International Core Journal of Engineering Volume 11 Issue 12, 2025
ISSN: 2414-1895 DOI: 10.6919/ICJE.202512_11(12).0013

Study on Cutter Torque Prediction based on Bayesian Optimisation Random Forests

Yifei Song[1], Liying Deng[1], Shutao Yu[2], and Haoqiang Zhang[3]

[1] Shenyang University of Technology, Liaoyang 110000, China

[2] Hefei Vocational College of Science and Technology, Hefei 231201, China

[3] CCCC Tianhe Machinery Equipment Manufacturing Co., Ltd., Changshu 215500, China

Abstract

During shield tunnelling operations, cutterhead torque serves as a critical parameter reflecting ground conditions, equipment performance, and construction risks. Accurate torque prediction provides operators with technical guidance, enables early warning of potential anomalies, and reduces failure risks. Traditional machine learning methods have been applied in engineering prediction, but limitations remain in their predictive accuracy and parameter dependency. This paper constructs a Bayesian-optimised Random Forest (BO-RF) model based on actual engineering data to predict cutterhead torque, comparing it with conventional RF, BR, and SVR models. Unified evaluation metrics (MAE, RMSE, R^2) are established to analyse the predictive performance of the four models. Results demonstrate that the RF-BO model significantly enhances prediction performance through automated hyperparameter optimisation, exhibiting superior fitting accuracy and robustness compared to other models. This study confirms that Bayesian optimisation effectively leverages the strengths of random forests, providing a feasible and efficient technical approach for high-precision cutterhead torque prediction and construction safety in shield tunnelling.

Keywords

Railway Tunnels; Bayesian Optimisation; Random Forests; Machine Learning; Construction Safety.

1. Introduction

In recent years, the accelerated advancement of artificial intelligence, coupled with national industrial digitalisation strategies, has seen AI progressively emerge as a new productive force within societal production. The concept of intelligent construction has also gradually permeated multiple sectors, including tunnel engineering[1] .Taking full-face tunnel boring machines (TBMs) as an example, their construction processes continuously and densely record operational information such as advance rate, total thrust, and cutterhead torque, generating typical multi-source time-series data. Such data not only encompasses multifaceted information including construction conditions and geological factors but also often exhibits pronounced nonlinear relationships and mutual coupling between different variables.Constrained by multiple factors, traditional models exhibit suboptimal predictive accuracy, struggling to fully reflect actual variations in complex operating conditions, thus maintaining limited applicability and precision. Unlike conventional methods reliant on empirical assumptions, machine learning can autonomously capture intricate mapping patterns from measured data, yielding more reliable outcomes in underground engineering prediction and analysis.

Leveraging these advantages, numerous scholars have applied machine learning methods to TBM construction data analysis and excavation parameter prediction, yielding positive outcomes[2-3] , providing crucial references for this study.Salimi et al.[4] developed an adaptive neural fuzzy inference system (ANFIS) and support vector regression (SVR) model based on principal component analysis (PCA) using Iranian hard rock tunnel engineering data. Final results demonstrated a high correlation between PCA-SVR predictions and TBM performance. Zhang Yan et al.[5] constructed a nonlinear relationship between characteristic factors and excavation speed based on support vector machines, thereby providing a novel approach for predicting TBM excavation speed. Zhu Mengqi et al.[6] employed a coupled CART-RF model to predict TBM parameters, achieving a prediction accuracy of 0.9. Kong et al.[7] employed the random forest method to forecast cutterhead torque; Yu et al.[8] considered the correlation and divergence between thrust and cutterhead torque, designing a multi-channel decoupled neural network model to predict both thrust and cutterhead torque; Yin Zhiqing et al.[9] analysed the influence patterns of thrust, soil chamber pressure, penetration rate variation, and cutter replacement on cutterhead torque based on the sign regression algorithm.

In summary, existing research has demonstrated the feasibility of machine learning in predicting TBM excavation parameters, achieving progress in feature selection, model construction, and prediction accuracy enhancement. However, current work exhibits the following limitations: (1) most studies rely on empirical methods or grid search for hyperparameter tuning, which is inefficient and prone to suboptimal solutions; (2) systematic comparative analyses of different models regarding accuracy, stability, and generalisability are lacking.

To address these issues, this paper proposes a Bayesian Optimisation Random Forest (BO-RF) model using real engineering data, comparing it with traditional Random Forest (RF) and Bayesian Regression (BR). This method employs the Optuna framework for efficient global optimisation of random forest hyperparameters, while integrating automatic time-series feature engineering (lag, sliding window statistics, and differencing) to enhance the model's ability to capture the nonlinear and dynamic characteristics of the tunnelling process.During the prediction phase, the model further outputs uncertainty bands based on ensemble variance, providing quantitative grounds for result interpretation and risk warning.

2. Model Development

2.1 Principles of Bayesian Optimisation

Bayesian Optimization (BO) is a black-box optimization method based on probabilistic modeling [10], suitable for efficient optimization in scenarios where the analytical form of the objective function is unknown and computational costs are high. Its core principle involves constructing surrogate models within the hyperparameter space, achieving a dynamic equilibrium between exploration and exploitation through function sampling. As the surrogate models are continuously updated, the search process progressively converges toward the global optimum. Compared to traditional grid search methods, Bayesian optimization demonstrates superior advantages in optimization efficiency and global convergence capability.

Optuna is an automated hyperparameter optimization framework proposed by Akiba et al.[11] in 2019. It implements Bayesian optimization based on the TPE (Tree-structured Parzen Estimator) algorithm, enabling adaptive adjustment between exploration and exploitation to achieve efficient hyperparameter search. Compared to traditional methods, Optuna exhibits faster convergence and stronger scalability, demonstrating particularly outstanding performance in complex model tuning scenarios.

Therefore, this study introduces Bayesian optimization and the Optuna framework to enhance the optimization efficiency of the random forest model across multiple hyperparameter spaces, thereby avoiding the issues of dimensionality disaster and computational resource waste commonly encountered in traditional grid search. Leveraging Optuna's automated search and pruning mechanisms, the model training process becomes more standardized and reproducible, facilitating the acquisition of optimal model configurations within a shorter timeframe. This lays the foundation for subsequent performance improvements in the TBM data prediction model.

Table 1. Comparison of Hyperparameter Optimisation Methods.

Parameter	Optuna Method	Grid Search Method
Flexibility	High	Low
Hyperparameter search range	Continuous, discrete parameters, etc. automatically defined	Manually defining parameter combinations
Search Method	TPE	Exhaustive Iterative Search
Convergence Speed	Rapid	Slow
Applicability	Diverse Model Optimisation	Simple model tuning

2.2 Principles of Random Forest.

Random Forest (RF) is a quintessential ensemble learning approach, fundamentally designed to enhance predictive accuracy and model robustness by combining multiple decision trees. During training, RF introduces randomness at both the sample and feature levels: firstly, employing bootstrap sampling to generate multiple subsamples from the original training set, upon which multiple decision trees are trained independently;Secondly, at each node split within a tree, the optimal splitting variable is selected from a randomly sampled subset of features rather than comparing all features. This "dual randomisation" mechanism ensures sufficient diversity among individual trees, effectively mitigating the high variance inherent in single decision trees.

For regression tasks, the random forest averages the predictions from all decision trees, using this final average as the output. This achieves low overall variance and strong generalisation capability:

$$y = \frac{1}{n}\sum_{i=1}^{n} y_i(x) \tag{1}$$

n denotes the number of trees.

yi(x): The predicted value from the ith tree.

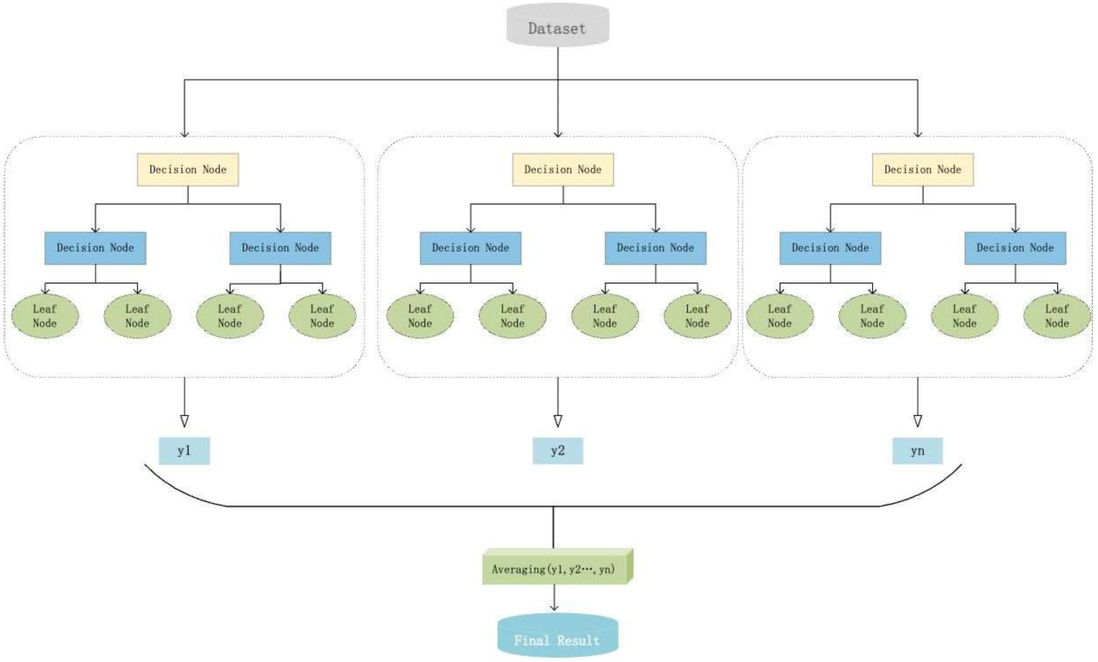


Fig. 1 Schematic diagram of the RF algorithm structure

Compared to single-tree decision trees, random forests better capture complex nonlinear relationships between input features and output variables, while demonstrating greater stability in handling noise and preventing overfitting. By evaluating each feature's contribution to improving predictive performance during splitting, corresponding importance metrics are provided, offering valuable reference for engineering problems.

2.3 Bayesian Optimisation for Random Forest Model Construction

Random forests exhibit strong nonlinear fitting capabilities and robust performance in engineering prediction. However, their effectiveness heavily relies on the appropriate tuning of hyperparameters (such as the number of trees, maximum depth, and minimum leaf node sample size). Relying solely on manual experience or fixed values often fails to strike a balance between prediction accuracy and generalisation ability, potentially leading to underfitting or overfitting.

To address these issues, this paper employs the Optuna framework to implement Bayesian optimisation, constructing a Bayesian Optimised Random Forest (BO-RF) model for TBM cutterhead torque prediction. The implementation involves feature engineering of TBM construction data, including the construction of multi-source time-series features such as lagged terms, sliding statistics, and differences. Subsequently, cross-validation is applied to the training set to select optimal hyperparameters.2as the objective function, iteratively updating the hyperparameter combinations of the random forest through Optuna. Each generation comprises a cycle of "parameter sampling – model training – performance evaluation – proxy model update" until the optimal configuration is attained. Finally, the optimal model is employed for prediction on the test set, outputting evaluation metrics, feature importance, and uncertainty intervals for prediction results, thereby achieving high-precision prediction and robustness verification for TBM cutterhead torque.

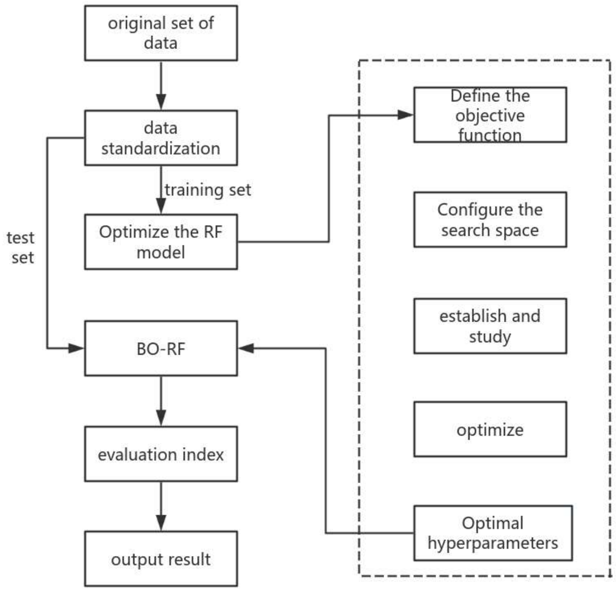


Fig. 2 BO-RF Modelling Process

3. Data Preprocessing

3.1 Data Selection

Data quality is a critical determinant of predictive performance during machine learning model training. Raw data collected at TBM construction sites often contains noise, missing values, and outliers. Without processing, these issues severely impact model convergence speed and prediction accuracy. Therefore, systematic data preprocessing must be conducted prior to modelling to ensure the reliability and validity of input data.

During TBM tunnelling, alongside primary excavation operations, auxiliary processes occur such as replacing support shoes, performing rock support work, changing cutting tools, and halting operations for equipment inspection and maintenance. Data generated from these non-excavation activities, when mixed with data collected during tunnelling, may cause confusion or mislead predictions of TBM excavation parameters, thereby affecting the accuracy of cutterhead torque forecasting.Therefore, special treatment is required to prevent interference with data analysis results. Using cutterhead thrust F, cutterhead torque T, cutterhead rotational speed N, advance rate V, and penetration rate P as indicators to determine whether the TBM is in a non-excavation state, a binary state discrimination function is constructed. When the constructed discrimination function equals 0, the TBM is deemed to be in a non-excavation state, and the data must be discarded;When the function equals 1, the TBM is deemed operational, and the data is retained. The discriminant function is defined as:

$$D = f(F)\,f(V)\,f(T)\,f(N) \tag{2}$$

$$f(x) = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0 \end{cases} \tag{3}$$

In the formula, F represents the total thrust of the TBM in kN; V denotes the advance speed in $\mathrm{m \cdot s^{-1}}$; T indicates the cutterhead torque in $\mathrm{kN \cdot m}$; and N signifies the cutterhead rotational speed in $\mathrm{r \cdot min^{-1}}$.

To eliminate non-steady-state data from start-up and shutdown phases, this study applies temporal rules at the cycle granularity: discarding short cycles with total duration <10 minutes; removing the first 5 minutes before start-up and the last 3 minutes before shutdown within retained cycles, retaining only stable excavation samples from the central segment.Following this screening, auxiliary columns generated during intermediate calculations were removed, and the data exported as the _cleaned dataset. A cleaning report (original rows, processed rows, total cycles, valid cycles, final sample size) and column name mapping were generated to facilitate experiment replication and auditing.

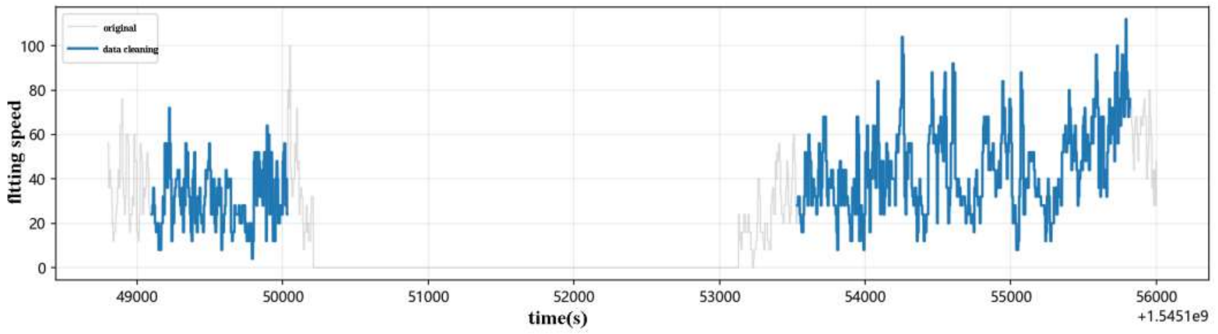


Fig.3 Data Processing Diagram

3.2 Data Outlier Handling

Given the complex and variable construction environment, collected signals often contain noise and outliers. During data preprocessing, statistical methods are employed to identify and handle outliers, preventing these from distorting overall distribution characteristics. A common approach involves setting the data range between [μ−3σ, μ+3σ], where μ is the feature mean and σ is the standard deviation. Records falling outside this interval are deemed outliers and excluded.The processed data provides a reliable foundation for subsequent modelling.

Table 2. Selected measured tunnelling parameters

Cutterhead torque	Tunnelling speed	Thrust
1191	48	8903
1128	48	7043
1042	48	8372
1887	44	7508
1598	40	6777
1240	44	7242
2041	24	7973
1738	52	7442
2323	28	9103

The aforementioned preprocessing workflow significantly enhances data quality, establishing a robust data foundation for subsequent modelling and forecasting. Selected measured tunnelling parameters:

3.3 Wavelet Noise Reduction

Processed data typically contains noise. To ensure the accuracy of parametric analysis, further denoising is required. This paper employs discrete wavelet transform for denoising the cleaned data. This signal processing technique, based on wavelet analysis, decomposes the signal using wavelet bases. By exploiting the distinct characteristics of noise and useful signals in wavelet coefficients, it processes these coefficients via thresholding methods before reconstructing the signal. This achieves the goal of removing noise while preserving signal features.The wavelet denoising process is illustrated in Fig. 4.

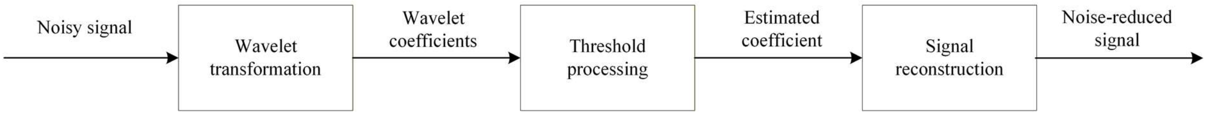


Fig. 4 Wavelet Noise Reduction Process

The Daubechies wavelet (sym8) function is employed to perform wavelet decomposition on the signal, yielding wavelet coefficients. The decomposed result can be expressed as:

$$x(n)=\sum_{k} c_{j,k}\,\varphi_{j,k}(n)+\sum_{j=1}^{J}\sum_{k} d_{j,k}\,\phi_{j,k}(n) \tag{4}$$

where denotes the original signal; represents the scale coefficient; signifies the wavelet coefficient; denote the scale index and position index respectively; and denote the scale function and wavelet function respectively. The wavelet basis determines the wavelet function.

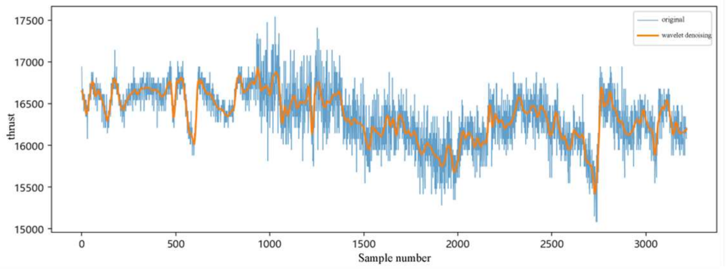


Fig. 5 Thrust wavelet denoising diagram

4. Performance Evaluation

4.1 Evaluation Metrics

To comprehensively evaluate the performance of the Bayesian Optimisation Random Forest (BO-RF) model, this paper selects Bayesian Ridge (BR), Support Vector Machine Regression (SVR), and Random Forest (RF) without Bayesian optimisation as baseline models.

To systematically evaluate each model's performance in predicting TBM cutterhead torque, all models employed consistent data and feature engineering configurations. Assessments were conducted using identical training and testing set divisions, with comparative analysis conducted through both visualisation and quantitative metrics. Visualisation aspects included:

(1) Actual-predicted scatter plots: examining overall fit consistency and outliers.

(2) Time-series overlay plots (with uncertainty bands): Comparing prediction accuracy and interval stability at each time step across the test set.

(3) Quantitative metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2) are employed. These measure the average absolute deviation between predicted and actual values:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

where $\hat{y}_i$ represents the predicted value for the ith sample, and yi denotes the actual value for the ith sample.

RMSE amplifies the influence of outliers by squaring errors before averaging, restoring the same units as the target variable:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

The coefficient of determination measures a model's ability to explain the variance in the target variable:

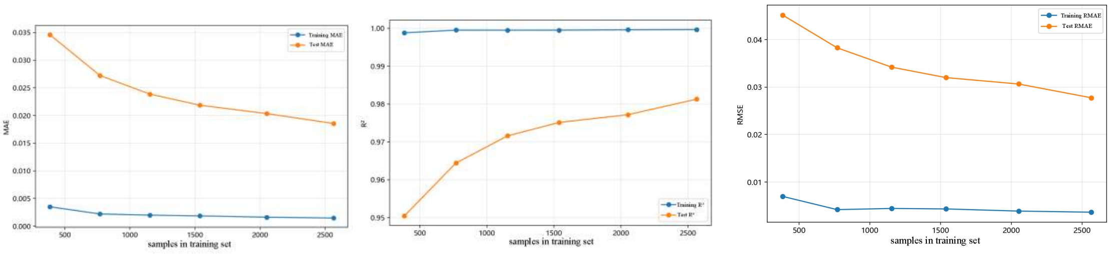
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7}$$

R^2= 1: Perfect fit.

R2= 0: The model fails to explain variation in the target variable.

4.2 Visualisation of Results Comparison and Analysis

4.2.1 Model Learning Process and Generalisation Capacity Analysis



(a) Learning Curve (MAE) (b) Learning Curve (R^2) (c) Learning Curve (RMSE)

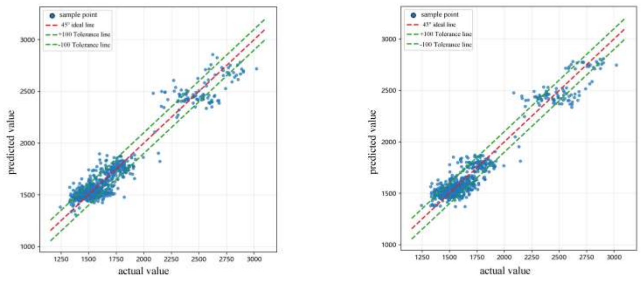
Fig. 6 Learning Curve

Fig. 6 demonstrates that as the number of training samples increases, both the training and validation sets exhibit synchronous decreases in MAE and RMSE, ultimately stabilising. The minimal disparity between these metrics indicates that the model maintains robust generalisation performance while enhancing its fitting capability, without exhibiting pronounced underfitting or overfitting. TheR^2curve continues to rise steadily with increasing samples before levelling off, confirming the BO-RF model's effective explanation of target variable variations.Collectively, all three metrics demonstrate the model's training effectiveness and generalisation capability.

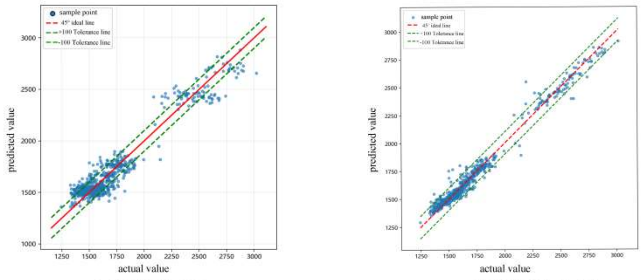
4.2.2 Overall Fitting Comparison

Fig. 7 presents the scatter plots for Bayesian Ridge, Random Forest (RF), and Bayesian Optimised Random Forest (BO-RF) in cutter disc torque prediction. Overall, differences exist in the degree of fit between predicted and actual values across models.

The scatter plot for Bayesian Ridge exhibits greater dispersion, with some points noticeably deviating from the ideal 45° line and the ±100 tolerance band. This indicates that while the model captures the general trend of cutter disc torque variations with respect to features, its predictive accuracy remains insufficient, resulting in significant fluctuations in outcomes.



(a) RF Model (b) BR model



(c) SVR Model (d) BO-RF Model

Fig.7 Model Scatter Plots

Random Forest model exhibits greater stability. The majority of data points cluster near the 45° ideal line, with most samples falling within the tolerance range. This demonstrates the Random Forest's strong fitting capability when handling non-linear relationships. However, some deviation is observed in locally high-fluctuation intervals.

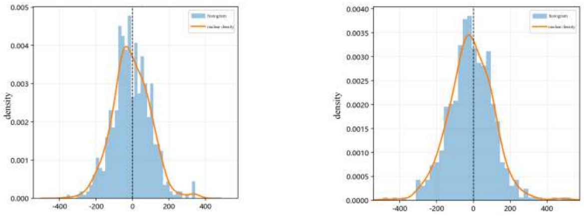
The Support Vector Regression (SVR) model also demonstrated strong overall nonlinear handling capability, with scatter points largely distributed along the 45° ideal line. Nevertheless, result fluctuations were more pronounced, and some points deviated from the tolerance band, indicating insufficient stability when processing complex multidimensional features.

By contrast, the BO-RF model delivered superior prediction performance. Scatter points clustered almost entirely near the ideal 45° line, highly concentrated within the centre of the tolerance band with uniform and dense distribution. This demonstrates that the Bayesian-optimised random forest ensures model robustness, significantly enhances prediction accuracy, and captures the variation patterns of cutter disc torque with greater precision.

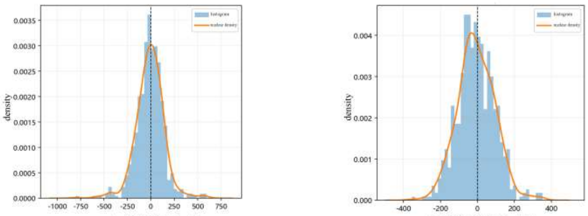
Overall, all three models can characterise the trend of cutter disc torque to some extent, but their precision and stability differ markedly, with the BO-RF model demonstrating the best overall performance.

4.2.3 Residual Analysis

Residual histograms aggregate the prediction errors (residual = predicted value − actual value) for each sample into a distribution, enabling a visual assessment of the magnitude and stability of model errors. A narrower distribution with a higher peak indicates greater stability, accompanied by relatively smaller RMSE and MAE values. A distribution skewed to the left or right suggests systematic bias, i.e., consistent overestimation or underestimation. A distribution exhibiting long tails indicates poorer model robustness compared to.



(a) RF Model (b) BR model



(c) SVR Model (d) BO-RF Model

Fig. 8 Residual Histogram

Fig. 8 reveals: BR (Bayesian Regression): Exhibits the broadest distribution with relatively thicker tails, indicating greater overall error and occasional substantial deviations; the peak slightly deviates from zero, suggesting a mild systematic bias.

RF (Random Forest): Narrower than BR with a peak closer to zero, indicating improved overall error and stability; however, a relatively long tail persists on one side, suggesting occasional large errors remain on that side.

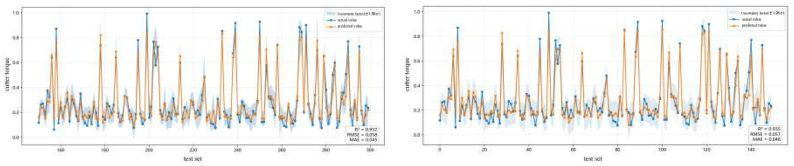
SVR (Support Vector Machine Regression): Although residuals cluster near zero, the horizontal range spans broadly (-1000 to 750) with high dispersion. Significant probability density persists far from zero, indicating substantial prediction bias and poor stability.

BO-RF (Bayesian Optimisation Random Forest): The curve is narrowest, with the highest peak and near-perfect symmetry. The majority of samples cluster around zero, indicating the highest accuracy, best stability, and virtually no systematic bias. Compared to the baseline model, outlier errors are markedly reduced.

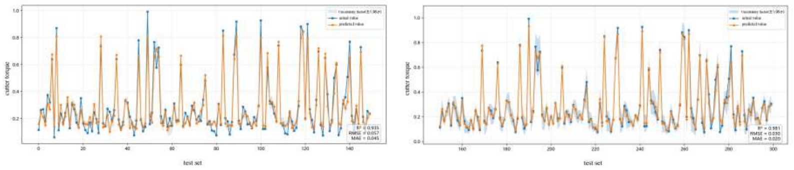
4.2.4 Comparison of Model Prediction Curves and Fitting Effect Analysis

Given the substantial dataset, segmented output is employed for each model's prediction results to comprehensively illustrate fitting performance.

As shown in Fig. 9, the RF model effectively tracks overall trends and most peaks/troughs, though it exhibits slight underfitting in steeply changing segments: extreme values are generally "suppressed", with occasional spikes showing lag or insufficient amplitude.



(a) RF Model (b) BR model



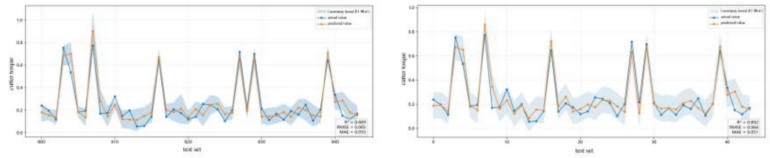
(c) SVR Model (d) BO-RF Model

Fig. 9 Model Results

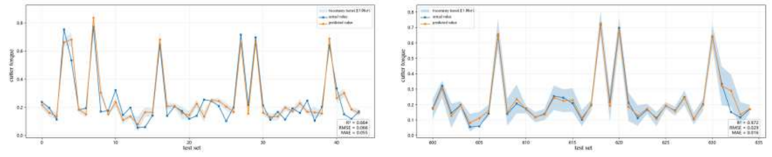
he BR model exhibits more pronounced contraction effects and demonstrates the weakest capability in capturing strong nonlinearities and spikes: peaks are significantly underestimated while troughs are elevated, resulting in systematic amplitude contraction. In regions of frequent fluctuations, greater deviations occur between predicted and actual values. Although its uncertainty coverage is generally reasonable, out-of-band points become more prevalent in areas of intense volatility.

The SVR model fits smoothly in steady segments, accurately capturing medium-to-low amplitude fluctuations. However, its performance deteriorates near extreme points, with some peaks and troughs markedly weakened or displaced. This leads to concentrated residual dispersion and insufficient stability in high-frequency and strongly fluctuating regions.

By contrast, the BO-RF model exhibits the highest degree of alignment between its forecast curve and actual values, demonstrating the most precise capture of extremes and minimal phase lag. Error fluctuations within the same interval are also reduced. Its uncertainty band is narrower with superior coverage, indicating robust stability and generalisability.



(a) RF Model (b) BR model



(c) SVR Model (d) BO-RF Model

Fig. 10 Segmented forecast map

As shown in Fig. 10, the comparison of the four models within sub-intervals further corroborates the overall conclusion: RF maintains fundamental stability but exhibits slight conservatism during certain periods of sharp volatility; the BR model demonstrates inadequate performance in handling non-linear fluctuations, resulting in relatively weaker fitting outcomes;The SVR model effectively captures trends and moderate-to-low amplitude fluctuations across most intervals, though it exhibits deviations in depicting extreme points, with relatively larger errors in localised fluctuation segments and inferior stability compared to RF. The BO-RF model demonstrates optimal performance in accuracy, extreme point capture, and uncertainty control.

4.2.5 Quantitative Metric Comparison of Different Models

To further quantify the performance comparison of different models, this paper calculates the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) for the BO-RF model, BR model, RF model, and SVR model. The results are shown in Table 3:

Table 3. Evaluation Metrics for Each Model

Model	R^2	RMSE	MAE
BO-RF Model	0.9794	0.0291	0.0190
BR model	0.9258	0.0557	0.0440
RF model	0.9213	0.0573	0.0442
SVR model	0.9229	0.0567	0.0450

It is evident that different models exhibit significant variations in accuracy when predicting cutter torque. The BO-RF model demonstrates the highest level of fit and the smallest error across all three metrics compared to other models, exhibiting superior stability and generalisation capability.In contrast, the BR model exhibits acceptable overall explanatory power but demonstrates an "amplitude contraction effect", leading to increased overall error. The RF model shows relatively high RMSE and MAE values; while maintaining a degree of stability, it exhibits significant shortcomings in capturing extreme values. The SVR model's R^2 is slightly higher than that of RF, but its RMSE and MAE are greater.

Overall, all four models adequately reflect the variation patterns of cutter disc torque to varying degrees. However, the BO-RF model demonstrates superior performance across all metrics, achieving the highest predictive accuracy.

5. Conclusion

Comparative analysis indicates that while the RF model demonstrates relative stability in fitting overall trends, it struggles to capture fluctuations within extreme value intervals. The BR model exhibits greater deficiencies in accuracy and stability when handling complex non-linear data. The SVR model performs reasonably well in predicting medium-to-low amplitude fluctuations during stable periods but shows significant deviation in fitting extreme points.In contrast, the proposed BO-RF model effectively mitigates reliance on hyperparameter tuning. It maintains robust overall trend fitting while more precisely capturing high-low peaks in cutter torque, significantly narrowing the range of prediction uncertainty.

The findings demonstrate that the BO-RF model outperforms RF, SVR, and BR in overall prediction accuracy, exhibiting superior robustness and generalisation capabilities across diverse operating conditions. By employing Bayesian optimisation, it overcomes the limitation of traditional random forests relying on empirically set hyperparameters, enabling the model to adapt more effectively to variable conditions. This methodology provides a reliable reference for setting TBM excavation parameters, holding significant implications for construction safety and optimising control strategies.

References

[1] Chuanqi L ,Jian Z ,Ming T , et al.Developing hybrid ELM-ALO, ELM-LSO and ELM-SOA models for predicting advance rate of TBM[J].Transportation Geotechnics,2022,36

[2] Mounib K ,Shervin S .Applied AI in instrumentation and measurement: The deep learning revolution[J].IEEE INSTRUMENTATION & MEASUREMENT MAGAZINE,2020,23(6):10-17.

[3] ZHANG Qinglong, ZHU Yanwen, MA Rui, et al.Study on Prediction of TBM Tunnelling Parameters Based on Attention-en-hanced Bi-LSTM Model[J].Modern Tunnelling Technology, 2022, 59(4): 69-80.

[4] Salimi A, Rostami J, Moormann C, et al. Application of non-linear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs[J]. Tunnelling and Underground Space Technology, 2016, 58: 236-246.

[5] Zhang Yan, Wang Wei, Deng Xueqin. Prediction Model of TBM Advance Rate Based on Relevance Vector Machine[J]. Modern Tunnel Technology, 2020, 57(3): 108-114.

[6] Zhu Mengqi, Zhu Hehua, Wang Xin, et al. Study on CART-based ensemble learning algorithms for predicting TBMtunneling parameters and classing surrounding rockmasses[J]. Chinese Journal of Rock Mechanics and Engineering , 2020, 39(9): 1860-1871.

[7] Kong X, Ling X, Tang L, et al. Random forest-based predictors for driving forces of earth pressure balance (EPB) shield tunnel boring machine (TBM)[J]. Tunnelling and underground space technology, 2022, 122: 104373

[8] Yu H, Qin C, Tao J, et al. A multi-channel decoupled deep neural network for tunnel boring machine torque and thrust prediction[J]. Tunnelling and Underground Space Technology, 2023, 133: 104949.

[9] Yin Zhiqing, Zhai Weijun, Han Aimin, et al. Predictive Research of Cutterhead Torque for Metro ShieldBoring Machine Based on Symbolic Regression Algorithm[J]. Urban Mass Transit, 2021, 24(01):127-131.DOI:10.16037/j.1007-869x.2021.01.027.

[10] Jia Haoyang, Qian Yu. Diagnosis on Transformer Fault Based on Bayesian Optimization XGBoost Algorithm[J]. Journal of Yellow River Conservancy Technical Institute, 2023, 35(02): 37-43. DOI:10.13681/j.cnki.cn41-1282/tv.2023.02.008.

[11] Akiba T, Sano S, Yanase T, et al. Optuna: A next-generation hyperparameter optimization framework[C]//Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019: 2623-2631.