Received: 26 February 2025 | Accepted: 26 September 2025

DOI: 10.1111/mice.70096

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

RESEARCH ARTICLE

Real-time prediction of tunnel boring machine thrust based on multi-resolution analysis and online learning

Yongxin Wu[1] | **Hanzhi Yang**[2] | **Houle Zhang**[1] | **Yue Hou**[3] | **Shangchuan Yang**[4]

[1]State Key Laboratory of Precision Blasting, Hohai University, Nanjing, China

[2]Key Laboratory of the Ministry of Education for Geomechanics and Embankment Engineering, Hohai University, Nanjing, China

[3]Department of Civil Engineering, Faculty of Science and Engineering, Swansea University, Swansea, UK

[4]State Key Laboratory of Intelligent Geotechnics and Tunnelling, Southwest Jiaotong University, Chengdu, China

Correspondence
Yue Hou, Department of Civil Engineering, Faculty of Science and Engineering, Swansea University, Swansea SA1 8EN, UK.
Email: yue.hou@swansea.ac.uk

Abstract

This study introduces a novel integrated framework for real-time tunnel boring machine (TBM) thrust prediction, addressing critical limitations in handling non-stationarity, complex spatiotemporal dependencies, and abrupt disturbances. First, a real-time windowed multi-resolution analysis process, which performs decomposition strictly within each segmented sample window, is presented to explicitly disentangle the latent multi-scale dependencies embedded in the thrust data. This ensures strict causality (using only current/historical data), prevents information leakage, and enhances resolution adaptability by capturing local dynamics specific to each data segment, overcoming global averaging effects. Second, a novel synergistic prediction architecture, integrating a hybrid static model with dynamic online residual correction, is proposed. A specifically optimized CNN-LSTM-attention primary model learns complex long-term global patterns. Crucially, an efficient random Fourier features-based online module is dedicated solely to real-time learning of the primary model's residual dynamics, acting as a dynamic corrector rather than an independent predictor. This targeted residual correction significantly enhances robustness against non-stationarity and disturbances. These innovations form an integrated solution and systematically address real-time capability, local adaptability, complex pattern learning, and dynamic error correction. The results indicate that the presented method reduces the mean absolute percentage error from 2.84% to 1.89% and increased R^2 from 0.901 to 0.953. The generalizability of the model was further confirmed through the application of diverse datasets obtained from various chainages along the route. The proposed machine learning–based model can provide guidance for operators in real-time TBM parameter adjustment during construction

1 | INTRODUCTION

Tunnel boring machines (TBMs) are sophisticated machines employed pervasively in excavation of tunnels for both transit infrastructure and water conveyance systems (Melis et al., 2002). TBMs offer advantages over traditional construction methods due to their high efficiency, improved safety, minimal surface disturbance, and adaptability to complicated geological conditions (Liu & Ding, 2020). The operating parameters of the TBM need dynamic adjustment by the operator to match the geological profile ahead of the cutterhead during the construction (Guo et al., 2022; Jin et al., 2022). However, real-time parameter adjustment remains challenging in practice due to the difficulty in accurately predicting the geological conditions ahead of the TBM (Chen et al., 2021; Gao et al., 2019; Y. P. Zhang et al., 2023). Thrust serves as a critical indicator for the TBM performance, directly correlating with the geological state encountered ahead of the cutterhead. Analyzing historical thrust time series allows for predicting future thrust values, enabling optimized TBM operation parameters for more efficient tunnel construction.

Traditional methods for predicting TBM parameters primarily involve a rock–soil mechanics model and numerical simulations. For instance, theoretical models have been created to estimate TBM loads under different ground conditions, including homogenous soil (Q. Zhang et al., 2014) and mixed rock–soil interfaces (Qi et al., 2022). These methods typically simplify the complex cutterhead load into several components based on the dynamic interplay between the TBM and its geological environment. However, the simplifying assumptions inherent in these theoretical models often limit their accuracy, providing only a rough range estimate of cutterhead loads, making them incompetent for accurate TBM parameter prediction. Numerical simulation techniques, such as the discrete element method (DEM) (Chen & Zhao, 2002; Wu et al., 2013) and finite element method (FEM) (Han et al., 2017; Huang et al., 2024), are also applied for predictions of TBM parameters. Unlike theoretical models, numerical simulations can effectively model complex geological conditions, heterogeneous soils, and fluid–solid coupling effects. However, capturing subtle changes in TBM parameters remains challenging (Finger & Dixon, 1989; Lislerud, 1988). Furthermore, both theoretical and numerical methods rely on detailed geological information, which is typically obtained through sparsely spaced sampling points. This lack of geological data hinders their application in real-time parameter forecasting.

Machine learning methods have shown promise in addressing the limitations of traditional TBM prediction methods (Kubota et al., 2024; Tondini et al., 2015; S. Y. Wang et al., 2024). These data-driven approaches leverage machine learning to discover complex patterns from data, while their development is fundamentally guided by physical knowledge in feature selection, data conditioning, and model design inspiration (C. Shi et al., 2021; Zhao et al., 2021). For instance, R. H. Wang et al. (2021) established a dynamic regulation model based on wavelet transform and bidirectional long short-term memory method to predict advance rate and torque. Other deep learning techniques, including causal explainable gated recurrent unit (GRU) (K. Y. Wang et al., 2023), deep belief networks (Feng et al., 2021), and graph convolutional networks (Pan et al., 2022), have demonstrated effectiveness in TBM performance evaluation and prediction. Researchers have also explored optimization algorithms for machine learning model hyperparameters.

Signal decomposition techniques such as empirical wavelet transform (Altan et al., 2021), empirical mode decomposition (Ali & Parasad, 2019), and variational mode decomposition (Abdoos, 2016) are frequently employed in time series prediction to enhance models' representational capacity for characterizing and predicting complex parameter fluctuations in geologically heterogeneous environments. Within the TBM parameter prediction domain, Qin et al. (2022, 2024) and G. Shi et al. (2021) achieved high-precision cutterhead torque forecasts through multimodal decomposition of input signals, thereby validating the efficacy of decomposition methods for processing the TBM operational data. However, the conventional "global decomposition → sample segmentation → independent multi-model prediction → result summation" paradigm requires a complete time series for global decomposition, actually utilizing future data. This violates real-time forecasting's causality principle, induces information leakage, and lacks dynamic execution capability with only historical data, thus failing to meet real-time temporal constraints. To fill this gap, a real-time windowed multi-resolution analysis (MRA) is proposed. This method decomposes thrust data strictly within sequential windows to capture local dynamics (like sudden changes) adaptively, ensuring causality, and preventing information leakage.

Additionally, while existing models support multistep prediction, their reliance on historical training datasets hinders their ability to dynamically adapt to real-time changes in geological conditions encountered during tunneling (such as undetected minor-scale geological anomalies) and the systematic drift induced by progressive cutter wear. Batch learning models, relying on static historical data, fail to dynamically adapt to real-time variations in geological conditions (e.g., undetected sudden rock layer changes) and progressive cutter wear during shield tunneling. This results in delayed or inaccurate parameter

predictions, potentially triggering engineering failures such as equipment overload damage and unplanned downtime, ultimately leading to project delays and resource wastage. Consequently, online learning methods need to be introduced to adjust the model dynamically. However, prevailing advanced techniques face critical limitations in tunneling operations: continual learning (e.g., elastic weight consolidation (EWC)) (Gepperth & Wiech, 2019), which mitigates catastrophic forgetting via regularization or knowledge distillation during full-model updates, introduces prohibitive computational overhead incompatible with real-time control. Dynamically expanding architectures (e.g., progressive nets) (Mehul et al., 2023), which add new modules for new tasks, escalate inference latency and memory consumption beyond embedded hardware limits. Online meta-learning (e.g., model-agnostic meta-learning (MAML) variants; Park et al., 2021), requiring rapid adaptation from few examples via gradient-based updates, necessitates extensive adaptation samples, delaying response to emergent anomalies. Consequently, a synergistic prediction architecture that combines a hybrid convolutional neural network-long short-term memory (CNN-LSTM)-attention (CLA) model learning global patterns with a dedicated online residual corrector is presented. This method utilizes a kernel approximation-based module dynamically to adjust the primary model's output in real-time by learning its residual errors (unmodeled effects, noise), significantly enhancing robustness.

In this study, a deep learning method incorporating real-time decomposition and an online learning section is proposed for predicting the TBM thrust. In Section 2, the raw data are cleansed, normalized, and processed into a format compatible with deep learning models. Then, each input time window is decomposed into several subsequences. In Section 3, a CLA model is established to generate the preliminary prediction, and an online learning section is utilized to forecast the residuals between target values and predicted values. Finally, the performance of the proposed model is evaluated in Section 4.

2 | MATERIALS AND CONSTRUCTION OF DATASET

2.1 | Project review

The dataset utilized in this research originates from the Dali segment of the water conveyance line in the Dianzhong Water Diversion Project. This tunnel section exhibits significant topographic relief, with a burial depth ranging from 87 to 1335 m (averaging 400–950 m). The surrounding rock mass primarily consists of medium-to-thickly bedded limestone and dolomitic limestone. The rock mass integrity is rated as fair to poor (Class III–IV), resulting in overall moderate tunnel-forming conditions.

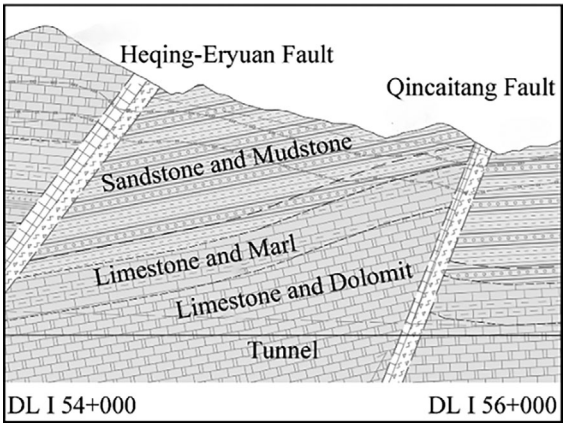

FIGURE 1 Geological profile of the project.

TABLE 1 Details of each dataset.

Dataset number	Recording date	Chainage	Size
Dataset 1	July 5, 2021	55959-55953	9,300
Dataset 2	May 15, 2021	56558-56547	10,000
Dataset 3	June 21, 2021	56101-56089	10,000
Dataset 4	August 21, 2021	55623-55614	10,000
Dataset 5	October 27, 2021	55117-55171	10,000

The key geological risk is concentrated within the Qincaitang fault zone. Within this zone, the rock mass is highly fractured (Class V), exhibiting poor tunnel-forming conditions and prominent issues with surrounding rock stability. The structural planes predominantly display planar surfaces with straight traces, often infilled with clayey material and rock fragments. During construction, geohazards such as water inrush and mud outbursts are frequently encountered/accompanied. The geological profile is shown in Figure 1.

2.2 | Data resource

Operational and state parameters of key TBM subsystems were recorded every second during excavation using onboard sensors. These data, comprising approximately 400 items, were collected by a programmable logic controller and stored in a database. A total of approximately 86,000 data points were gathered daily, with a sampling frequency of 1 Hz. Five segments of data, each recorded on a different day, were selected to evaluate the proposed method comprehensively. These segments were processed into five distinct datasets following the

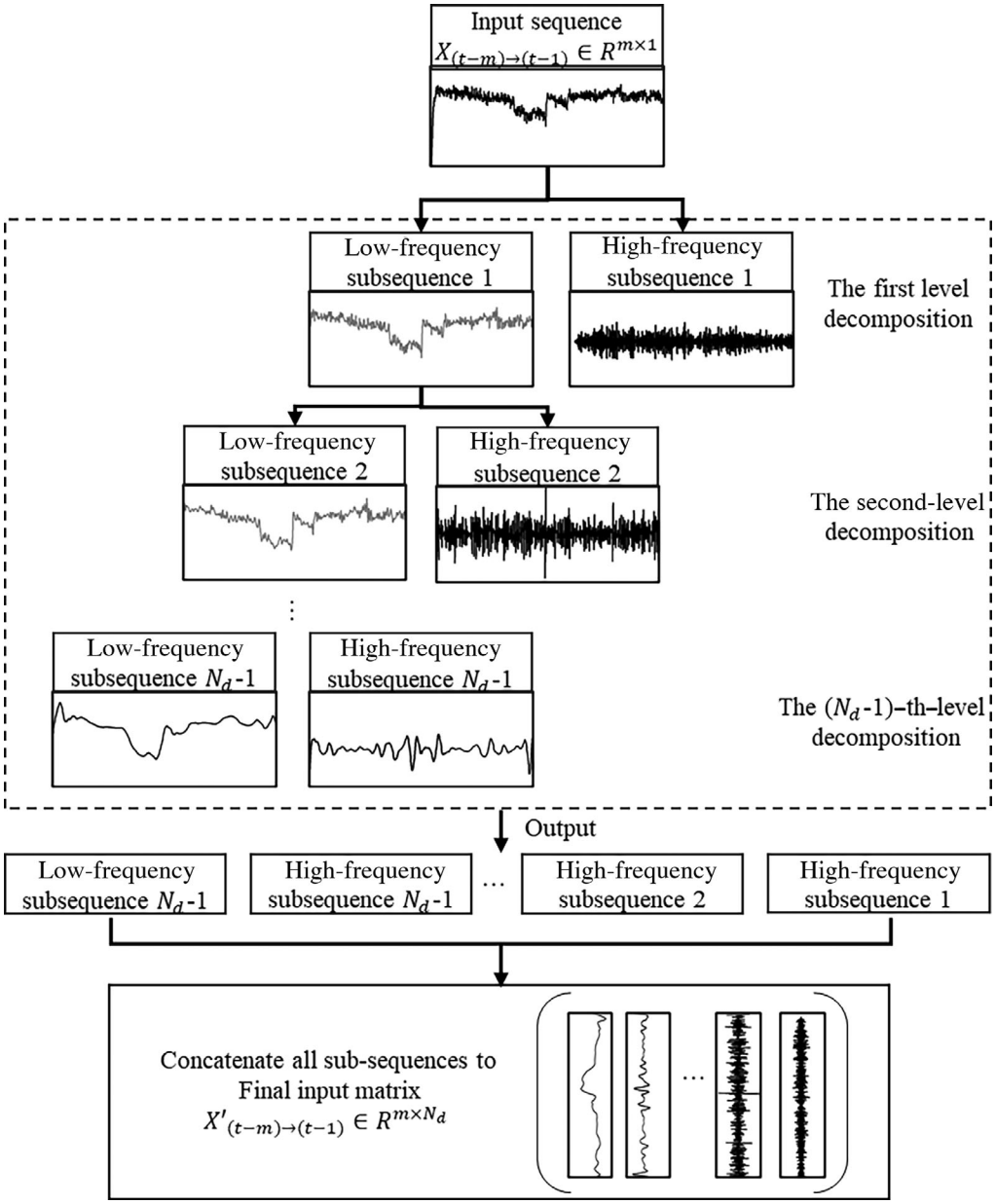

FIGURE 2 Details of multi-resolution analysis.

procedure described below. Dataset 1 was used to analyze the effects of the decomposition method and the online learning section. Datasets 2 through 5 were employed to evaluate the generalizability of the model. The details of these datasets are summarized in Table 1.

2.3 | Data preprocessing

The operational states of the TBM can be categorized as running or stopped. The boring cycle, defined as the interval between machine startup and shutdown, typically lasts hundreds to thousands of seconds. Invalid data recorded between boring cycles should be removed. The dataset is cleansed by the following method:

$$F = \prod_{i=1}^{K} f(x_i) = \begin{cases} 1, & \text{Boring state} \\ 0, & \text{Shutdownstate} \end{cases}, \tag{1}$$

$$f(x_i) = \begin{cases} 1, x_i > 0 \\ 0, x_i < 0 \end{cases}, \tag{2}$$

where x_i represents the ith operational parameter obtained from TBM's data acquisition system. A binary method was employed to assess the boring states by defining a state function $F(x)$, which contains several sub-items $f(x_i)$. Each $f(x_i)$ is a determining function for a specific data point. Using this method, a full day's recording is segmented into distinct boring cycles.

Data standardization is essential to mitigate the negative impact of varying magnitudes of TBM operational parameters on predictive performance. In this study, the min–max scaling is adopted to rescale all attributes to the range [0, 1], which is formulated as follows:

$$x_j = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \tag{3}$$

where x_j denotes the scaled value; x_i represents the observed value from the original dataset; and $x_{\max}$ $x_{\min}$ represent the maximum and minimum values, respectively.

After standardization of the raw data, the dataset is divided into a training set and a testing set. In this research, 70% of the dataset is allocated for training, with the remaining 30% reserved for testing. Before training, the data are transformed into input–output pairs to be compatible with the TensorFlow Keras framework. The training sample for the deep learning predictor can be expressed as $(\{x_{t-m}, x_{t-m+1}, \dots, x_{t-1}\}, \{x_t, x_{t+1}, \dots, x_{t+n-1}\})$, which can be simplified as $(X_{(t-m)\to(t-1)}, X_{t\to(t+n-1)})$. m and n represent the length of the input series and the length of the output series, respectively.

2.4 | Multi-resolution analysis

MRA is a computational technique designed to analyze signals or data across multiple scales or resolutions. In this study, for each input–output pair $(X_{(t-m)\to(t-1)}, X_{t\to(t+n-1)})$, the MRA is applied to the input sequence $X_{(t-m)\to(t-1)}$. The real-time and causal nature of this approach relies fundamentally on this forward-looking windowing strategy. The general procedure is illustrated in Figure 2 and outlined below:

1. Decompose the input sequence $X_{(t-m)\to(t-1)}$ into a low-frequency component and a high-frequency component through a group of orthotropic wavelet basis functions.
2. The decomposition is iterated on successive low-frequency components until a predefined level $N_d - 1$ is reached.
3. All high-frequency sub-signals, along with the ultimate low-frequency sub-signal, are combined to create the input matrix $X'_{(t-m)\to(t-1)}$ for the neural network.

The process of decomposition and reconstruction of the signal can be mathematically illustrated as

$$x(t) = A_{N_d-1}(t) + \sum_{j}^{N_d-1} D_j(t), \tag{4}$$

where $A_{N_d-1}(t)$ is the approximation (low-frequency subsequence) and $D_j(t)$ is the detail (high-frequency subsequence). $N_d - 1$ represents the maximum level of decomposition. The approximation and detail can be expressed as

$$A_{N_d-1}(t) = \sum_{k} a_{N_d-1,k} \cdot \phi_{N_d-1,k}(t), \tag{5}$$

$$D_{N_d-1}(t) = \sum_{k} d_{j,k} \cdot \phi_{j,k}(t), \tag{6}$$

where $\phi_{j,k}(t)$ is the scale function; $\psi_{j,k,t}$ is the wavelet basis function; a_k and $d_{j,k}$ are the scale coefficient and wavelet coefficient; j and k represent the decomposition level and the position index within that level, respectively. In this study, the MRA is applied to each time window of input data, which provides both rough and detailed information of thrust signal. This process yields a causally valid multi-scale representation of the thrust signal, thus producing a feature set of dimension N. The final input matrix for the neural network consequently has the shape of $m \times N_d$.

3 | ESTABLISHMENT OF THE MODEL

3.1 | Architecture of the proposed model

The intricate interplay between the TBM and surrounding geological conditions poses substantial challenges in predicting TBM's thrust. To uncover the underlying patterns in the tunnel boring process and achieve accurate predictions of thrust, an efficient machine learning–based model is proposed. The architecture of the proposed model, illustrated in Figure 3, primarily comprises three essential components: signal decomposition, deep learning prediction, and online learning prediction. Initially, the input signal is decomposed into subsequences of varying frequencies. Subsequently, a hybrid deep learning model, designated as the CLA model, is developed. This model integrates a convolution layer, a long short-term memory (LSTM) layer, and an attention mechanism to generate multistep predictions. Finally, an online section is incorporated to predict the residuals between the deep learning results and actual values. The final predictions are obtained by summing the preliminary predictions with these residuals. The CLA model serves as the baseline model. When integrated with the signal decomposition method, it is referred to as the SD-CLA model. By integrating the online learning module, this framework is termed the SD-CLA-OL model.

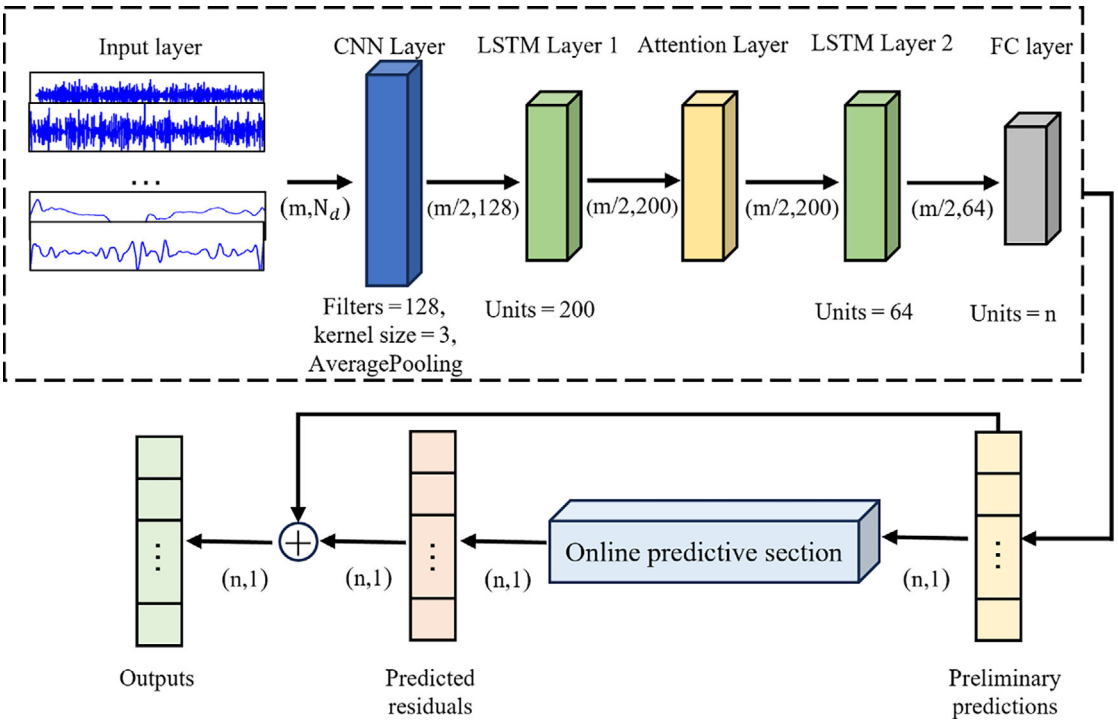

FIGURE 3 The architecture of the proposed model. FC, fully connected; LSTM, long short-term memory.

3.2 | Details of the CLA model

A hybrid architecture is designed to hierarchically capture and prioritize relevant patterns in the data. The process begins with a one-dimensional CNN layer, which operates along the temporal dimension to extract salient local features from the input time series (R. Y. Zhang, Liu, et al., 2020). Using a kernel size of 3 and 128 filters, followed by rectified linear unit (ReLU) activation and average pooling, this layer identifies short-term, high-frequency patterns within small sliding windows of the signal. By condensing the input into a more informative feature representation, the CNN not only reduces computational complexity but also provides enhanced input to subsequent layers (Cui et al., 2021).

These extracted features are then passed to the first LSTM layer, comprised of 200 units. This layer processes the sequence step-by-step, capturing medium- and long-range temporal dependencies through its gated mechanism—regulating the flow of information via input, forget, and output gates (P. Zhang, Yin, et al., 2020). Crucially, this LSTM returns the hidden state at each time step, preserving the temporal structure of the features and enabling finer-grained analysis in subsequent layers.

An attention mechanism is then applied over these step-wise LSTM outputs. It computes adaptive weights to dynamically highlight the most relevant time steps for the prediction task, effectively addressing the limitation of treating all historical points equally. This results in a weighted context vector that emphasizes critical periods within the sequence (e.g., moments of abrupt change or high variability), thereby minimizing the loss of salient information over long sequences (Kang et al., 2023). The calculation process is illustrated in Figure 4.

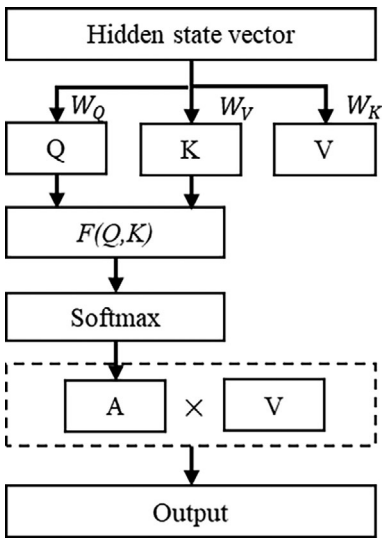

FIGURE 4 The calculation process of the attention mechanism.

Finally, the context vector produced by the attention layer is fed into a second LSTM layer. This layer synthesizes the attended information and further models complex temporal interactions, serving as an aggregator that refines

the representation before prediction. The output from this LSTM is passed through a fully connected dense layer to produce the final prediction.

Throughout this architecture, information is transformed sequentially from local feature maps (CNN) to temporal encoding (first LSTM), to attention-weighted summarization, and ultimately to context-aware temporal integration (second LSTM). This multistage design ensures that both fine-grained variations and overarching trends are captured effectively, significantly enhancing predictive performance in complex time-series forecasting tasks such as the TBM thrust prediction.

3.3 | Online learning section

This research introduces a hybrid framework integrating the predictive strengths of batch learning with the adaptive nature of online learning, aiming to enhance the efficacy of current methods. To enable online learning, this study employs a kernel approximation method based on random fourier features (RFF) to perform online kernel regression for predicting the residuals between the deep learning model's outputs and the true values. Specifically, while a stochastic gradient descent (SGD) regressor serves as the underlying linear model, it is applied within a high-dimensional feature space transformed by RFF. This transformation approximates the radial basis function (RBF) kernel, effectively enabling nonlinear modeling. It is crucial to highlight that the linear prediction model acts upon the high-dimensional representation generated by RFF mapping, not the original input directly. This approach thus enables the model to capture underlying nonlinear patterns while preserving the computational advantages of a linear model. The linear relationship of inputs and predictions can be expressed as

$$\hat{y} = \omega^T x + b, \tag{7}$$

where $\hat{y}$ is the predicted value, ω is the weight parameter, and b is the bias term.

The renewing rule of SGD is as follows:

$$\begin{aligned} \omega &:= \omega - \eta \cdot \nabla L(\omega, x_i, y_i), \\ b &:= b - \eta \cdot \nabla L(\omega, x_i, y_i), \end{aligned} \tag{8}$$

where the loss function is represented by L, which is squared error in this study; $\nabla L(\omega, x_i, y_i)$ denotes the gradient of the loss function concerning the parameters, computed at the current given data point; the learning rate is denoted by η. While the RFF method is leveraged to extend the linear model to the nonlinear condition, which provides an efficient approximation for the RBF kernel. It uses random features to estimate the kernel function, which effectively maps raw inputs to an elevated dimensional representation, simplifying computations and enabling nonlinear modeling (Engel & Mannor, 2004; Kivinen et al., 2001; Qin et al., 2017). The details of this method are presented in Rahimi and Recht (2007). This approach is grounded in Bochner's theorem, which asserts that any continuous, shift-invariant kernel $k(x, x') = k(x - x')$ on R^d can be expressed as the Fourier transform of a nonnegative measure. Leveraging this property, the RBF kernel can be approximated using a finite sum of random features. First, sample random vectors ω_i from the Gaussian distribution $\rho(\omega) = N(0, \frac{1}{\sigma^2})$. Then, for a data point x, the random feature mapping is constructed as

$$z(x) = \sqrt{\frac{2}{D}} \begin{bmatrix} \cos\left(\omega_1^T x + b_1\right) \\ \cos\left(\omega_2^T x + b_2\right) \\ \cdots \\ \cos\left(\omega_D^T x + b_D\right) \end{bmatrix}, \tag{9}$$

where b_i are random offsets sampled uniformly from $[0, 2\pi]$; D is the number of random features, chosen to control the approximation accuracy, which is 100 in this passage. Finally, the original kernel function is approximated by the inner product within the transformed feature space through feature mapping.

$$k(x, x') \approx z(x)^T z(x'). \tag{10}$$

It approximates the RBF kernel's nonlinear capabilities while maintaining computational efficiency. This makes it particularly suitable for large-scale machine learning tasks. The integration of a linear predictor with the RFF technique enables real-time updates while enhancing predictive performance and model robustness.

The details of the online learning module at time t are illustrated in Figure 5. Initially, the residuals R between predicted values $\widehat{X}$ and actual values X from time $t-n$ to $t-1$ are calculated. Then, R is mapped to a higher dimension using a kernel approximation method and input to a linear predictor. The predicted residuals $\widehat{R}$ from time t to $t+n-1$ are given by the predictor. After n seconds from time t, the $R_{t\to(t+n-1)}$ becomes available as known data. Thereafter, $R_{(t-n)\to(t-1)}$ and $R_{t\to(t+n-1)}$ are utilized as an input–output pair to update the weights of the online learning section.

4 | RESULT ANALYSIS

4.1 | Evaluating predictive performance

To comprehensively assess the efficacy of the proposed model, the mean absolute percentage error (MAPE) is

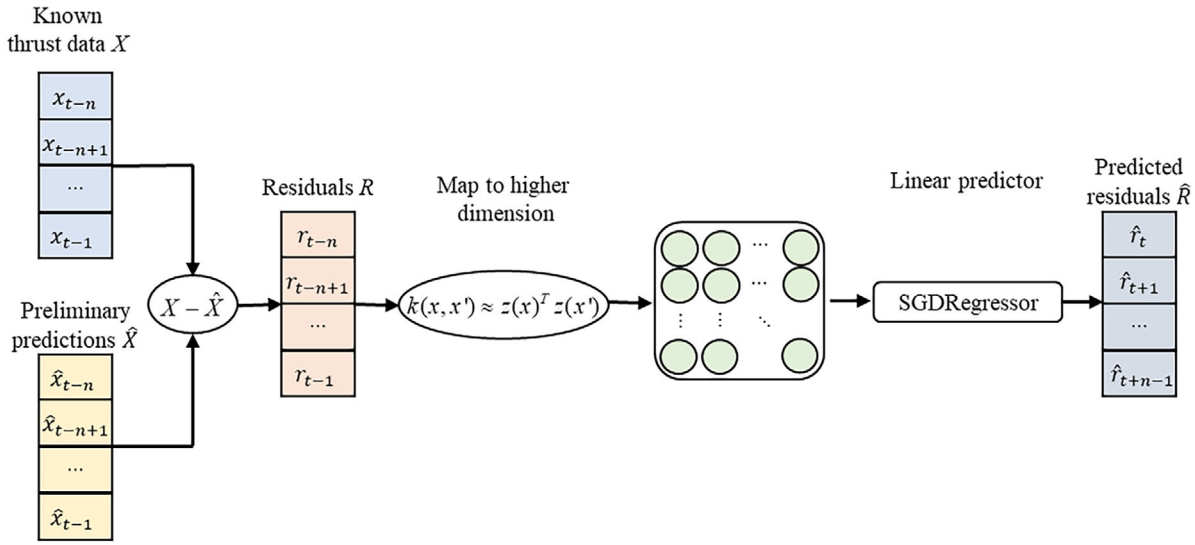

FIGURE 5 The details of the online learning section at time *t*.

introduced in this study to assess the accuracy and reliability of the model.

$$MAPE = \frac{100\%}{l}\sum_{i=1}^{l}\left|\frac{\hat{y}_i - y_i}{y_i}\right|, \tag{11}$$

where l is the number of data points; $\hat{y}_i$ is a predicted value from the proposed model, y_i is the observed data obtained from the data acquisition system. MAPE reflects the accuracy of the predictive model, with lower values indicating more accurate predictions. R^2 is used in this study to further assess model performance. The R^2 value measures the fraction of variability in the dependent variable that can be explained by the independent variables. R^2 value of 1 signifies perfect agreement between predicted and observed data. The formula for calculating R^2 value is

$$R^2 = 1 - \frac{\sum_{i=1}^{l}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{l}(\bar{y}_i - y_i)^2}, \tag{12}$$

where $\bar{y}_i$ denotes the mean of the observed data.

4.2 | Effect of signal decomposition

The efficacy of the introduced decomposition approach is evaluated on Dataset 1. As the majority of existing multi-step prediction models adopt a five-step prediction horizon (Qin et al., 2022; G. Shi et al., 2024), this study follows the same approach and focuses on analyzing the predicted results at the fifth step. Figure 6 compares the predicted thrust at the fifth step between the CLA model and the SD-CLA model with different decomposition levels. It can be observed that as the number of decomposition levels increases, the predicted values align more closely with the actual values, and the error at jump points is significantly reduced.

Figure 7 illustrates the predictive accuracy at the fifth step across varying decomposition levels of input signals. The MAPE value decreases from 2.72% to 2.44% when the decomposition level increases from 1 to 7. Concurrently, the R^2 value increases steadily from 0.911 at level 1 to a peak value of 0.930 at level 7, demonstrating an increasing correlation degree between the predicted and actual values. The improvement rate gradually diminishes as the decomposition level increases. This progressive improvement indicates that an increase in decomposition level enhances the model's capability to discern and exploit meaningful signal patterns. Beyond decomposition level 7, the R^2 and MAPE reach a plateau and begin to fluctuate within a narrow range, which indicates that further decomposition does not significantly enhance prediction performance. This fluctuation suggests that excessive decomposition may introduce redundant information or distorted signals, which will not make contributions to the overall predictive performance. Based on this analysis, decomposition level 7 is selected as the optimal level, balancing improved prediction accuracy and computational efficiency. At this stage, the model attains minimal prediction error alongside maximal correlation, avoiding unnecessary complexity while ensuring reliable performance.

The normalized error distributions for the fifth-step predictive results of the CLA model and the SD-CLA model with seven levels of decomposition are presented in Figure 8. The error distribution for the SD-CLA model exhibits a sharper peak and a noticeably narrower spread compared to those of the CLA model. This indicates that the decomposition model achieves reduced prediction

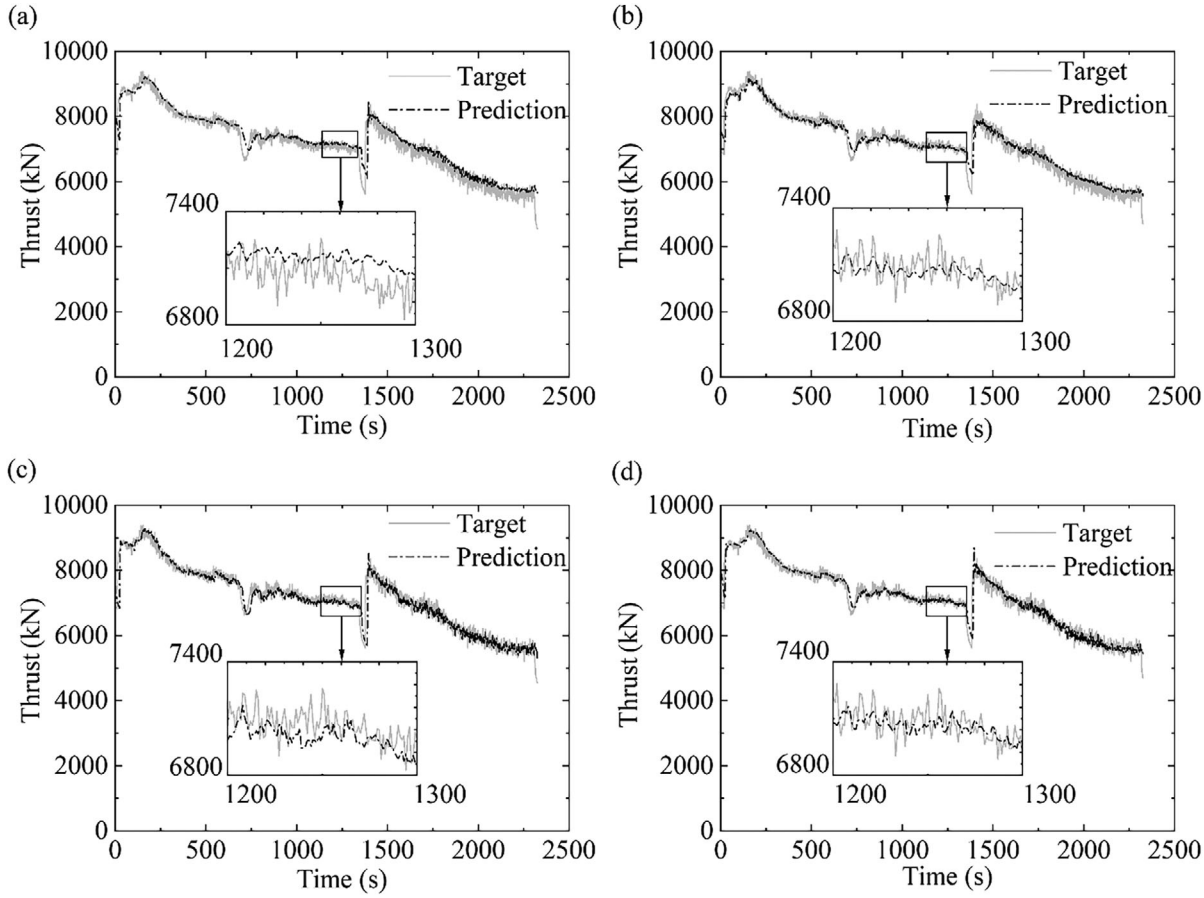

FIGURE 6 Predictive performance of the CLA model and the SD-CLA model with different decomposition levels, (a) CLA model, (b) SD-CLA model with decomposition level 1, (c) SD-CLA model with decomposition level 4, (d) SD-CLA model with decomposition level 7.

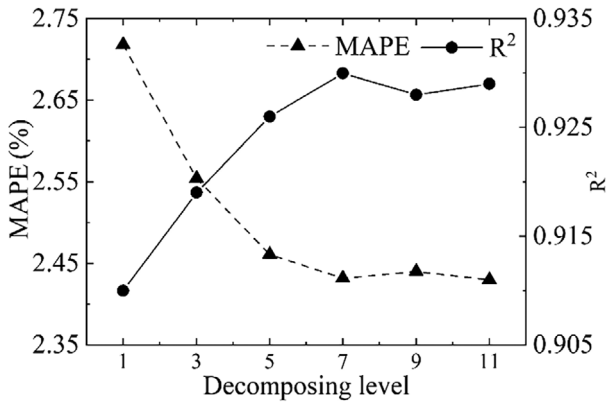

FIGURE 7 The relationship between decomposition levels of input signals and the prediction performance. MAPE, mean absolute percentage error.

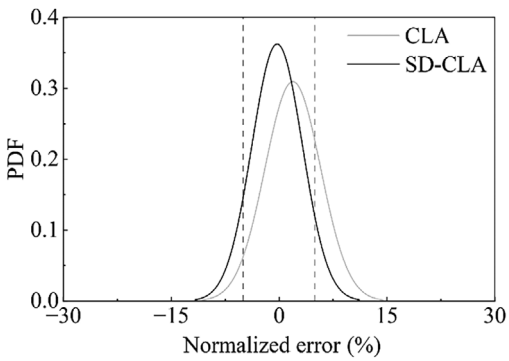

FIGURE 8 Error distributions of the prediction for the CLA model and SD-CLA model.

error and enhanced accuracy. The confidence intervals (CI), defined with a 5% margin of error, for the CLA model and SD-CLA model are 91.98% and 82.70%, respectively. This suggests that the decomposition model achieves better reliability and consistent performance. Furthermore, the mean values of the error distribution curves for the SD-CLA model and CLA model are 0.27% and 1.89%, respectively. This disparity highlights a notable reduction in systematic bias when signal decomposition is employed. In addition, the decomposition model has shorter tails in the error distribution curve, signifying fewer extreme errors. This feature underscores its ability to produce more reliable and consistent predictions, thereby reducing both systematic and random errors.

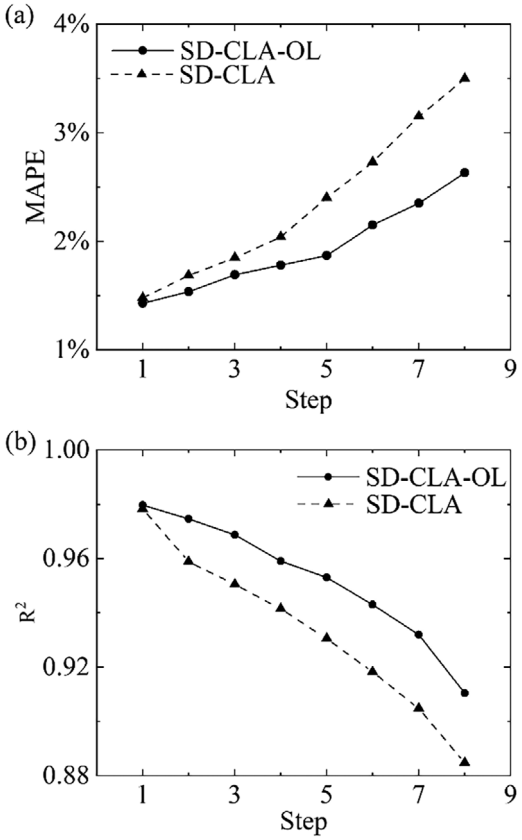

FIGURE 9 MAPE and R^2 for multistep predictions of the SD-CLA model and SD-CLA-OL model.

4.3 | Effect of the online learning section

The effectiveness of the online learning section is evaluated based on Dataset 1. Figure 9 illustrates the multistep predictive performance of the SD-CLA model and the SD-CLA-OL model. In Figure 9a, it is evident that when predicting only one step ahead, the MAPE values of both models are nearly identical. As the output horizon increases, the MAPE values of both models increase. However, the SD-CLA-OL model exhibits a slower rate of MAPE increase compared to the SD-CLA model. At the first step, the difference in MAPE values between the two models is merely 0.05%, but by the eighth step the difference grows to 0.87%. From the first to the eighth step, the MAPE increases by 1.2% and 2.07% for the two models, respectively. This demonstrates that the incorporation of the online learning section effectively mitigates the cumulative error, thereby improving the model's multistep predictive accuracy. Figure 9b reveals a similar trend for R^2, where the model with the online learning section consistently maintains higher R^2 values across all prediction steps. At the first step, both models exhibit high R^2 values near 0.98, indicating strong predictive performance. However, as the prediction horizon extends, the R^2 for both models declines, with the SD-CLA model experiencing a sharper decrease. At the eighth step, the R^2 for the SD-CLA-OL model remains 0.911, whereas the R^2 of the SD-CLA model drops below 0.885. From the first to the eighth step, the R^2 growth for the two models is 0.069 and 0.094, respectively. This underscores that the SD-CLA-OL module enhances the model's ability to explain variance over extended prediction horizons, thereby improving its robustness and reliability. After the fifth step, both the MAPE and R^2 curves of the SD-CLA-OL model exhibit a marked increase in steepness, suggesting a rapid deterioration in the model's predictive accuracy beyond this point. To ensure the model's precision and reliability, an output horizon of five steps is thus selected.

TABLE 2 Relative predictive error of each time step.

Relative error range	Proportion of each time step (%)				
	Step 1	Step 2	Step 3	Step 4	Step 5
<1%	48.11	44.08	41.77	41.26	40.10
<5%	97.73	97.17	96.19	95.97	95.71

The relative predictive error of each time step is displayed in Table 2. The proportion of predictions with relative error less than 1% decreases from 48.11% at Step 1 to 40.10% at Step 5, indicating an 8% decline. Similarly, the proportion of predictions with relative error less than 5% decreases from 97.73% at Step 1 to 95.71% at Step 5, reflecting 2% decline. This indicates that the model exhibits higher accuracy in short-term predictions (Step 1), while the stability of long-term predictions (Step 5) slightly decreases as the time steps progress. The proportion of predictions with relative error less than 5% consistently exceeds 95%, indicating robust overall predictive reliability within an acceptable error margin. The proportion of predictions with relative error less than 1% shows a more pronounced decline, reflecting that the challenge for the model to achieve high-precision predictions (strict error control) increases as the steps progress.

The fifth-step predictive results of the SD-CLA-OL model and SD-CLA model are compared in Figure 10. Overall, both models' predictions closely follow the target values. However, they exhibit notable discrepancies in capturing fine-grained details. The SD-CLA model demonstrates limitations in accurately capturing local thrust variations, particularly in regions with significant fluctuations. The predicted values deviate more noticeably from the target values, showing a degree of lag and instability. Furthermore, the model without the online section struggles to fit high-frequency variations, leading to larger local prediction errors. The introduction of the online learning section allows the model to dynamically adjust its outputs based on real-time feedback. In terms of performance metrics, the final model delivers the more accurate

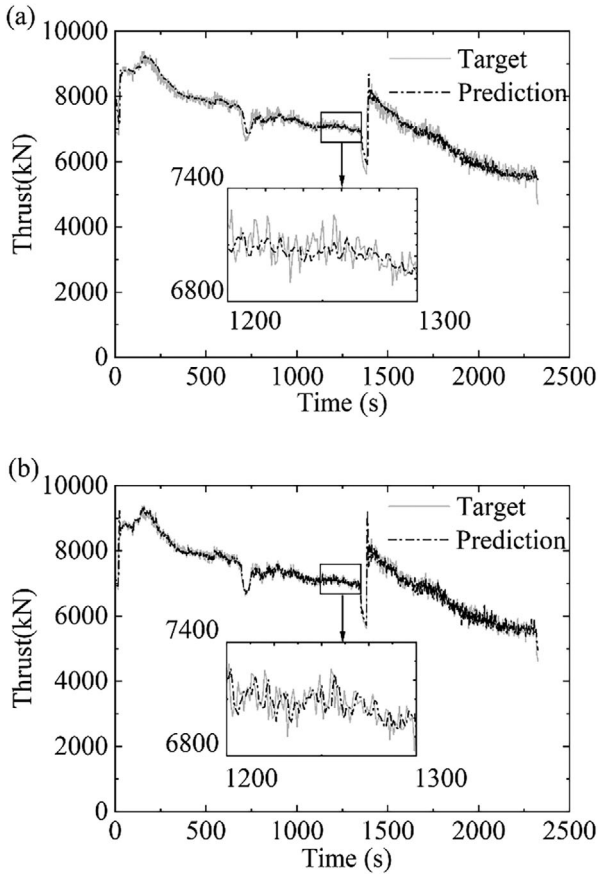

FIGURE 10 Comparison of target value and predictive value generated by different models:(a) SD-CLA model and (b) SD-CLA-OL model.

results, with an R^2 of 0.953 and an MAPE of 1.89%. The R^2 increased 2.03%, and the MAPE reduced 20.83% compared with the SD-CLA model. The predictions indicate that the model excels in capturing high-frequency features of thrust. However, it can be seen that it is still possible to generate abnormal values when abrupt changes occur in input data. This issue arises as the online learning section is highly sensitive to noisy data, which may be influenced by anomalous values, leading to potential degradation in overall prediction accuracy. In future research, it is advisable to embed a lightweight online outlier detection mechanism as a preprocessing layer in the model, thus enabling selective parameter updates or data rejection.

The normalized error distributions for the fifth-step prediction results of the SD-CLA model and SD-CLA-OL model are compared in Figure 11. A notable improvement can be observed in the error distribution with the introduction of the online learning section. Specifically, the distribution curve of the SD-CLA-OL model is significantly narrower and exhibits a sharper peak, which indicates a considerable enhancement in prediction accuracy and precision. The CIs of the normalized error with

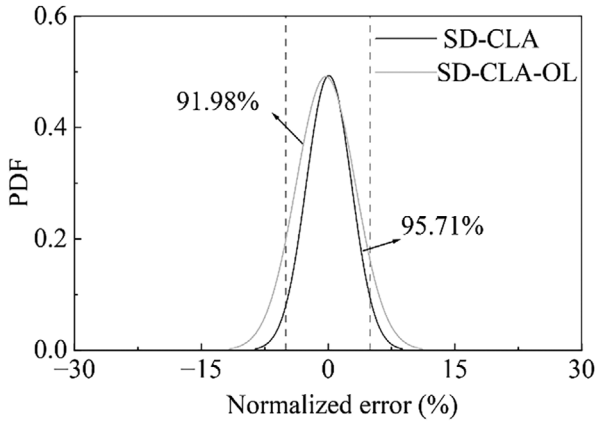

FIGURE 11 Error distributions of the prediction for the SD-CLA model and SD-CLA-OL model.

TABLE 3 Impact of the RFF dimension D on predictive performance.

Evaluation metrics	Value of D				
	50	**75**	**100**	**125**	**150**
MAPE (%)	2.03	1.94	1.89	1.89	1.89
R^2	0.944	0.949	0.953	0.953	0.953

the margin of error at 5% for the SD-CLA model and SD-CLA-OL model are 91.98 and 95.71, respectively. Furthermore, the tails of the distribution curve of the model with the online section extend further from the zero-error line, indicating that the SD-CLA model is more prone to large errors. Whereas the SD-CLA-OL model effectively reduces these extreme deviations. Additionally, the mean values of the two distribution curves are 0.0087% and 0.28%, respectively, indicating that the introduction of the online learning module notably decreases systematic bias in predictions.

The single-step inference time of the model was measured on a hardware configuration equipped with an Intel i5-12600K CPU and an NVIDIA RTX 4060 Ti GPU, representing a medium-computational-power setup typical of industrial field environments. The results demonstrate that the model (SD-CLA-OL) achieves a single-step prediction time of 50.6 ± 2.1 ms. This comfortably satisfies the real-time constraint—given the TBM thrust data sampling interval of 5 s, the model computation occupies only 1.19% of the time window. This leaves sufficient margin for dynamic control. Furthermore, the resource consumption of this method is manageable, with a peak GPU memory usage of 1.2 GB and stable CPU utilization below 23%. These data confirm the model's engineering deployability.

Furthermore, the impact of the dimension of the RFF is systematically evaluated in Table 3. Experiments show that as D increases from 50 to 100, model performance improves

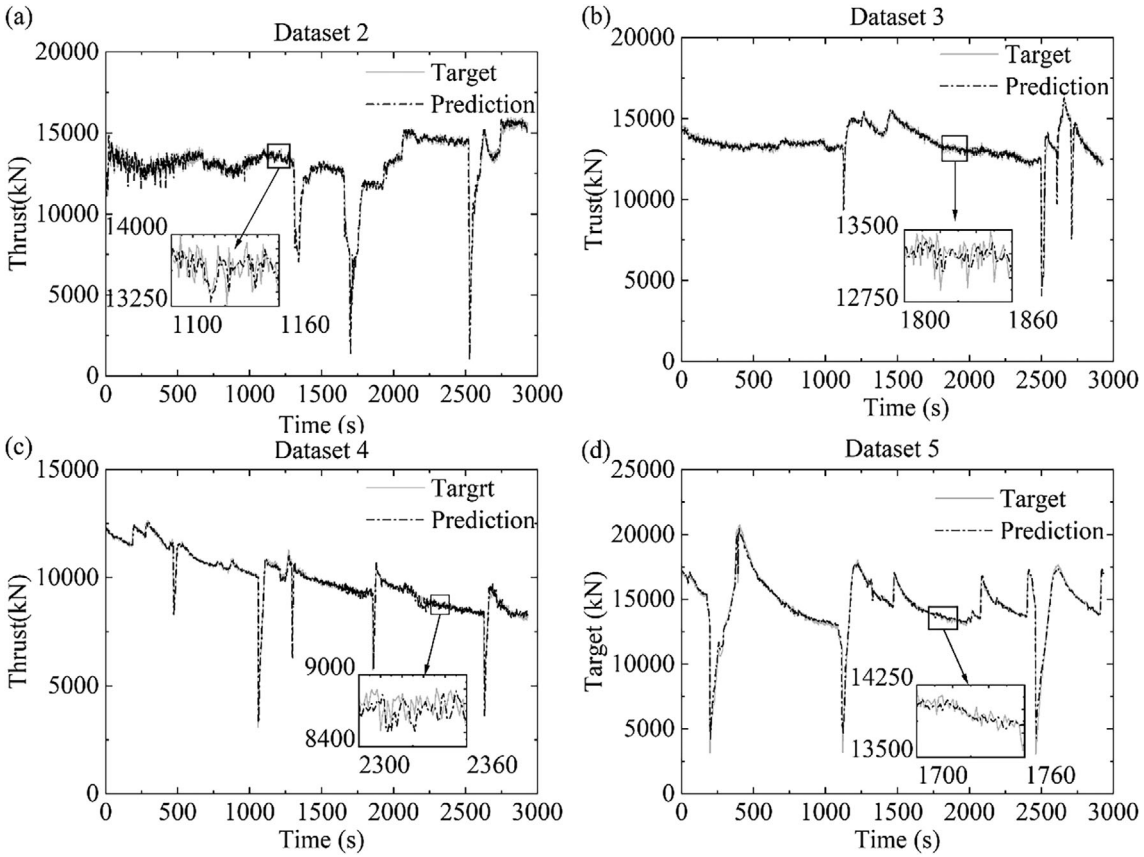

FIGURE 12 Fifth-step predictive result of the SD-CLA-OL model across different datasets.

TABLE 4 Impact of different kernels on predictive performance.

Evaluation metrics	Kernel		
	Matern	Laplacian	RBF
MAPE (%)	2.03	2.36	1.89
R^2	0.951	0.944	0.953

consistently, with R^2 rising from 0.944 to 0.953, and MAPE decreasing from 2.03% to 1.89%. Beyond $D = 100$, performance gains saturate and stabilize. This trend indicates that the method is insensitive to the specific value of D as long as the RFF dimension is sufficiently large to adequately approximate the target kernel function, demonstrating favorable robustness and reducing the difficulty of parameter tuning in practice.

Additionally, the influence of different kernel functions on the performance of the online module is evaluated in Table 4. The RBF and polynomial kernels demonstrated excellent and comparable efficacy under the same conditions; both attained R^2 values exceeding 0.95. The Laplacian kernel, however, exhibited a slight performance degradation, with an R^2 of 0.944 and a MAPE of 2.36%. These findings underscore the stability of the online module when used with suitable kernels and indicate that the widely applicable RBF kernel is a dependable choice for this architecture.

4.4 | Effect of the proposed model on different datasets

The predictive capability of the proposed framework is confirmed through a comparison of fifth-step forecasts with target values using Dataset 2–4. Table 5 displays the MAPE values and R^2 values at the fifth step of the CLA model, SD-CLA model, and SD-CLA-OL model. The average MAPE values across the four datasets are 2.55%, 2.39%, and 2.18%, respectively, while the average R^2 values are 0.885, 0.901, and 0.914, respectively. Compared to the CLA model, the SD-CLA model demonstrates an average MAPE reduction of 6.4% and an average R^2 improvement of 1.45%. In comparison to the CLA model, the SD-CLA-OL model achieves an average MAPE reduction of 14.4% and an average R^2 improvement of 3.26%. These results confirm the effectiveness of both the signal decomposition module and the online learning module in enhancing prediction accuracy and model robustness.

The fifth-step predictive performance of the SD-CLA-OL model in predicting thrust values across four datasets

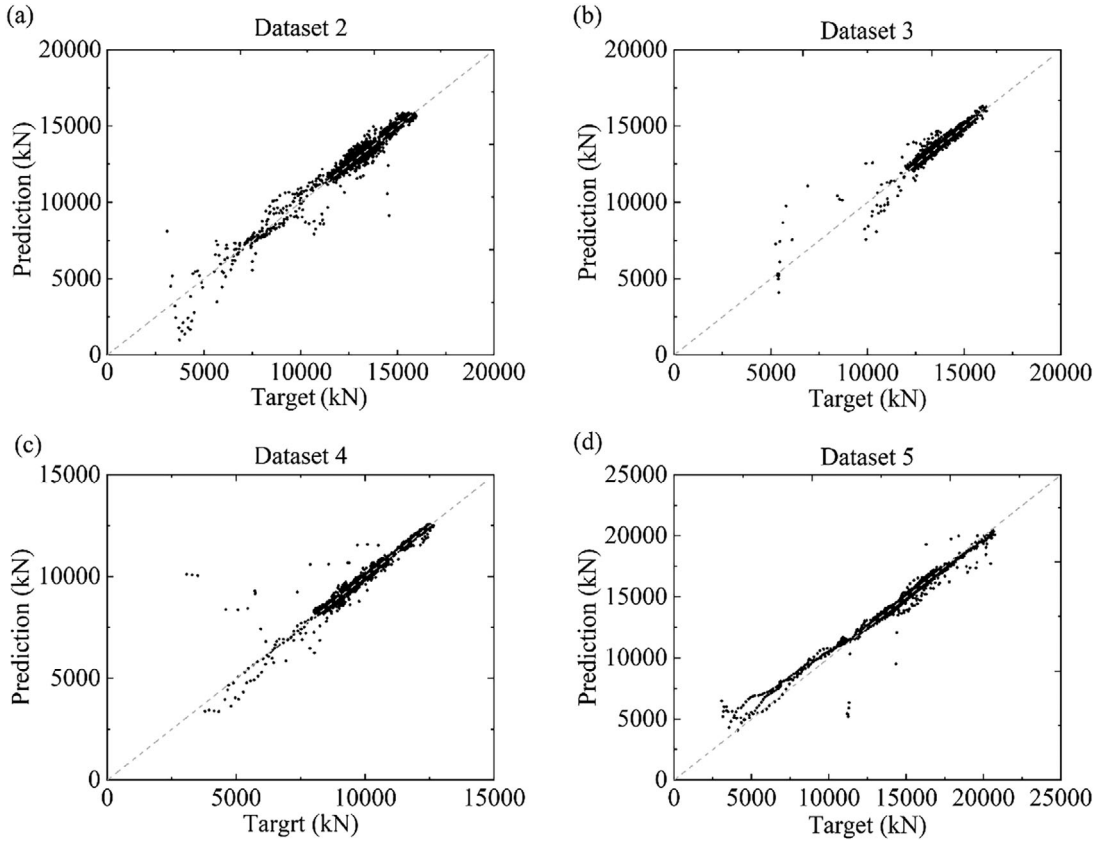

FIGURE 13 Scatter plots of the fifth-step prediction result for the SD-CLA-OL model across different datasets.

TABLE 5 Fifth-step predictive performance of three models in four additional datasets.

Dataset number	Models	Performance index	
		MAPE (%)	R^2
Dataset 2	CLA	3.12	0.885
	SD-CLA	2.85	0.907
	SD-CLA-ON	2.61	0.914
Dataset 3	CLA	1.82	0.896
	SD-CLA	1.73	0.912
	SD-CLA-ON	1.59	0.924
Dataset 4	CLA	2.41	0.863
	SD-CLA	2.23	0.877
	SD-CLA-ON	1.98	0.902
Dataset 5	CLA	2.88	0.897
	SD-CLA	2.75	0.908
	SD-CLA-ON	2.57	0.916

is displayed in Figure 12. The proposed model excels in both predicting steady-state phases and capturing sudden changes in thrust, such as sharp drops or spikes, highlighting its robustness in learning dynamic system behaviors. The MAPE values of each dataset are 2.61%, 1.59%, 1.98%, and 2.57% respectively. The R^2 values of each dataset are 0.914, 0.914, 0.902, and 0.916, respectively. In all datasets, the MAPEs are less than 3%, and the R^2 values exceed 0.9, which indicates the strong forecasting ability of the proposed model. The proposed model maintains a similar level of accuracy regardless of the complexity or variability of the data. The results validate the model's ability to deliver reliable and accurate predictions across multiple datasets, further highlighting its potential to be applied effectively in various tunneling projects with different geological conditions.

The scatter plot of the fifth-step prediction results for the SD-CLA-OL model across different datasets is shown in Figure 13. The perfect fitting line in the graph describes situations in which the predicted value equals the target value. Points located near the line suggest that the predicted values closely align with the recorded values. The majority of the predicted points are closely aligned with the perfect fitting line, indicating strong predictive performance. Additionally, most points are located symmetrically about the perfect fitting line, manifesting the absence of systematic errors. Furthermore, the clustering is consistent in regions corresponding to extreme thrust values, such as high or low outputs, which highlights the model's ability to effectively manage edge cases. Furthermore, the low occurrence of outliers demonstrates the

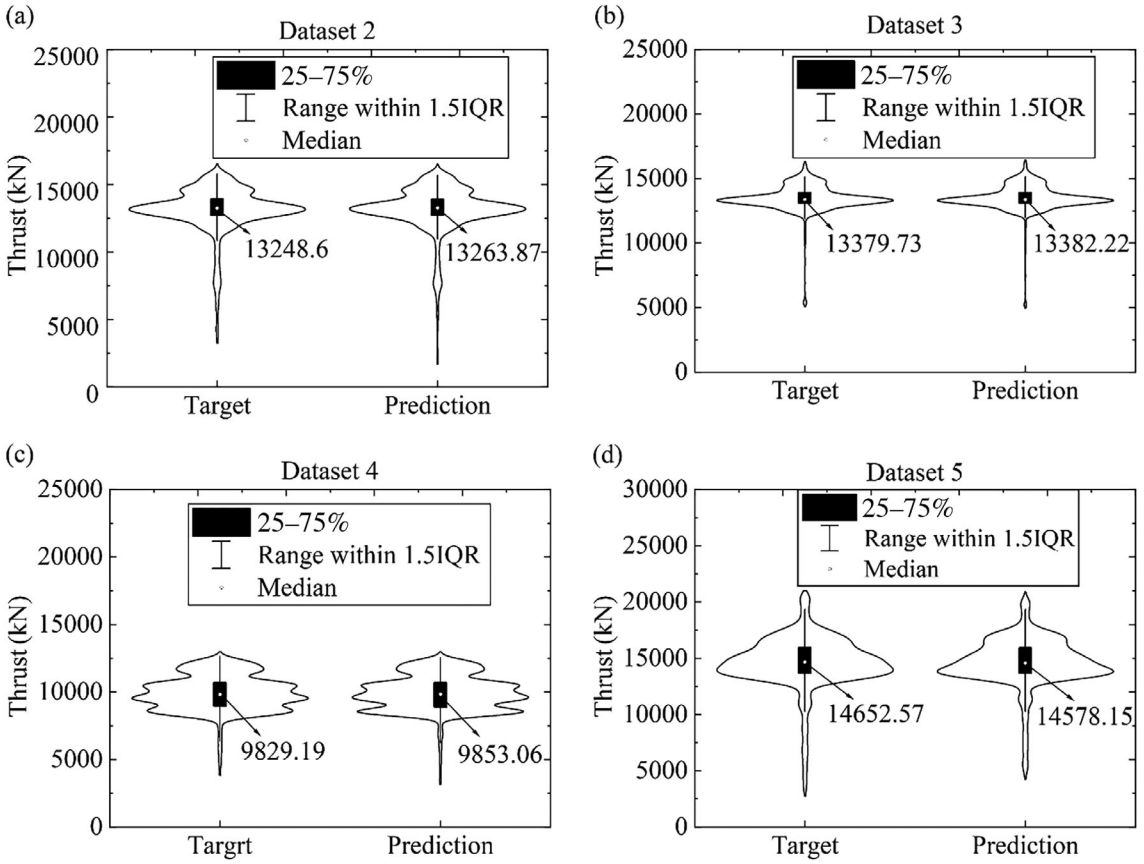

FIGURE 14 Violin plots of the fifth-step predictive results for the SD-CLA-OL model across different datasets. IQR, interquartile range.

model's reliability in avoiding large errors. These scatter plots also reveal the model's performance consistency across diverse datasets, which underscores its robustness, reliability, and adaptability in handling varying data conditions. The violin plot of the fifth-step prediction results for the SD-CLA-OL model across different datasets multiple datasets is shown in Figure 14. The median thrust values are represented by white dots, while the black bars indicate the interquartile range.

The violin shape displays the data's overall density distribution. Across all datasets, the medians of the predicted values closely match those of the actual values. The median differences between the predictive values and actual values are 15.27, 2.49, 23.87, and 74.42, respectively, highlighting the model's capability to maintain the statistical integrity of the data. The violin shapes show that the model captures both average trends and the distribution's spread and tails, which are crucial for detecting rare or extreme events. Additionally, the consistency in distribution replication across datasets suggests that the model generalizes effectively under varying conditions. This property is particularly valuable in real-world scenarios where understanding variability and uncertainty is equally important as predicting mean outcomes.

5 | CONCLUSION

This study proposes a real-time thrust prediction method for shield tunneling in the Dianzhong Water Diversion Project. A wavelet-based MRA first decomposes thrust signals into frequency components to enhance feature extraction. A hybrid CLA model then generates initial predictions, followed by an online learning module using SGD with RFF to predict residuals between preliminary estimates and actual values. The key findings are outlined below:

1. MRA decomposition (optimal level = 7) improves prediction accuracy, reducing MAPE from 2.84% to 2.44% and increasing R^2 from 0.901 to 0.930.
2. The online learning module further reduces MAPE to 1.89% and elevates R^2 to 0.953, confirming its effectiveness in multistep forecasting with an optimal output step of five.
3. Validation across diverse tunneling sections verifies the model's generalization capability under varying geological conditions.

This improvement of thrust prediction accuracy enables proactive safety control and dynamic optimization of

excavation parameters, effectively mitigating equipment risks, extending service life, and minimizing environmental disturbances. This technology, underpinned by a data-driven closed-loop decision framework, facilitates efficient tunneling in complex geology, while its accumulated data assets further advance the digital transformation of tunneling engineering.

Although the proposed method has reached satisfying accuracy in multistep real-time predictions, there is potential for further enhancement in future research. First, the proposed framework needs to be deployed in new tunneling projects with diverse geological conditions, utilizing multi-source engineering data to further verify its robustness and adaptability. Additionally, this study focuses solely on predicting thrust data. In future research, the interdependencies among various types of information could be considered in multi-output models.

ACKNOWLEDGMENTS

The authors acknowledge the support of the National Natural Science Foundation of China under grant number 42377140.

REFERENCES

Abdoos, A. A. (2016). A new intelligent method based on combination of VMD and ELM for short term wind power forecasting. *Neurocomputing*, *203*, 111–120. https://doi.org/10.1016/j.neucom.2016.03.054

Ali, M., & Parasad, R. (2019). Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. *Renewable and Sustainable Energy Reviews*, *104*, 281–295. https://doi.org/10.1016/j.rser.2019.01.014

Altan, A., Karasu, S., & Zio, E. (2021). A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. *Applied Soft Computing*, *100*, 106996. https://doi.org/10.1016/j.asoc.2020.106996

Chen, Z. Y., Zhang, Y. P., Li, J. B., Li, X., & Jiang, L. J. (2021). Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: A case study on the Yinsong Project. *China. Tunnelling and Underground Space Technology*, *108*, 103700. https://doi.org/10.1016/j.tust.2020.103700

Chen, S. G., & Zhao, J. (2002). Modeling of tunnel excavation using a hybrid DEM/BEM method. *Computer-Aided Civil and Infrastructure Engineering*, *17*(5), 381–386. https://doi.org/10.1111/1467-8667.00284

Cui, M. D., Wu, G., Chen, Z. Q., Dang, J., Zhou, M. H., & Feng, D. M. (2021). Geometric attention regularization enhancing convolutional neural networks for bridge rubber bearing damage assessment. *Journal of Performance of Constructed Facilities*, *35*(5), 0001634. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001634

Engel, Y., & Mannor, S. (2004). The Kernel recursive least-squares algorithm. *IEEE Transactions on Signal Processing*, *52*(8), 2275–22285. https://doi.org/10.1109/TSP.2004.830985

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. https://doi.org/10.1016/j.tust.2020.103636

Finger, S., & Dixon, J. R. (1989). A review of research in mechanical engineering design. Part I: Descriptive, prescriptive, and computer-based models of design processes. *Research in Engineering Design*, *1*, 51–67. https://doi.org/10.1007/BF01580003

Gao, X. J., Shi, M. L., Song, X. G., Zhang, C., & Zhang, H. W. (2019). Recurrent neural networks for real-time prediction of TBM operating parameters. *Automation in Construction*, *98*, 225–235. https://doi.org/10.1016/j.autcon.2018.11.013

Gepperth, A., & Wiech, F. (2019). Simplified computation and interpretation of Fisher matrices in incremental learning with deep neural networks. Paper presented at the 28th International Conference on Artificial Neural Networks, Munich.

Guo, D., Li, J. H., Jiang, S. H., Li, X., & Chen, Z. Y. (2022). Intelligent assistant driving method for tunnel boring machine based on big data. *Acta Geotechnica*, *17*, 1019–1030. https://doi.org/10.1007/s11440-021-01327-1

Han, M. D., Cai, Z. X., Qu, C. Y., & Jin, L. S. (2017). Dynamic numerical simulation of cutterhead loads in TBM tunnelling. *Tunnelling and Underground Space Technology*, *70*, 286–298. https://doi.org/10.1016/j.tust.2017.08.028

Huang, H., Ruan, B., Wu, X. G., & Qin, Y. W. (2024). Parameterized modeling and safety simulation of shield tunnel based on BIM-FEM automation framework. *Automation in Construction*, *162*, 105362. https://doi.org/10.1016/j.autcon.2024.105362

Jin, Y. R., Qin, C. J., Tao, J. F., & Liu, C. L. (2022). An accurate and adaptative cutterhead torque prediction method for shield tunneling machines via adaptative residual long-short term memory network. *Mechanical Systems and Signal Processing*, *165*(15), 108312. https://doi.org/10.1016/j.ymssp.2021.108312

Kang, Q., Chen, E. J., Li, Z. C., Luo, H. B., & Liu, Y. (2023). Attention-based LSTM predictive model for the attitude and position of shield machine in tunneling. *Underground Space*, *13*, 335–350. https://doi.org/10.1016/j.undsp.2023.05.006

Kivinen, J., Smola, A. J., & Williamson, R. C. (2001). Online learning with kernels. In T. G. Dietterich, S. Becker, Z. Ghahramani (Eds.), *Proceedings of the 14th international conference on neural information processing systems: natural and synthetic* (pp. 785–792). MIT Press.

Kubota, Y., Yabuki, N., & Fukuda, T. (2024). Autopilot model for shield tunneling machines using support vector regression and its application to previously constructed tunnels. *Computer-Aided Civil and Infrastructure Engineering*, *39*(1), 46–62. https://doi.org/10.1111/mice.13019

Lislerud, A. (1988). Hard rock tunnel boring: Prognosis and costs. *Tunnelling and Underground Space Technology*, *3*(1), 9–17. https://doi.org/10.1016/0886-7798(88)90029-6

Liu, W. L., & Ding, L. Y. (2020). Global sensitivity analysis of influential parameters for excavation stability of metro tunnel. *Automation in Construction*, *113*, 103080. https://doi.org/10.1016/j.autcon.2020.103080

Mehul, S., Shrid, P., Priety, Y., Deepak, K., Nitin, G., & Gautam, S. (2023). Advancing security in the industrial Internet of Things using deep progressive neural networks. *Mobile Networks and Applications*, *28*, 782–794.

Melis, M., Medina, L., & Rodriguez, J. (2002). Prediction and analysis of subsidence induced by shield tunnelling in the Madrid Metro extension. *Canadian Geotechnical Journal*, *39*(6), 1273–1287. https://doi.org/10.1139/t02-073

Pan, Y., Fu, X. L., & Zhang, L. M. (2022). Data-driven multi-output prediction for TBM performance during tunnel excavation: An attention-based graph convolutional network approach. *Automation in Construction*, *141*, 104386. https://doi.org/10.1016/j.autcon.2022.104386

Park, S., Jang, H., Simeone, O., & Kang, J. (2021). Learning to demodulate from few pilots via offline and online meta-learning. *IEEE Transactions on Signal Processing*, *69*, 226–239. https://doi.org/10.1109/TSP.2020.3043879

Qi, W. C., Wang, L. H., Zhou, S. Y., Kang, Y. L., & Zhang, Q. (2022). Total loads modeling and geological adaptability analysis for mixed soil-rock tunnel boring machines. *Underground Space*, *7*(3), 337–351. https://doi.org/10.1016/j.undsp.2021.09.002

Qin, C. J., Huang, G., Yu, H. G., Zhang, Z. N., & Tao, J. F. (2024). Adaptive VMD and multi-stage stabilized transformer-based long-distance forecasting for multiple shield machine tunneling parameters. *Automation in Construction*, *165*, 105563. https://doi.org/10.1016/j.autcon.2024.105563

Qin, C. J., Shi, G., Tao, J. F., Yu, H. G., Jin, Y. R., Xiao, D. Y., Zhang, Z. N., & Liu, C. L. (2022). An adaptive hierarchical decomposition-based method for multi-step cutterhead torque forecast of shield machine. *Tunnelling and Underground Space Technology*, *175*, 109148.

Qin, Z. D., Chen, B. D., & Zheng, N. N. (2017). Random Fourier feature kernel recursive least squares. In *Proceedings of the 2017 International joint conference on neural networks (IJCNN)*, Anchorage (pp. 2881–2886). IEEE. https://doi.org/10.1109/IJCNN.2017.7966212

Rahimi, A., & Recht, B. (2007). Random features for large-scale kernel machines. In *Proceedings of the 20th international conference on neural information processing systems* (pp. 1177–1184). Curran Associates.

Shi, C., & Wang, Y. (2021). Development of subsurface geological cross-section from limited site-specific boreholes and prior geological knowledge using iterative convolution XGBoost. *Journal of Geotechnical and Geoenvironmental Engineering*, *147*(9), 0002583. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002583

Shi, G., Qin, C. J., Tao, J. F., & Liu, C. L. (2021). AVMD-EWT-LSTM-based multi-step prediction approach for shield tunneling machine cutter head torque. *Knowledge-Based Systems*, *228*, 107213. https://doi.org/10.1016/j.knosys.2021.107213

Shi, G., Qin, C. J., Zhang, Z. N., Yu, H. G., Tao, J. F., & Liu, C. L. (2024). A novel decomposition and hybrid transfer learning-based method for multi-step cutterhead torque prediction of shield machine. *Mechanical Systems and Signal Processing*, *214*, 111362. https://doi.org/10.1016/j.ymssp.2024.111362

Tondini, N., Bursi, O. S., Bonelli, A., & Fassin, M. (2015). Capabilities of a fiber Bragg grating sensor system to monitor the inelastic response of concrete sections in new tunnel linings subjected to earthquake loading. *Computer-Aided Civil and Infrastructure Engineering*, *30*(8), 636–653. https://doi.org/10.1111/mice.12106

Wang, K. Y., Zhang, L. M., & Fu, X. L. (2023). Time series prediction of tunnel boring machine (TBM) performance during excavation using causal explainable artificial intelligence (CX-AI). *Automation in Construction*, *147*, 104730. https://doi.org/10.1016/j.autcon.2022.104730

Wang, R. H., Li, D. Q., Chen, E. J., & Liu, Y. (2021). Dynamic prediction of mechanized shield tunneling performance. *Automation in Construction*, *132*, 103958. https://doi.org/10.1016/j.autcon.2021.103958

Wang, S. Y., Yuan, X., & Qu, T. M. (2024). Machine learning–informed soil conditioning for mechanized shield tunneling. *Computer-Aided Civil and Infrastructure Engineering*, *39*(17), 2662–2682. https://doi.org/10.1111/mice.13152

Wu, L., Guan, T. M., & Lei, L. (2013). Discrete element model for performance analysis of cutterhead excavation system of EPB machine. *Tunnelling and Underground Space Technology*, *37*, 37–44. https://doi.org/10.1016/j.tust.2013.03.003

Zhang, P., Yin, Z. Y., Zhang, Y. Y., & Gao, F. P. (2020). A LSTM surrogate modelling approach for Caisson foundations. *Ocean Engineering*, *204*, 107263. https://doi.org/10.1016/j.oceaneng.2020.107263

Zhang, Q., Qu, C. Y., Cai, Z. X., Kang, Y. L., & Huang, T. (2014). Modeling of the thrust and torque acting on shield machines during tunneling. *Automation in Construction*, *40*, 60–67. https://doi.org/10.1016/j.autcon.2013.12.008

Zhang, R. Y., Liu, Y., & Sun, H. (2020). physics-guided convolutional neural network (PhyCNN) for data-driven seismic response modeling. *Engineering Structures*, *215*, 110704. https://doi.org/10.1016/j.engstruct.2020.110704

Zhang, Y. P., Chen, Z. Y., JIN, F., Jing, L. J., Xing, H., & Li, P. Y. (2023). Cross-project prediction for rock mass using shuffled TBM big dataset and knowledge-based machine learning methods. *Science China Technological Sciences*, *66*, 751–770. https://doi.org/10.1007/s11431-022-2290-7

Zhao, S., Zhang, D. M., Xue, Y. D., Zhu, M. L., & Huang, H. W (2021). A deep learning-based approach for refined crack evaluation from shield tunnel lining images. *Automation in Construction*, *132*, 103934. https://doi.org/10.1016/j.autcon.2021.103934

How to cite this article: Wu, Y., Yang, H., Zhang, H., Hou, Y., & Yang, S. (2025). Real-time prediction of tunnel boring machine thrust based on multi-resolution analysis and online learning. *Computer-Aided Civil and Infrastructure Engineering*, 1–16. https://doi.org/10.1111/mice.70096