samples whose error for the  $j$ -th target does not exceed  $\pm 20\%$  of the measured values.

DL algorithms with black-box mechanisms offer advantages due to their intricate network structures, enabling them to effectively fit a given data set. However, this complexity comes at the cost of reduced model interpretability. Understanding how complex models make decisions is not only a relevant problem in data science applications but also a focal point in engineering (Tahmassebi et al., 2022). To address the interpretability problem of online PDNN models, the partial dependence plot (PDP) and Shapley additive explanations (SHAP) analysis methods are introduced in this study, aiming to visualize the prediction results and enhance the model interpretability.

PDP analysis is a powerful technique in ML and statistical analysis, primarily used to understand the relationship between specific input features and model predictions. PDP allows for the examination of how variations in a

selected feature influence the predicted outcomes while holding other features constant. This approach enables a focused analysis of individual or combined effects of features on model behavior, making it particularly useful for assessing feature importance and interpreting complex models. Typically, PDP can be one-dimensional (1D PDP) or two-dimensional (2D PDP), which depends on whether one or two input parameters are varied at a time.

In SHAP, an efficient way is provided for estimating the marginal contribution of each input feature to the model. The Shapley value  $\phi_{x_i}$ , which represents the contribution of the feature  $x_i$ , is determined by:

$$\phi_{x_i} = \sum_{x' \subseteq x \setminus \{x_i\}} \frac{|x'|!(m-1-|x'|)!}{m!} (\mathfrak{R}(x' \cup \{x_i\}) - \mathfrak{R}(x')), \quad (25)$$

where  $x'$  is the feature subset that does not contain  $x_i$ ;  $\mathfrak{R}(x' \cup \{x_i\})$  and  $\mathfrak{R}(x')$  represent the model's outputs on the  $x'$  and the feature set  $x' \cup \{x_i\}$  obtained by merging  $x'$  with  $x_i$ , respectively. SHAP analysis is based on the additive feature attribution method, which can simplify a complex model into a linear function:

$$\xi(x_i) = \phi_0 + \sum_{t=1}^m \phi_{x_t} x_t, \quad (26)$$

where  $\phi_0$  is computed as the mean of the estimated values; the coalition vector  $x_t \in \{0, 1\}^m$  describes whether feature  $x_t$  is observed.  $x_t = 0$  and  $x_t = 1$  represent the feature is absent or present, respectively.

### 3.4 Multi-objective optimization for shield tail clearance control

To avert excessive extrusion between the shield shell and the concrete lining, it is necessary to consider multi-objective optimization and control for the shield tail clearance. Generally, a minimization MOO problem needs to consider four components, including the  $r$ -dimensional target vector  $y(x)$ , the  $M_{\text{ineq}}$  inequality constraints,  $M_{\text{eq}}$  equality constraints, and the  $m$ -dimensional decision vector  $x$ , whose mathematical representations are presented by Eqs. (27)–(29).

$$\text{Minimize } y(x) = (y_1(x), y_2(x), \dots, y_r(x)), r \geq 2, \quad (27)$$

$$\text{subject to } g_\rho(x) \leq 0, \rho = 1, 2, \dots, M_{\text{ineq}}, \quad (28)$$

$$h_\mu(x) = 0, \mu = 1, 2, \dots, M_{\text{eq}}$$

$$x = (x_1, x_2, \dots, x_m) \in \Omega, \quad (29)$$

where  $y_r(x)$  is the  $r$ -th objective function;  $g_\rho(x)$  represents the  $\rho$ -th inequality constraint;  $h_\mu(x)$  is the  $\mu$ -th equality constraint;  $\Omega$  is the decision space of independent variables. In view of this, the overall objective of the shield tail clearance optimization problem in this study is to determine appropriate ranges for key TBM mechanic parameters by minimizing the shield tail clearance in

different directions. This approach helps to prevent excessive extrusion between the shield shell and the concrete linings, thus reducing safety hazards and avoiding engineering accidents. It is worth noting that the optimized shield tail clearance is not only constrained by the geometries of TBM shield shell and lining rings, but also must not fall below the minimum clearance specified by industry standards. Referring to the literature (Han, 2014), this study sets the warning threshold for shield tail clearance at 10 mm.

According to the existing literature (Chai et al., 2017), the NSGA-II algorithm has good applicability to solve a MOO problem. Compared to NSGA, the main innovation of the NSGA-II algorithm adopts three special principles, including nondominated sorting, crowding distance, and crowd comparison operator, which significantly reduces the complexity of the algorithm and speeds up the computation. In the nondominated sorting procedure, the predominant concept is introduced to search for the optimal solutions to the MOO problem. For example, assuming  $x^{\tau 1} = \{x_1^*, \dots, x_m^*\}$  and  $x^{\tau 2} = \{x_1, \dots, x_m\}$  be two solution vectors of MOO problem (Eq. (27)),  $x_i^*$  dominates  $x_i$  (denoted as  $x_i^* < x_i$ ) if  $x_i^* \leq x_i$  for all  $t = 1, 2, \dots, m$  and  $x_i^* < x_i$  for at least one  $t = 1, 2, \dots, m$ . When a solution vector  $x$  of Eq. (27) is not dominated by any other solution vectors, it is called Pareto optimal. The set of all Pareto optimal solutions is denoted as a Pareto set.

Once the sorting is complete, the crowding distance will be considered to maintain the diversity and uniform distribution of the solution set. The crowded distance of the  $\tau$ -th solution is formulated as follows:

$$T = \sum_{j=1}^r \frac{y_j^{\tau+1} - y_j^{\tau-1}}{y_j^{\max} - y_j^{\min}}, \quad (30)$$

where  $y_j^{\tau+1}$  and  $y_j^{\tau-1}$  present the  $j$ th objective function values of the  $(\tau+1)$ th and  $(\tau-1)$ th individuals, respectively;  $y_j^{\max}$  and  $y_j^{\min}$  refer to the maximum and minimum values of the  $j$ th objective function.

After the fast nondominated ordering and calculations of crowding distance are completed, all population members in the set are assigned two attributes, which are nondomination rank and crowding distance. Based on these two properties, we can utilize the crowded comparison operator to perform the comparison of two solutions. Specifically, if two solutions are of different ranks, we desire to select the solution with the lower (or better) rank. If two solutions are of the same front, the solution with the larger crowding distance is preferred.

The aforementioned steps are iteratively repeated until the maximum number of iterations is reached, at which point the algorithm is terminated, and the Pareto front is generated. To determine the best solution from the Pareto front, it is necessary to calculate the distance between the Pareto optimal solution vector and the ideal solution vector by Eq. (31), where the ideal solution vector is

composed of the optimum for all objective functions.

$$d_{\tau} = \sqrt{\sum_j^r (y - o_j)^2}, \quad (31)$$

where  $y$  is the value of the  $j$ th objective function in the  $\tau$ th Pareto solution;  $o_j$  is the ideal value of the  $j$ th objective function. Compare the values of  $d_{\tau}$  calculated by different Pareto solutions, and choose the minimum of them as the best solution.

The pseudocode for optimizing the shield tail clearance based on the hybrid approach of online PDNN and NSGA-II is shown in Algorithm 3. The main steps are as follows: First, select multiple conflicting objectives (minimizing shield tail clearance in different directions) and set appropriate parameter ranges and constraints. Then, apply NSGA-II for MOO to generate a Pareto front. In this process, the online PDNN model acts as a surrogate model to quickly predict shield tail clearance values. Finally, by calculating the distance between each Pareto optimal solution and the ideal point, the best solution with evenly reduced shield tail clearance across all directions can be identified.

## 4 Case study

### 4.1 Background

To verify the effectiveness of the proposed online PDNN

model, we have applied it to the tunneling project of Wuhan Metro Line 12 (i.e., the first loop line of Wuhan Metro in China). The target section focused on in this case is the single-bore, two-lane tunnel between the Guobo Center South Station and Lingwucun Station. The design length and the segment width of the tunnel are 3,373.67 m and 2 m, respectively. For the excavation of this tunnel, a mud-water balanced TBM with a cutter diameter of 12.56 m is employed. The engineering floor plan, lining rings, and photos of the TBM's data acquisition system are shown in Fig. 8(a)–8(c), respectively.

During the tunnel excavation process, the data acquisition system is used to monitor the performance parameters of the TBM in real-time, which can contribute to ensuring the safety and controllability of construction. Out of all the parameters collected from the TBM, 25 machine parameters closely associated with shield tail clearance are relatively easy to obtain, such as shield attitude parameters, cylinder stroke, chamber pressure, etc. Accordingly, these parameters are considered input features of the model to reliably forecast the shield tail clearance. A detailed description of all TBM operational parameters is given in Section 4.2.

In this case study, the data set is taken from the first 553 concrete lining rings (Ring No. 0–552) with a diameter of 12.1 m and a width of 2 m, respectively. To mitigate the risk of segment damage caused by minimal local shield tail clearance, the onsite workers measured the

---

**Algorithm 3** Pseudocode for optimizing shield tail clearance with online PDNN and NSGA-II

---

**Input:** Operational parameters  $x = (\alpha, x_{\text{other}})$ ; Developed online PDNN model  $\mathfrak{R}$

**Optimization:**

1. Define a MOO problem:

Minimize shield tail clearance  $y$ ;

Define the lower and upper limit of each parameter in  $x$ ,  $x_t^l \leq x_t \leq x_t^u$ ,  $t = 1, 2, \dots, m$ .

2. Initialize the population  $P^0$ , and set the iteration number  $\zeta = 1$

3. Calculate the value of the fitness function

4. Perform crossover and mutation operations for the first-generation population

5. Integrate the parent and offspring population

6. Perform nondominated sort

7. Calculate crowding distance

8. Produce a new population from integrated populations  $P^1$

9. Select the new parent population based on the crowded comparison operator

**If**  $\zeta < \zeta_{\text{max}}$  :

$\zeta = \zeta + 1$

Repeat procedures 4-9

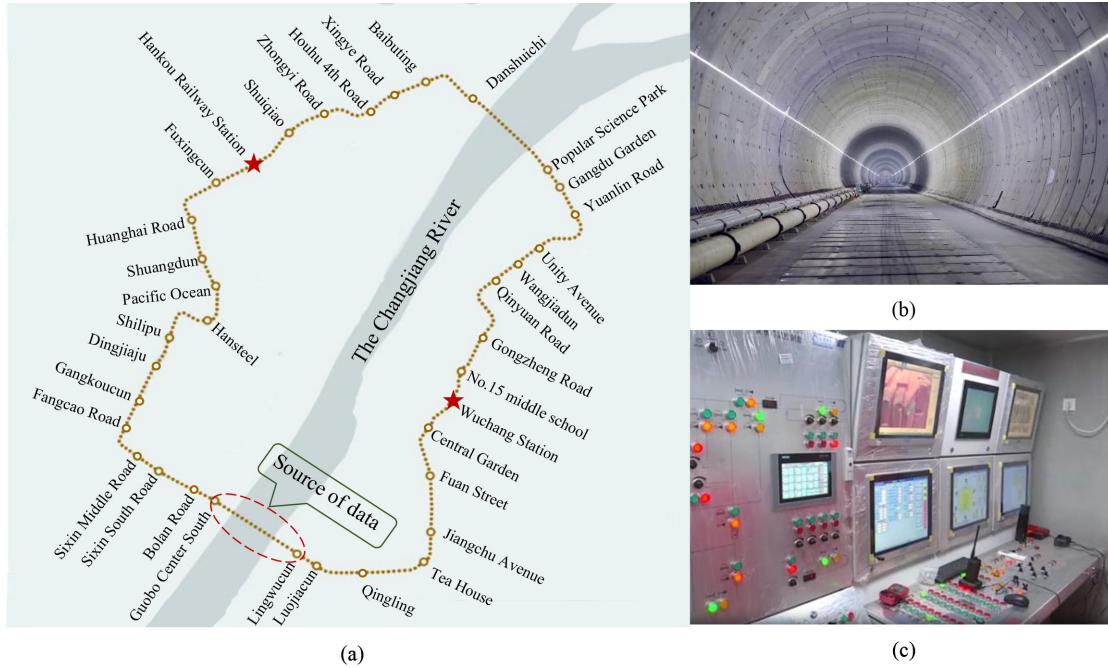
**else:**

**Return** a set of Pareto optimal solutions

Determine the best solution by calculating the distance from the ideal point

**Output:** Optimized parameters  $x_{\text{opt}}$ ; Optimized objectives  $y_{\text{obj}}$

---



**Fig. 8** Project profile: (a) The layout plan of Wuhan Metro Line 12, (b) Lining rings (CSTNET, 2021), and (c) TBM's data acquisition system (WMG, 2021).

shield tail clearance at three specific positions, including directly above, above left ( $\beta = 135^\circ$ ), and above right ( $\beta = 45^\circ$ ), which are represented by  $y_1$ ,  $y_2$ , and  $y_3$ , respectively. As a result, the acquired data set comprises 553 rows and 28 columns, with each row containing the aforementioned 25 input features and 3 output targets. It is worth noting that, for the sake of consistency in variable naming, the pitch angle  $\alpha$  will be replaced with  $x_1$ .

#### 4.2 Data resources

Table 1 provides a detailed overview of the 25 input features and 3 output targets, with their numerical distribution shown in Fig. 9. It can be observed that the data set is well-distributed in the sample space, with no extreme or anomalous data points. To clarify whether there is a linear correlation between all input features and each output target, Fig. 10 illustrates the results of the Mantel test. Based on Mantel's  $r_{\text{men}}$  and  $p_{\text{men}}$  values calculated from the feature and target matrices, it is clear that 25 input features play a significant role in determining the magnitude of the shield tail clearance. Nevertheless, the Pearson values between all input features and each output target are less than 0.75, which indicates that there is no strong correlation present in the prepared data set. In other words, it is almost impossible to accurately predict the shield tail clearance by relying solely on manual analysis. In light of this, this study aims to utilize DL technology to fully exploit the potential value of the acquired data to realize the dynamic prediction and control of the shield tail clearance.

As shown in Table 1, due to the dimensions of each feature being different, the data of different input variables vary greatly. However, this difference in attributes may have an adverse effect on the convergence and prediction of the model. Thus, we perform Min-Max normalization on the data set before developing a DL model. That is to say, without changing the distribution of the original data, the original data are mapped to the interval of  $[0,1]$  by Eq. (32).

$$x_{\text{normal}} = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}, \quad (32)$$

where  $x_{\text{normal}}$  is the normalized value,  $\min(x_i)$  and  $\max(x_i)$  denote the minimum and maximum values of the feature  $x_i$  in the data set, respectively. After data preprocessing is completed, gray relational grade (GRG) is applied to further illustrate the rationality of the selected input features. The definition of GRG is given in Eq. (33), and the result is shown in Fig. 11.

$$\gamma(x_{\text{ref}}(k), x_{\text{comp}}(k)) = \frac{1}{Q} \sum_{k=1}^Q \frac{\min_{\text{comp } k} |x_{\text{ref}}(k), x_{\text{comp}}(k)| + \zeta \max_{\text{comp } k} |x_{\text{ref}}(k), x_{\text{comp}}(k)|}{|x_{\text{ref}}(k), x_{\text{comp}}(k)| + \zeta \max_{\text{comp } k} |x_{\text{ref}}(k), x_{\text{comp}}(k)|}, \quad (33)$$

where  $x_{\text{ref}}$  is the reference sequence;  $x_{\text{comp}}$  refers to the comparability sequence;  $k$  is an indicator used to represent the time period;  $\zeta$  is the distinguishing coefficient,  $\zeta \in (0, 1)$ . As presented in Fig. 11, the GRG values

**Table 1** Details of the prepared data set for shield tail clearance evaluation

Variable	Description	Min	Max	Mean	Std	Unit
x1	Pitch angle	-31.60	4.10	-21.83	5.70	°
x2	Horizontal trend	-7.10	6.30	1.87	2.23	mm
x3	Cut horizontal deviation	-248	66	-22.60	57.29	mm
x4	Shield tail horizontal deviation	-268	69	-43.04	58.97	mm
x5	Cut vertical deviation	-205	54	-19.74	43.49	mm
x6	Shield tail vertical deviation	-256	61	-41.92	55.64	mm
x7	Vertical trend	-6.20	7.50	2.11	2.61	mm
x8	Chamber pressure (CP)	0.25	4.27	3.10	0.47	bar
x9	Advance rate (AR)	1	41.10	19.57	7.99	mm/min
x10	Cutter head rotational speed (CHD RPM)	-0.83	3.13	1.16	0.46	r/min
x11	Cutter head torque (CHD torque)	1067	11352	5311	1865	kN·m
x12	Thrust force (TF)	17814	129971	75631	16127	kN
x13	Total extrusion force (TEF)	563	13451	5669	3547	kN
x14	Starting stroke of group A cylinder	492	808	619	43	mm
x15	Starting stroke of group B cylinder	533	808	646	44	mm
x16	Starting stroke of group C cylinder	533	821	659	45	mm
x17	Starting stroke of group D cylinder	268	831	645	49	mm
x18	Starting stroke of group E cylinder	245	841	625	48	mm
x19	Starting stroke of group F cylinder	215	932	609	49	mm
x20	Ending stroke of group A cylinder	596	2799	2617	95	mm
x21	Ending stroke of group B cylinder	2203	2802	2648	48	mm
x22	Ending stroke of group C cylinder	2380	2857	2664	49	mm
x23	Ending stroke of group D cylinder	2527	3664	2656	62	mm
x24	Ending stroke of group E cylinder	2317	3668	2636	63	mm
x25	Ending stroke of group F cylinder	2467	2807	2613	42	mm
x1	Shield tail clearance at directly above	35	95	68.71	10.20	mm
x2	Shield tail clearance at the upper left	30	94	60.53	10.50	mm
x3	Shield tail clearance at the upper right	40	93	64.26	9.47	mm

Note: Min, Max, Mean, and Std represent the minimum, maximum, mean, and standard deviation of the observed values for each TBM operational parameter.

between the selected features and the output targets are all greater than 0.5, which demonstrates using the mentioned 25 features to train the DL model for shield tail clearance estimation is reliable. In particular,  $x1-x7$  and  $x14-x25$  tend to have a more direct impact on the variation tendency of the shield tail clearance, while the features  $x8-x13$ , which are directly related to TBM operation, have some indirect relationships with the shield tail clearance.

#### 4.3 Implementation details

In terms of small-scale data sets, DNN models trained using traditional methods usually exhibit worse performance than machine learning models (Feng et al., 2019). To break this inherent limitation, an online version of the physics-informed DL model is developed in this study,

which is expected to have higher accuracy and better generalization performance. According to Subsection 4.2, the number of selected feature variables is 25, and the total number of samples collected is 553. However, it is considered that the shield attitude of the current ring may be affected by the shield attitude of the adjacent ring. For this reason, when evaluating the shield tail clearance of the current ring, not only the feature data of the current ring should be taken into account, but also the feature and target data of the adjacent ring should be input into the model as feature variables. Thus, the number of samples used to train models is 552, and the data composition of each sample is 50 feature variables and 3 target variables.

According to Eq. (9), it is evident that the values of the parameter  $\beta$  differ when measuring clearances in different orientations, which makes the online PDNN models used to forecast the shield tail clearances at different locations

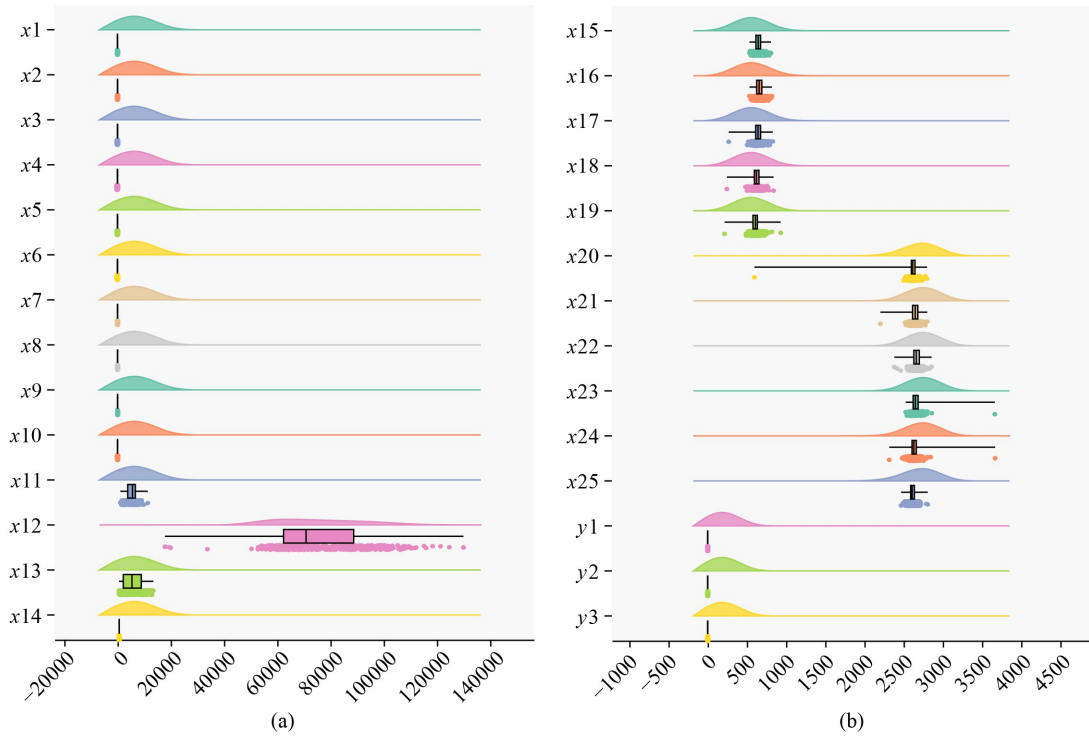


Fig. 9 The numerical distributions of (a) the feature variables  $x_1$ – $x_{14}$  and (b) the feature variables  $x_{15}$ – $x_{25}$  and target variables  $y_1$ – $y_3$ .

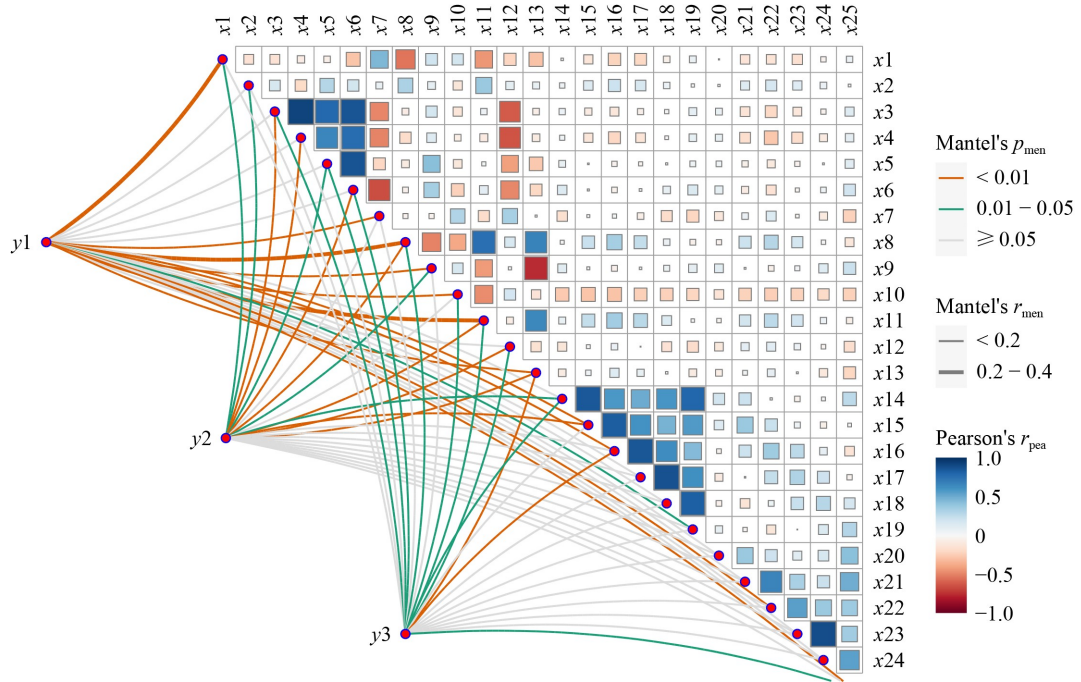


Fig. 10 Pearson correlation analysis between features and objectives.

have different loss functions. Therefore, for  $y_1$ ,  $y_2$ , and  $y_3$ , we design separate online PDNN models. The development process of the online PDNN models for predicting the shield tail clearance is as follows:

The first 501 samples are taken out from the 552 samples after preprocessing, and 498 of these samples are

randomly selected as the training data, while the remaining samples are used as the testing data. Considering the complexity of the input features and the scale of the sample set, the initial parameters of the online PDNN model are set as follows: 2 hidden layers, the number of neurons in each hidden layer, and the number of iterations

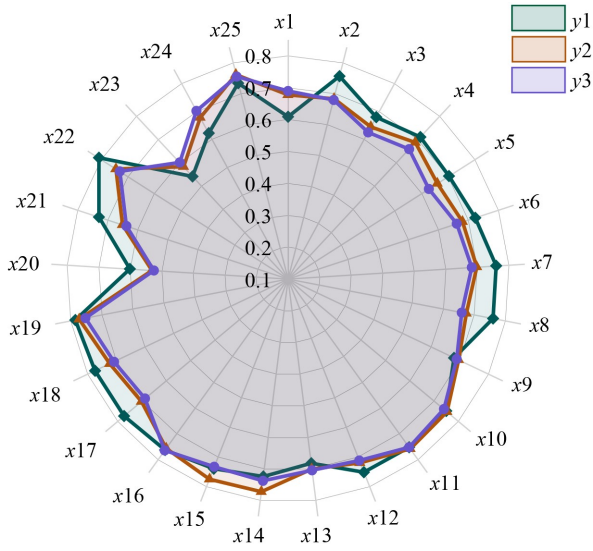


Fig. 11 GRG values between features and objects.

after each data set update are shown in Fig. 12, the learning rate is 0.001, the physical loss weight is 0.1, and a *ReLU* function is used to activate the neurons. In addition, the Adam optimizer is used for the optimization of the coefficients of the neural network to improve the convergence speed and performance of the model.

To assess the reliability of the online PDNN method, the model is applied to the testing data, and the performance of the model is evaluated using five evaluation indicators:  $R^2$ , *VAR*, *a20\_index*, *MSE*, and *RMSE*. Based on the resulting values of these evaluation metrics, if the model performs well, the model can be directly applied to predict newly arrived samples. Otherwise, it is required to retrain the model based on the optimal hyperparameter combination searched by the BO algorithm. The ideal result is that the initial model’s prediction accuracy on new samples remains high, which proves the model established has strong generalization performance. Nevertheless, if the model does not match the new samples well, it is necessary to integrate the new samples into the initial data set to obtain an updated data set. The

updated data set then is divided into the training data and testing data for model training and evaluation. The process is repeated until there are no new samples available. The constructing process of the data set used to train the online PDNN is depicted in Fig. 13. As an example, Fig. 14 presents the training loss of online PDNN models on (a)  $y_1$ , (b)  $y_2$ , and (c)  $y_3$  after the 1st, 6th, 12th, and 18th data set updates. It can be seen that the training loss decreases rapidly at the first 100 iterations and then gradually reaches a steady value.

After training the online PDNN model, it is integrated with the NSGA-II algorithm for shield tail clearance optimization. Based on the correlation between different TBM machine parameters,  $x_1$  to  $x_{25}$  are considered adjustable decision variables. However, the historical data from the previous moment are regarded as unadjustable fixed factors, such as the shield tail clearance  $y_1$ ,  $t-1$ , the starting stroke of group C cylinder  $x_{16}$ ,  $t-1$ , and the shield tail horizontal deviation  $x_4$ ,  $t-1$  from the previous moment. To clarify the impact of the adjustable parameter range on the optimization of the shield tail clearance, this study sets five constraint conditions, corresponding to 10%, 20%, 30%, 40%, and 50% of the up-down range of the current adjustable parameter self-values.

As is well-known, reasonable hyperparameters cannot only yield better optimization results but also significantly improve optimization efficiency. Therefore, Table 2 provides the hyperparameter analysis of the NSGA-II algorithm under the 20% constraint condition. It can be observed that while enlarging the population size appropriately can increase the hypervolume (*HV*) and overall improvement percentage of the Pareto solution set, it also increases the training time. For instance, when the population size increases from 50 to 100, the training time increases by 4.18 s, but the overall improvement percentage only improves by 0.37%. Compared to population size, mutation and crossover probabilities have a smaller effect on training time. Overall, the diversity of the Pareto solution set improves with an increase in these two parameters. Based on these observations, the population size is set to 50, with both crossover and mutation probabilities

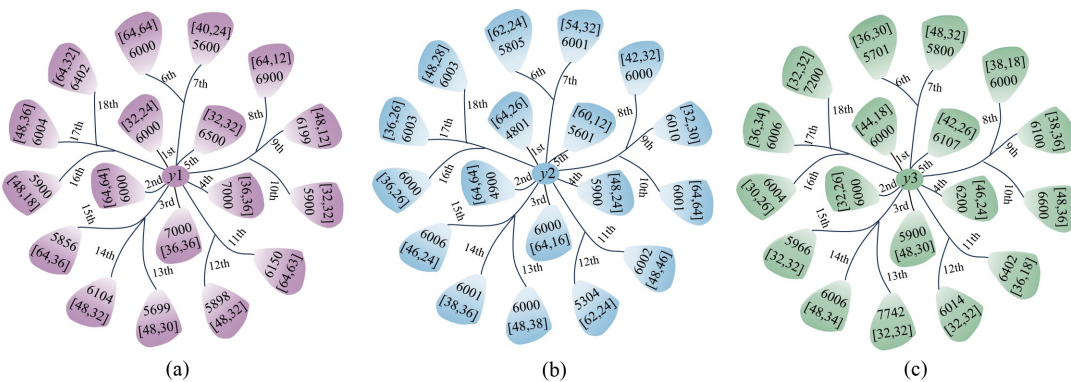


Fig. 12 Hyperparameters of online PDNN models on (a)  $y_1$ , (b)  $y_2$ , and (c)  $y_3$  after each data set update.

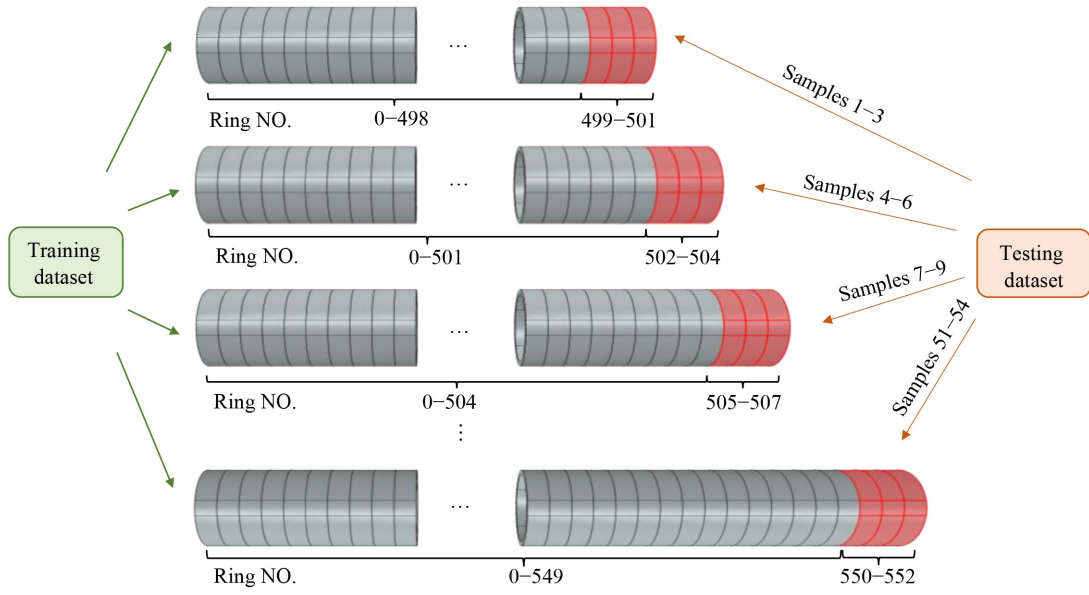


Fig. 13 Construction of the data set used to train the online PDNN model.

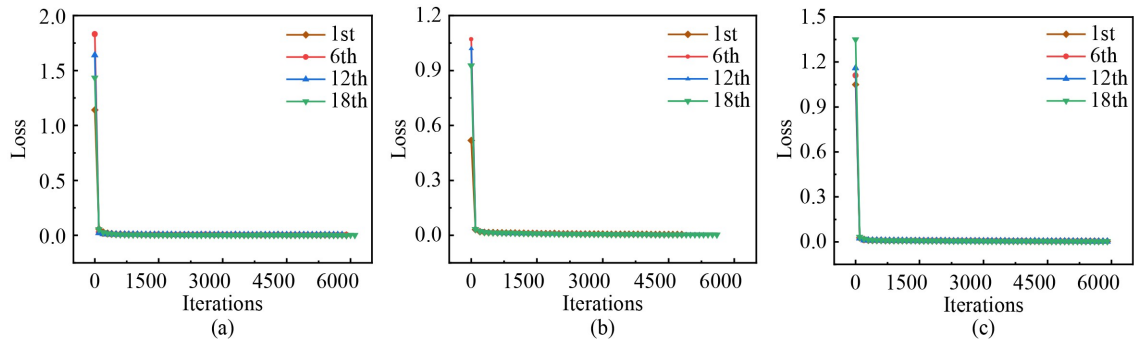


Fig. 14 Training loss of online PDNN models on (a)  $y_1$ , (b)  $y_2$ , and (c)  $y_3$  after the 1st, 6th, 12th, and 18th data set updates.

Table 2 Hyperparameter analysis of NSGA-II

Parameters	Overall optimization percentage	HV	Time (s)
[Population size, Mutation rate, Crossover rate]	14.34%	17	12.50
[20, 0.9, 0.9]	30.87%	32	14.70
[50, 0.9, 0.9]	31.24%	32	18.88
[100, 0.9, 0.9]	29.73%	31	14.52
[50, 0.5, 0.9]	30.09%	30	14.68
[50, 0.8, 0.9]	28.43%	23	16.49
[50, 0.9, 0.5]	28.95%	33	15.57
[50, 0.9, 0.7]			

set to 0.9 in this experiment. Additionally, this study employs parallel computing techniques and sets the convergence tolerance to  $1 \times 10^{-6}$  and the maximum number of iterations to 1000, which effectively reduces computation time.

#### 4.4 Results analysis

The proposed online PDNN is suitable for online regression

tasks on small sample data sets. It has demonstrated a strong capability in modeling the nonlinear relationship between TBM machine parameters and shield tail clearance. To demonstrate the superiority of the online PDNN, we adopt the traditional data set partitioning method (training set: testing set = 9:1) for training DNN and PDNN models. The performance comparison of online PDNN with the DNN and PDNN is shown in Table 3. To understand the effect of incorporating the physical laws into the algorithm, we used the SHAP and PDP methods to analyze the contribution of features, particularly the impact of pitch angle on model output  $y_1$ . Detailed results are analyzed as follows:

(1) The proposed online PDNN method can accurately predict the shield tail clearance with  $R^2$  scores higher than 0.9. As shown in Table 3, the online PDNN model performs best when predicting the shield tail clearance directly above ( $y_1$ ), achieving the lowest  $RMSE$  (2.97) and  $MAE$  (2.52), as well as the highest  $VAF$  (0.96) and  $R^2$  scores (0.93). For the upper left shield tail clearance ( $y_2$ ) and the upper right shield tail clearance ( $y_3$ ), the online PDNN model shows similar predictive performance.

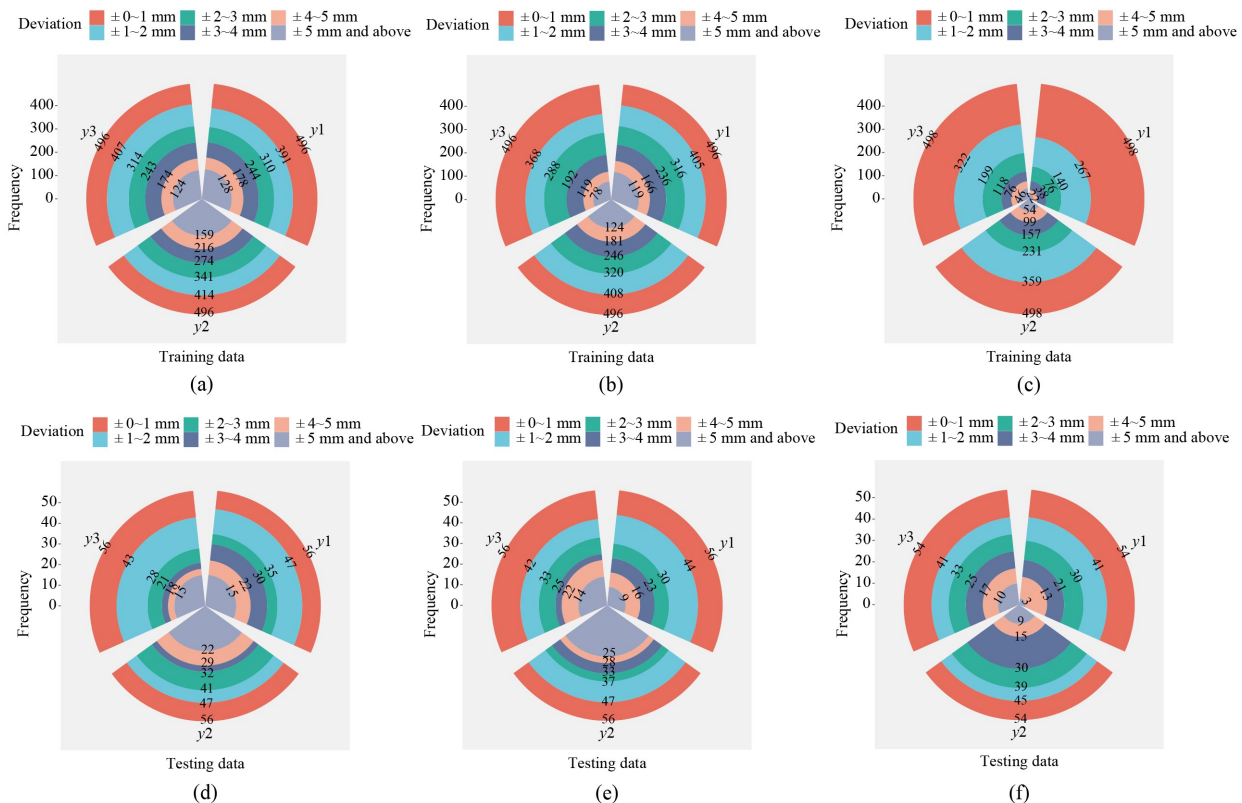
Specifically, both have an *RMSE* of 3.64, *MAE* values of 3.17 and 3.00, *VAF* values of 0.93 and 0.94, and  $R^2$  scores of 0.90. In summary, these low error metric values and high  $R^2$  scores indicate that the online PDNN method performs reliably. In addition, it can be observed that the *a20\_index* values of online PDNN models reach 1 when predicting the shield tail clearance at all three specified locations, which indicates that the proposed method has physical engineering significance.

(2) Compared with offline DNN and PDNN approaches, the online PDNN has better prediction

results. As presented in Table 3, the *a20\_index* metric values obtained based on the different DNN algorithms are close to 1, demonstrating that both the original and improved DL methods are reliable with 20% tolerance. From the four metrics of *RMSE*, *MAE*, *VAF*, and  $R^2$ , the prediction performance of the online PDNN significantly outperforms traditional DNN and PDNN methods. For example, the online PDNN exhibited a decrease of 2.66 mm in *RMSE* and 1.73 mm in *MAE*, coupled with an improvement of 0.18 in *VAF* and 0.14 in  $R^2$  compared with the second-best model named PDNN in forecasting

**Table 3** Performance comparison of the physics law and different DNN models

Target	Model	Training Data					Testing Data				
		<i>RMSE</i> /mm	<i>MAE</i> /mm	<i>VAF</i>	<i>a20_index</i>	$R^2$	<i>RMSE</i> /mm	<i>MAE</i> /mm	<i>VAF</i>	<i>a20_index</i>	$R^2$
y1	DNN	4.66	3.55	0.80	1.00	0.79	4.69	3.55	0.77	1.00	0.77
	PDNN	4.39	3.44	0.82	0.99	0.82	4.17	3.09	0.82	0.98	0.81
	Online PDNN (Average)	2.42	1.81	0.94	1.00	0.94	2.97	2.52	0.96	1.00	0.93
y2	DNN	6.12	4.32	0.65	0.96	0.65	6.71	4.89	0.64	0.95	0.64
	PDNN	5.66	3.98	0.70	0.95	0.7	6.36	4.92	0.68	0.95	0.68
	Online PDNN (Average)	2.89	2.11	0.92	1.00	0.92	3.64	3.17	0.93	1.00	0.90
y3	DNN	4.71	3.59	0.75	0.98	0.75	4.82	3.5	0.73	0.98	0.73
	PDNN	3.96	2.92	0.83	0.99	0.83	4.36	3.32	0.79	0.98	0.77
	Online PDNN (Average)	2.76	2.04	0.92	1.00	0.92	3.64	3.00	0.94	1.00	0.90



**Fig. 15** Deviation of measurements using DNN (Subgraph (a) and (d)), PDNN (Subgraph (b) and (e)), and online PDNN (Subgraph (c) and (f)) methods on training and testing data.

the upper left shield tail clearance. Furthermore, Fig. 15 demonstrates the shield tail clearance measurement deviations of the three models on the training and testing data. It can be seen that the online PDNN model can significantly increase the number of samples with a prediction deviation of less than 5 mm (an average increase of 14.99% over the PDNN), which again demonstrates the superiority of the online PDNN approach.

(3) The online PDNN shows strong robustness, with its  $R^2$  scores remaining at 0.9 and above after each addition of new samples to the original data set. Figure 16 exhibits the values of the five evaluation metrics for the online PDNN model on the training and testing data after each data set update. The online PDNN consistently achieves values above 0.90 for  $VAF$ , 1 for  $a20\_index$ , and 0.90 for the  $R^2$ , which indicates that the method has better generalization performance in adapting to new samples. Figure 17 illustrates the SHAP analysis result of the pitch angle. As presented in Fig. 17(a), the pitch angle ranks

second among the 51 feature variables contributing to the output  $y1$ . Figure 17(b) demonstrates a positive correlation between the pitch angle and the corresponding Shap value. To further explore the contribution of the interactions between input variables to shield tail clearance prediction, Fig. 18 presents a 2D PDP to reveal the interactions among the top three factors. All input feature values in the 2D PDP are standardized, and the contour lines represent the shield tail clearance directly above. For instance, in Fig. 18(a), when the standardized value of the previous moment's shield tail clearance  $y1, t-1$  is 0.07 and the pitch angle  $x1$  is 0.9, the model can minimize  $y1$  to 50 mm during the testing phase.

(4) The hybrid approach of online PDNN and NSGA-II performs well in searching for the best solution for MOO of shield tail clearance, with the optimal variability space of the adjustable parameters found to be in the 20% self-values up-down range. Figures 19 and 20 show optimized objectives values and the Pareto fronts obtained

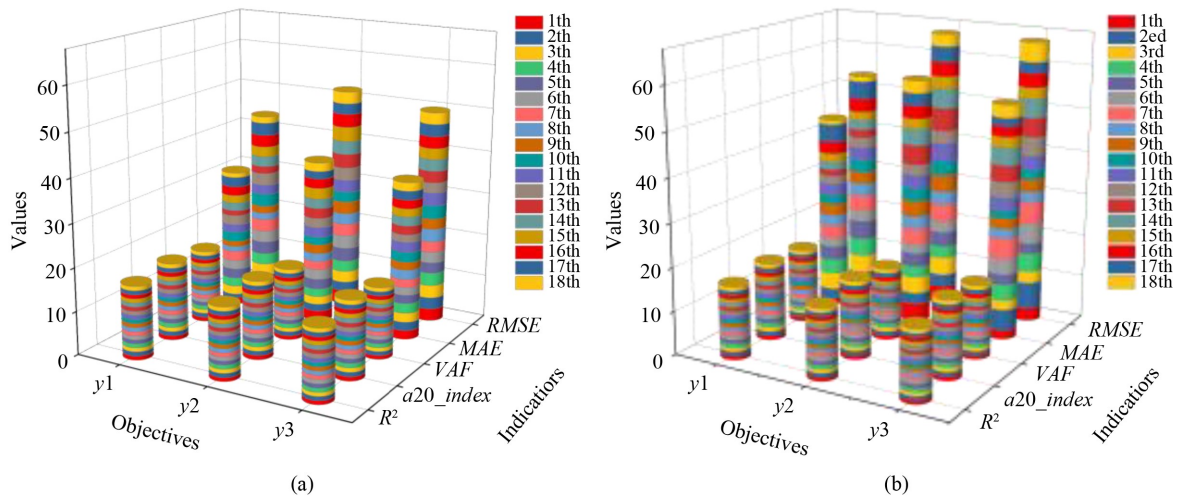


Fig. 16 Performance evaluation of online PDNN models on (a) training and (b) testing data after each data set update.

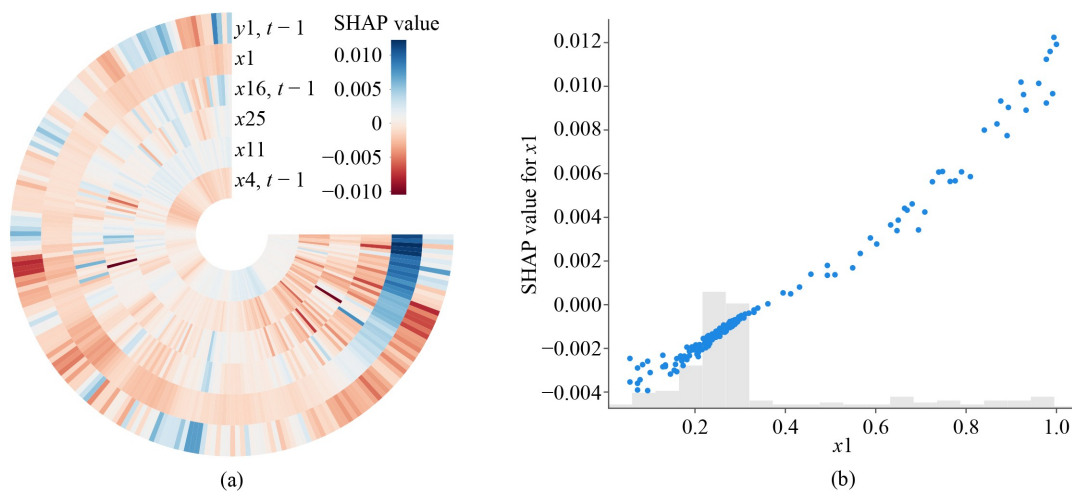
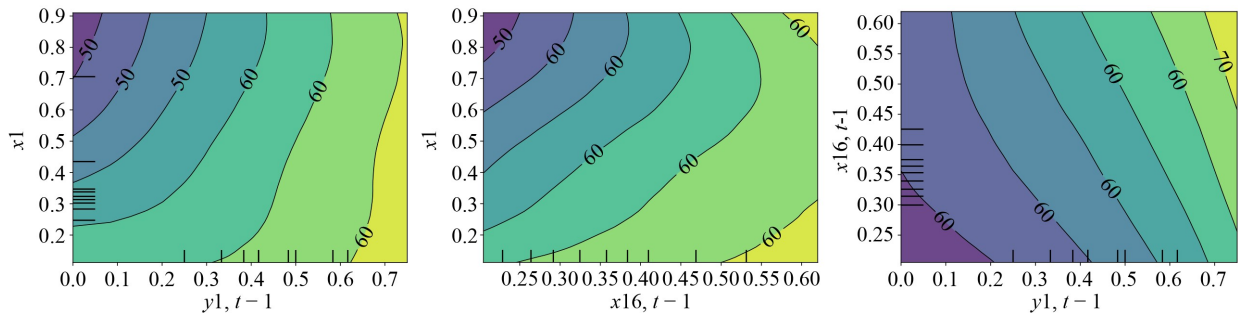
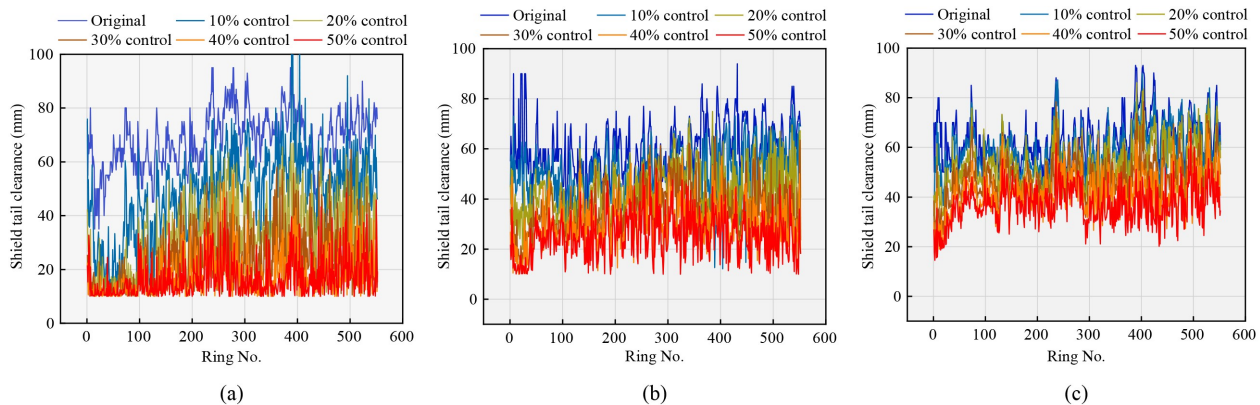


Fig. 17 SHAP analysis result of pitch angle in model trained by online PDNN, including (a) the heat map and (b) the dependency plot.



**Fig. 18** 2D PDPs of  $y_1$  with the top three parameters in accordance with their largest average SHAP value: (a)  $x_1$  and  $y_1, t-1$ , (b)  $x_1$  and  $x_{16}, t-1$ , and (c)  $y_1, t-1$  and  $x_{16}, t-1$ .



**Fig. 19** The optimized values and the original values for (a)  $y_1$ , (b)  $y_2$ , and (c)  $y_3$ .

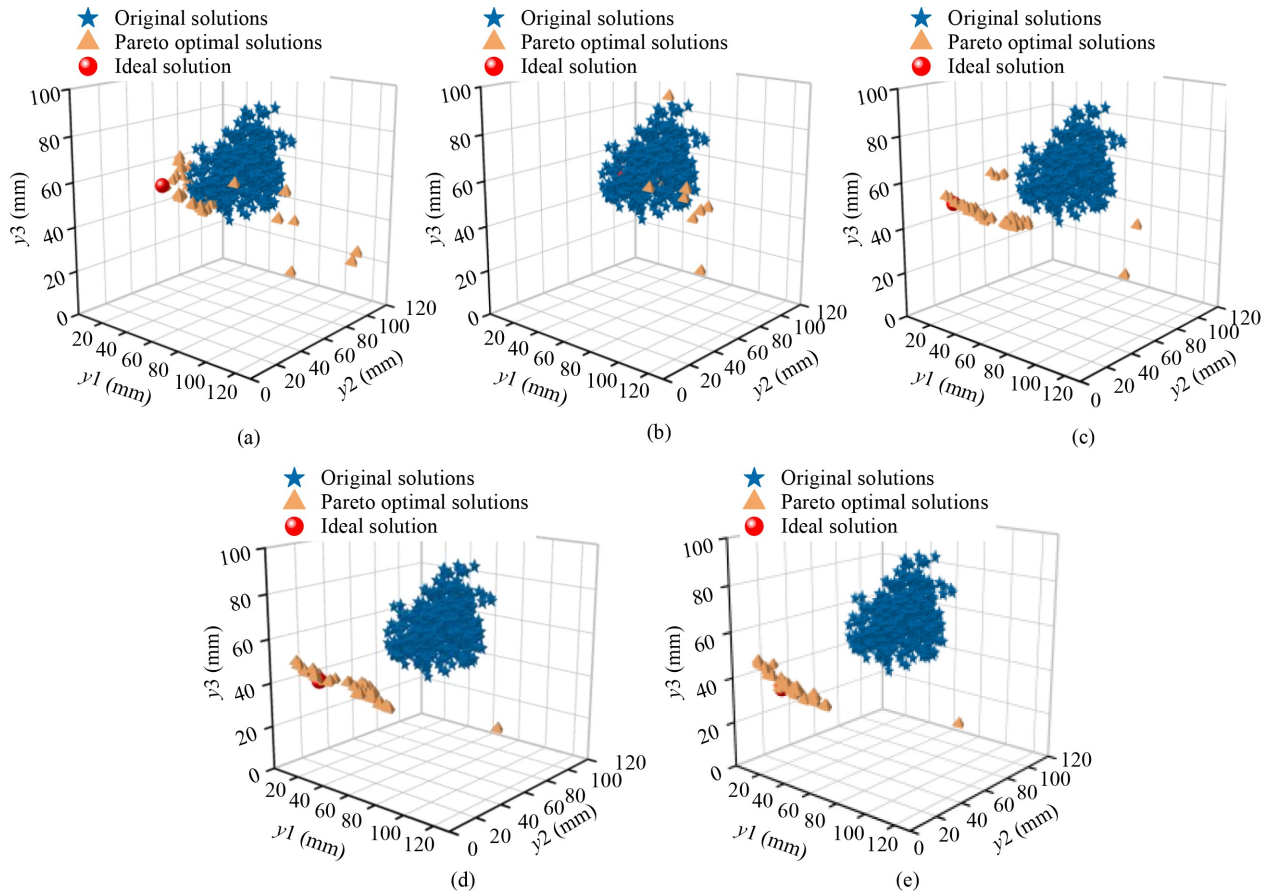
under the different constraints. It can be evident that as the range of the adjustable decision variables increases, the shield tail clearance gradually narrows, and the Pareto solution set is more evenly distributed in the entire non-inferior solution space. As depicted in Table 4, the overall optimized percentage of the shield tail clearance increases up to 58.37% under the 50% constraints than the original data. According to the reference (Han, 2014), the warning value for shield tail clearance is 10 mm, the normal range is 10–20 mm, and the preferable range is 20–30 mm. Hence, the optimized magnitude of shield tail clearance does not exceed the warning value, which indicates the optimized solutions obtained based on the MOO are reasonable. Extensive adjustments to the machine parameters during shield tunneling can potentially lead to significant increases in insecurity and wear and tear. Accordingly, the optimal variation space of adjustable parameters is confined at a 20% self-values up-down range.

## 5 Discussions

Through the case analysis, it is displayed that the online PDNN method is an effective tool for estimating and controlling shield tail clearance. To further highlight the advantages of online PDNN, we compared its performance

with other popular algorithms from two perspectives. From the prediction aspect, three powerful artificial intelligent models, Xgboost, LightGBM, and LSTM, are trained with the same online learning approach as well as training (and testing) data sets. The forecast performance comparison of different models is presented in Fig. 21. From the optimization aspect, this study conducts five comparative experiments. In the first and second groups, the combination of DNN and PDNN with NSGA-II is used to perform the MOO for shield tail clearance. In the third, fourth, and fifth groups, the online PDNN model is combined with the NSGA-III, multi-objective evolutionary algorithm based on decomposition (MOEA/D), and the adaptive geometry estimation based MOEA algorithm 2 (AGE-MOEA2) to control shield tail clearance. The optimization results obtained utilizing different MOO frameworks are shown in Table 5. In addition, this section conducts a comparison between the online PDNN approach and other non-contact measurement techniques. Findings from these results are summarized below.

(1) The online PDNN demonstrates superior performance compared to other state-of-the-art models across all five evaluation metrics, which include *RMSE*, *MAE*, *VAF*, *a20\_index*, and  $R^2$ . As shown in Fig. 21, the online PDNN model consistently achieves the lowest values of *RMSE* and *MAE*, as well as the highest values of *VAF*, *a20\_index*, and  $R^2$ , highlighting its remarkable accuracy



**Fig. 20** Pareto front of MOO on the specific case considering variation space of adjusted decision variables equaling to self-values up-down range for (a) 10%, (b) 20%, (c) 30%, (d) 40%, and (e) 50%.

**Table 4** Data statistics analysis between the original and the optimized values of different constraint conditions based on the MOO framework of online PDNN

Constraints	y1		y2		y3		Optimization percentage			Overall improvement percentage
	Mean/mm	SD/mm	Mean/mm	SD/mm	Mean/mm	SD/mm	y1	y2	y3	
Original data	68.71	10.20	60.57	10.50	64.24	9.50	-	-	-	-
10% control	49.23	15.81	51.63	10.23	59.71	9.14	29.25%	13.68%	6.66%	16.53%
20% control	33.77	13.92	44.09	9.73	54.17	9.26	51.57%	25.75%	15.29%	30.87%
30% control	23.96	11.25	36.27	10.00	48.25	8.68	65.13%	40.12%	24.90%	43.38%
40% control	18.42	8.08	30.03	9.08	41.93	8.48	73.33%	49.29%	34.05%	52.22%
50% control	16.67	7.17	25.15	8.62	36.90	7.75	75.70%	57.64%	41.78%	58.37%

in forecasting the shield tail clearance. Specifically, compared to the prediction performance of LightGBM on  $y_1$ , the online PDNN reduces the  $RMSE$  and  $MSE$  by 1.07 and 0.79, respectively, and improves the  $VAF$ ,  $a_{20}$  index, and  $R^2$  by 0.91, 0.02, and 0.12, respectively. In terms of predicting the upper-left shield tail clearance  $y_2$ , the priorities of the four models are online PDNN > LightGBM > Xgboost > LSTM. Particularly, in comparison to the LSTM model, the online PDNN exhibits a reduction of 1.94 in  $RMSE$ , a reduction of 1.58 in  $MAE$ , an increase of 0.20 in  $VAF$ , an increase of 0.04 in  $a_{20}$

index, and an increase of 0.29 in  $R^2$ . In terms of predicting the upper-right shield tail clearance  $y_3$ , the online PDNN outperforms the other three intelligent algorithms. It reduces the  $RMSE$  and  $MAE$  by at least 0.10 and 0.23, respectively, and improves the  $VAF$  and  $R^2$  by at least 0.03 and 0.03, respectively. Therefore, it can be concluded that the online PDNN model is a more appropriate choice for minimizing prediction errors and enhancing model performance, which is supported by the study by Feng et al. (2019) in 2019. That is, DNN can effectively deal with issues related to small data sets with

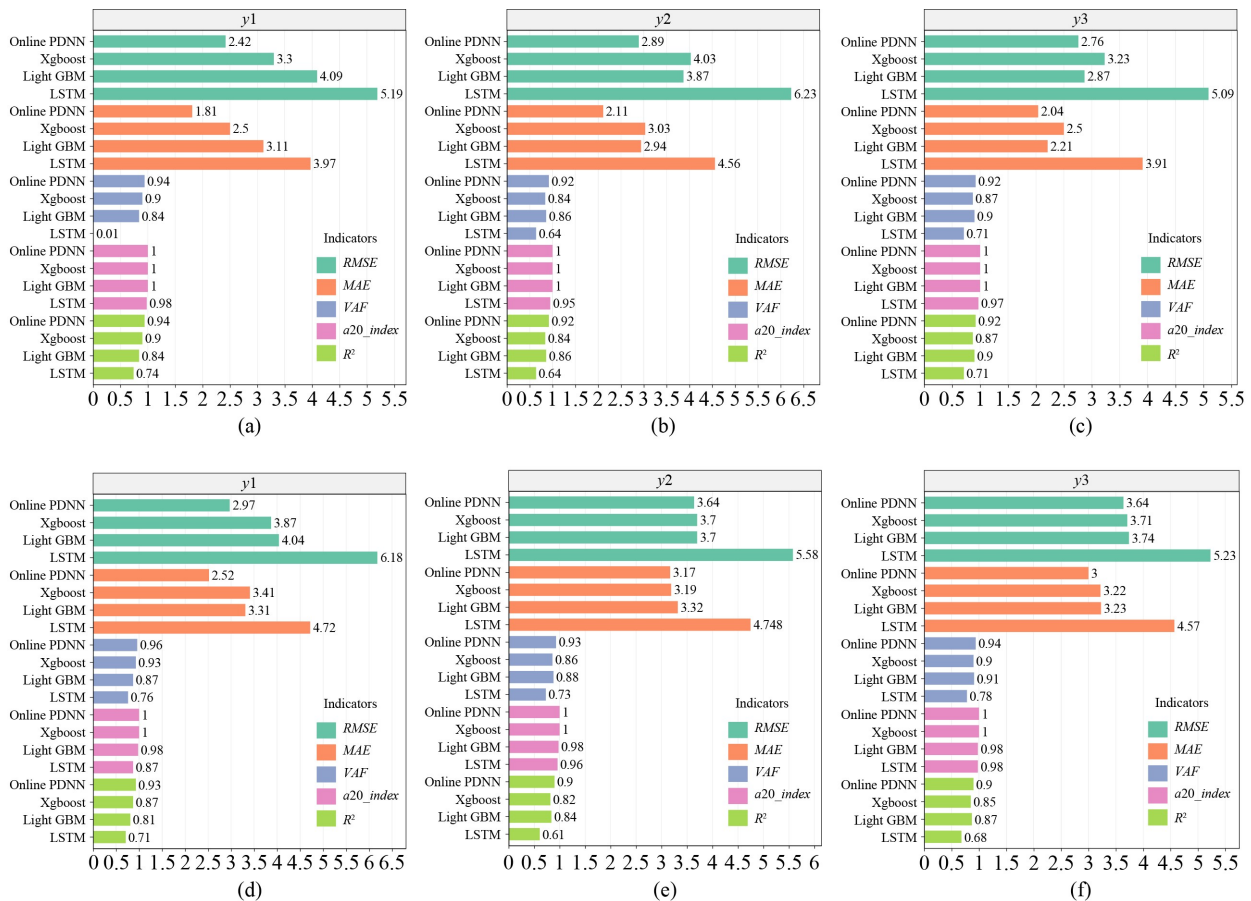


Fig. 21 Model performance on training data (subgraph (a)–(c)) and testing data (subgraph (d)–(f)) from the four intelligent algorithms.

Table 5 Performance comparison of different MOO frameworks

Evaluation indicators		DNN+ NSGA-II	PDNN+ NSGA-II	Online PDNN +NSGA-II	Online PDNN +NSGA-III	Online PDNN +MOEA/D	Online PDNN + AGE-MOEA2
Overall optimization percentage	10% constraint	8.71%	10.17%	16.53%	8.82%	10.61%	13.11%
	20% constraint	17.52%	21.43%	30.87%	26.57%	30.66%	30.75%
	30% constraint	26.41%	32.31%	43.38%	37.76%	39.44%	40.79%
	40% constraint	34.44%	42.88%	52.22%	47.75%	46.43%	47.95%
	50% constraint	42.04%	52.55%	58.37%	46.99%	57.69%	57.80%
HV	10% constraint	4	20	24	2	7	10
	20% constraint	10	2	32	9	6	18
	30% constraint	1	1	7	7	4	7
	40% constraint	3	6	14	7	5	7
	50% constraint	5	1	36	1	5	25
Time (s)	10% constraint	13.62	15.02	12.35	12.59	64.80	18.98
	20% constraint	14.29	14.26	14.70	14.72	64.61	20.37
	30% constraint	13.80	13.74	14.44	14.59	63.03	21.30
	40% constraint	13.86	13.40	14.43	14.76	35.29	17.85
	50% constraint	13.96	10.41	13.92	14.29	32.89	17.40

higher accuracy and superior generalization performance.

(2) Compared with other advanced MOO frameworks, the hybrid model with the integration of online PDNN

and NSGA-II exhibits better performance in performing MOO of shield tail clearance. Based on the average overall improvement percentages under different constraints

shown in Table 5, the priorities of the six MOO frameworks are Online PDNN + NSGA-II > Online PDNN + AGE-MOEA2 > Online PDNN + MOEA/D > Online PDNN + NSGA-III > PDNN + NSGA-II > DNN + NSGA-II. Under the constraints of currently adjustable parameter self-values up-down range for 10%, 20%, 30%, 40%, and 50%, the overall optimization percentage of the proposed hybrid method improves by at least 3.42%, 0.12%, 2.59%, 4.27%, and 0.57%, respectively. Furthermore, the proposed method achieves the Pareto optimal solution sets with the highest HVs. Compared to other MOO frameworks, the proposed method improves the average HV values by at least 9.20, indicating that the Pareto optimal solution sets obtained by this method have better distribution and diversity. Notably, a significant advantage of the online PDNN model over other non-contact measurement methods is that it does not require specialized equipment to measure shield tail clearance, making it more practical for tunnel engineering projects. For example, the VMT GmbH GB's Automatic Tailskin Clearance Measurement System – SluM (VMT, 2021) and ENZAN's Tail Clearance Control System (ENZAN, 2021) use ultrasonic sensors and industrial cameras to collect data, respectively. It makes the measurement process of shield tail clearance more complex and expensive. In summary, the proposed hybrid approach with the integration of online PDNN and NSGA-II can effectively realize the control of shield tail clearance, which may be valuable in facilitating the smooth advancement of the excavation process.

## 6 Conclusions and future works

To accurately predict and control the magnitude of the shield tail clearance, a hybrid approach integrating the online PDNN and NSGA-II is proposed. The contribution of this study is reflected in two aspects. From the prediction of the shield tail clearance aspect, first, the physical equations applicable to measuring shield tail clearance at any position are solved; secondly, the derived equations are integrated into the loss function of the DNN to develop PDNN models. These PDNN models are trained using online learning techniques, which assist in adaptively updating hyperparameters with the incoming stream of observations to realize the accurate estimation of shield tail clearance. From the MOO of the shield tail clearance aspect, this research enriches the area of optimization and control of shield tail clearance. With the proposed hybrid approach, shield tail clearance can be successfully controlled within a reasonable range by adjusting the appropriate TBM operating parameters. In addition, we verify the effectiveness of the method based on the measured data of the construction project of Wuhan Metro Line 12, Guobo Center South Station-Lingwucun

Station.

The conclusions obtained from this research can be summarized below. (1) Under the evaluation metric of  $R^2$ , the online PDNN model demonstrates superior performance, with  $R^2$  values reaching 0.93, 0.90, and 0.90 for the three objectives ( $y_1$ ,  $y_2$ , and  $y_3$ ), respectively. Compared with the other four approaches, including the DNN, PDNN, Xgboost, LightGBM, and LSTM, the prediction result of online PDNN is greater than the previous best result (Xgboost) of 0.06 in  $y_1$ , and the previous best result (LightGBM) of 0.06 and 0.03 in  $y_2$  and  $y_3$ , respectively. (2) Maintaining all other features constant, when the pitch angle value increases, it has a positive influence on the output of the shield tail clearance prediction model. (3) The optimal variability space of the TBM operating parameters used for shield tail clearance optimization is 20% self-values up-down range. By implementing this constraint, the optimized improvement percentage can reach up to 30.87%, with a hypervolume of 32, surpassing the average performance of other MOO frameworks by 5.48% and 23, respectively. (4) In comparison to the existing SluM system and ENZAN's Tail Clearance Control System, the online PDNN model can realize the prediction of shield tail clearance without sacrificing accuracy and efficiency. Additionally, the shield tail clearance measurement method based on the online PDNN eliminates the need for specialized equipment, which makes the measurement process easier and more cost-effective.

Although the results are generally satisfactory, there are still some limitations that need to be addressed in future work. For one thing, the model performance is dependent on the choice of hyperparameters, so more comparative experiments and hyperparameter optimization are needed in future studies to further validate the superiority of the proposed method. For another, since the innovation of the method primarily focuses on the development of the online PDNN model, it somewhat reduces the emphasis on the advanced nature of the MOO algorithm for shield tail clearance. Therefore, future research will focus on exploring the application of adaptive algorithms and multi-agent RL in the shield tail clearance control framework. Additionally, by integrating the proposed hybrid method with digital twin technology, it can be applied to more TBM construction projects with complex conditions, further verifying its applicability, generalizability, and repeatability.

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

## References

Afradi A, Ebrahimabadi A, Hallajian T (2020). Prediction of TBM penetration rate using support vector machine. *Geosaberes: Revista*

- de Estudos Geoeeducacionais, 11: 467–479
- Chai R, Guo Y, Zuo Z, Chen K, Shin H S, Tsourdos A (2024). Cooperative motion planning and control for aerial-ground autonomous systems: Methods and applications. *Progress in Aerospace Sciences*, 146: 101005
- Chai R, Savvaris A, Tsourdos A, Chai S (2017). Solving multi-objective aeroassisted spacecraft trajectory optimization problems using extended NSGA-II. *AIAA SPACE and Astronautics Forum and Exposition*: 5193
- Chai R, Savvaris A, Tsourdos A, Xia Y, Chai S (2020). Solving multi-objective constrained trajectory optimization problem by an extended evolutionary algorithm. *IEEE Transactions on Cybernetics*, 50(4): 1630–1643
- Chai R, Tsourdos A, Savvaris A, Chai S, Xia Y, Philip Chen C (2021). Multiobjective overtaking maneuver planning for autonomous ground vehicles. *IEEE Transactions on Cybernetics*, 51(8): 4035–4049
- Chai R, Tsourdos A, Savvaris A, Xia Y, Chai S (2020). Real-time reentry trajectory planning of hypersonic vehicles: A two-step strategy incorporating fuzzy multiobjective transcription and deep neural network. *IEEE Transactions on Industrial Electronics*, 67(8): 6904–6915
- Chakraborty S (2021). Transfer learning based multi-fidelity physics informed deep neural network. *Journal of Computational Physics*, 426: 109942
- Chen J, Liu S, Li W, Liu F, Zheng D (2021). Design of laser vision measurement system for shield tail clearance of shield tunneling machine. *Machine Tool & Hydraulics*, 49: 77–80
- Chen J, Zhou Z, Liu F, Zheng D (2020). Research on the visual measurement based on shield tail clearance space structure of the shield machine. *Machine Tool & Hydraulics*, 48: 116–121
- Chen K, Zhou X, Bao Z, Skibniewski M J, Fang W (2025). Artificial intelligence in infrastructure construction: A critical review. *Frontiers of Engineering Management*, 12(1): 24–38
- CSTNET (2021). The “Jiangcheng Pioneer” tunnel boring machine surpasses 1,000 rings while tunneling under the Yangtze River. Available at the website of [finance.sina.com.cn](http://finance.sina.com.cn)
- ENZAN (2021). Tail Clearance control system [EB/OL]. Available at the website of [enzan-k.com](http://enzan-k.com)
- Feng S, Zhou H, Dong H (2019). Using deep neural network with small dataset to predict material defects. *Materials & Design*, 162: 300–310
- Feng S X, Chen Z Y, Luo H, Wang S Y, Zhao Y F, Liu L P, Ling D S, Jing L J (2021). Tunnel boring machines (TBM) performance prediction: A case study using big data and deep learning. *Tunnelling and Underground Space Technology*, 110: 103636
- Feng X, Wang P, Liu S, Wei H, Miao Y, Bu S (2022). Mechanism and law analysis on ground settlement caused by shield excavation of small-radius curved tunnel. *Rock Mechanics and Rock Engineering*, 55(6): 3473–3488
- Gao X, Shi M, Song X, Zhang C, Zhang H (2019). Recurrent neural networks for real-time prediction of TBM operating parameters. *Automation in Construction*, 98: 225–235
- Han X, Zhang F, He Y, Zhong H (2021). Research on deformation treatment and control technology of tail shield of underwater large diameter slurry shield. In: *IOP Conference Series: Earth and Environmental Science* IOP Publishing, 783: 012027
- Han Y (2014). Selection of construction parameters for general segment of shield tunnel. *Theoretical Research in Urban Construction* (electronic version): 3446–3447
- He C, Xu C, Xiong D (2021). Development of an automatic measurement device for double laser shield tail gap based on image recognition technology. In: *IOP Conference Series: Earth and Environmental Science* IOP Publishing, 783: 012070
- Lin P, Zhang L, Tiong R (2023). Soil chamber pressure forecasting in tunnel construction using physics-informed deep learning. Available at SSRN 4614636
- Liu J, Bao Z (2021). Construction technology of shield tunnel in extremely-soft fluid plastic stratum. *Tunnel Construction*, 41: 527–532
- Liu J, Shi C, Wang Z, Lei M, Zhao D, Cao C (2021). Damage mechanism modelling of shield tunnel with longitudinal differential deformation based on elastoplastic damage model. *Tunnelling and Underground Space Technology*, 113: 103952
- Liu L, Zhou W, Gutierrez M (2023). Physics-informed ensemble machine learning framework for improved prediction of tunneling-induced short-and long-term ground settlement. *Sustainability*, 15(14): 11074
- Lou P, Li Y, Tang X, Lu S, Xiao H, Zhang Z (2023). Influence of double-line large-slope shield tunneling on settlement of ground surface and mechanical properties of surrounding rock and segment. *Alexandria Engineering Journal*, 63: 645–659
- Mahmoodzadeh A, Mohammadi M, Hashim Ibrahim H, Nariman Abdulhamid S, Farid Hama Ali H, Mohammed Hasan A, Khishe M, Mahmud H (2021). Machine learning forecasting models of disc cutters life of tunnel boring machine. *Automation in Construction*, 128: 103779
- Pan Y, Fu X, Zhang L (2022). Data-driven multi-output prediction for TBM performance during tunnel excavation: An attention-based graph convolutional network approach. *Automation in Construction*, 141: 104386
- Phoon K K, Zhang W (2023). Future of machine learning in geotechnics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 17: 7–22
- Raissi M, Perdikaris P, Karniadakis G E (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378: 686–707
- Rao C, Sun H, Liu Y (2020). Physics-informed deep learning for incompressible laminar flows. *Theoretical & Applied Mechanics Letters*, 10(3): 207–212
- Roy A M, Bose R, Sundararaghavan V, Arróyave R (2023). Deep learning-accelerated computational framework based on physics informed neural network for the solution of linear elasticity. *Neural Networks*, 162: 472–489
- Shaban W M, Elbaz K, Zhou A, Shen S L (2023). Physics-informed deep neural network for modeling the chloride diffusion in concrete. *Engineering Applications of Artificial Intelligence*, 125: 106691
- Shin S, Lee Y, Kim M, Park J, Lee S, Min K (2020). Deep neural network model with Bayesian hyperparameter optimization for prediction of NO<sub>x</sub> at transient conditions in a diesel engine.

- Engineering Applications of Artificial Intelligence, 94: 103761
- Sun L, Zhuang Q (2016). Experimental study on the measurement device of shield tail gap. *Modern Tunnelling Technology*, 53: 56–59
- Tahmassebi A, Motamedi M, Alavi A H, Gandomi A H (2022). An explainable prediction framework for engineering problems: Case studies in reinforced concrete members modeling. *Engineering Computations*, 39(2): 609–626
- Tang L, Kong X, Ling X, Zhao Y, Tang W, Zhang Y (2022). Deviation correction strategy for the earth pressure balance shield based on shield-soil interactions. *Frontiers of Mechanical Engineering*, 17(2): 20
- Verma S, Pant M, Snasel V (2021). A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems. *IEEE Access: Practical Innovations, Open Solutions*, 9: 57757–57791
- VMT (2021). Automatic Tailskin Clearance Measurement System SLuM Ultra. Retrieved from Available at the website of enzank.com
- WMG (2021). Infrastructure dragon! The “Wuhan-made” super-large diameter tunnel boring machine “Jiangcheng Pioneer” rolled off the production line. Available at the website of baijiahao.baidu.com
- Xu Z, Wang W, Lin P, Nie L, Wu J, Li Z (2021). Hard-rock TBM jamming subject to adverse geological conditions: Influencing factor, hazard mode and a case study of Gaoligongshan Tunnel. *Tunnelling and Underground Space Technology*, 108: 103683
- Yang H (2021). Computer vision shield interval automatic detection device development. *Building Technology*, 5: 24–27
- Yavas M S, Gao Z, Mekaoui N, Saito T (2023). A machine learning-based hybrid seismic analysis of a lead rubber bearing isolated building specimen. *Soil Dynamics and Earthquake Engineering*, 174: 108217
- Yin X, Liu Q, Huang X, Pan Y (2022). Perception model of surrounding rock geological conditions based on TBM operational big data and combined unsupervised-supervised learning. *Tunnelling and Underground Space Technology*, 120: 104285
- Yu W (2022). Analysis of shield tail clearance control. *Journal of East China Jiaotong University*, 39: 47–53 (in Chinese)
- Zhang L, Guo J, Fu X, Tiong R L K, Zhang P (2024a). Digital twin enabled real-time advanced control of TBM operation using deep learning methods. *Automation in Construction*, 158: 105240
- Zhang L, Lin P (2021). Multi-objective optimization for limiting tunnel-induced damages considering uncertainties. *Reliability Engineering & System Safety*, 216: 107945
- Zhang L, Wang Y, Fu X, Song X, Lin P (2024b). Geological risk prediction under uncertainty in tunnel excavation using online learning and hidden Markov model. *Frontiers of Engineering Management*, 1–20
- Zhang P, Li H, Ha Q P, Yin Z-Y, Chen R-P (2020a). Reinforcement learning based optimizer for improvement of predicting tunneling-induced ground responses. *Advanced Engineering Informatics*, 45: 101097
- Zhang P, Wu H N, Chen R P, Dai T, Meng F Y, Wang H B (2020b). A critical evaluation of machine learning and deep learning in shield-ground interaction prediction. *Tunnelling and Underground Space Technology*, 106: 103593
- Zhang Z, Pan Q, Yang Z, Yang X (2023). Physics-informed deep learning method for predicting tunnelling-induced ground deformations. *Acta Geotechnica*, 18(9): 4957–4972
- Zhou H, Wang H, Zeng W (2018). Smart construction site in mega construction projects: A case study on island tunneling project of Hong Kong–Zhuhai–Macao Bridge. *Frontiers of Engineering Management*, 5(1): 78–87