OPEN ACCESS

Artificial Intelligence COMMENTARY

Machine learning–based TBM parameter prediction: Methodological reflections from an engineering perspective

Xianzhou Lyu,[1,2,*] Haoran Wu,[1] Wei Lu,[3,*] Lucia Landaverde Robles,[2] Yayong Li,[4] Yaran Wu,[1] Xukui Fan,[1] Bincheng Liang,[3] Baocai Zheng,[5] and Bole Sun[5]

[1]College of Earth Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China
[2]Department of Structural Engineering, University of California, San Diego, La Jolla 92093, USA
[3]College of Engineering, Ocean University of China, Qingdao 266100, China
[4]School of Mathematics and Physics, The University of Queensland, St Lucia 4072, Australia
[5]China Railway No.3 Engineering Group Co., Ltd., Taiyuan 030006, China
*Correspondence: lyuxianzhou0608@sdust.edu.cn (X.L.); luwei@ouc.edu.cn (W.L.)

Received: January 4, 2026; Accepted: January 29, 2026; Published Online: January 30, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100019

Citation: Lyu X., Wu H., Lu W., et al. (2026). Machine learning–based TBM parameter prediction: Methodological reflections from an engineering perspective. *The Innovation Informatics* 2:100019.

Machine learning (ML) has been increasingly adopted for predicting tunnel boring machine (TBM) construction parameters such as thrust force, penetration rate, and cutterhead torque.[1] With the rapid accumulation of high-frequency monitoring data from modern TBM projects, data-driven models have demonstrated considerable potential in capturing complex nonlinear relationships between geological conditions, operational variables, and machine responses.[2] A wide range of ML techniques, including tree-based ensembles, support vector machines, and deep neural networks, have been reported to achieve encouraging prediction accuracy in individual case studies.

At the same time, the rapidly growing body of literature reveals substantial variability in reported prediction performance and modeling practices.[3] Models that perform well in one project are often difficult to reproduce or generalize in others, even when similar algorithms are employed. Comparisons across studies are further complicated by differences in geological conditions, data characteristics, feature selection strategies, and evaluation protocols. These observations suggest that the challenges faced by ML-based TBM parameter prediction extend beyond algorithm selection alone. They are closely related to how prediction tasks are formulated, evaluated, and interpreted in an engineering context.

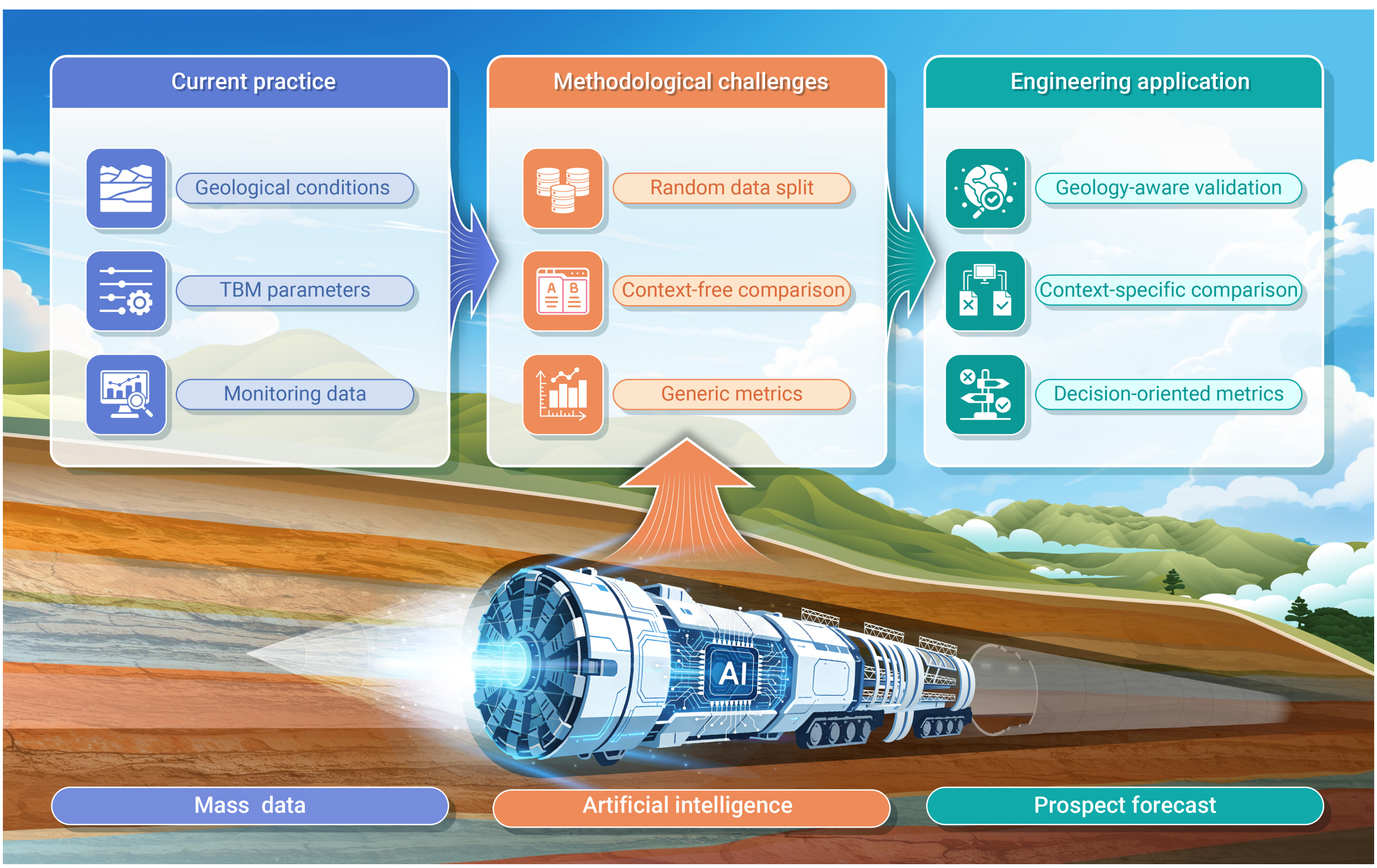

Figure 1. Conceptual framework for machine-learning-based TBM parameter prediction.

This commentary reflects on recent ML-based TBM studies by examining recurring methodological issues that may limit the interpretability, comparability, and engineering relevance of reported prediction results. By highlighting these issues, this commentary aims to contribute to a more informed and cautious use of ML techniques in mechanized tunneling applications. To clarify these issues, Figure 1 provides an overview of the data sources, modeling workflow, and key methodological challenges commonly encountered in machine-learning-based TBM parameter prediction, highlighting the importance of context-aware interpretation.

TBM PARAMETERS AS CONTEXT-DEPENDENT ENGINEERING QUANTITIES

In many ML applications, TBM parameters are treated as generic regression targets. However, from an engineering perspective, quantities such as thrust force or penetration rate are inherently context-dependent. They reflect the combined influence of geological conditions, operational decisions, and machine state, rather than representing intrinsic or standalone properties of the TBM system. For instance, in many single-project studies where thrust force is predicted using continuous monitoring data, the learned relationships often reflect a combination of geological conditions and operator-dependent control strategies, rather than a unique response to ground properties. As a result, statistically accurate predictions may still offer limited engineering insight if contextual factors are not explicitly considered.

Several recent studies implicitly acknowledge this conditional nature by incorporating geological and operational information as model inputs. Nevertheless, many studies differ markedly in how explicitly prediction targets and their applicable conditions are defined. In particular, the intended scope of model applicability is often only implicitly described, complicating the engineering interpretation of reported prediction accuracy. These observations highlight the importance of clearly framing TBM parameter prediction as a conditional engineering problem. Without explicit consideration of context, prediction results risk being interpreted as universally applicable, despite being derived from highly site-specific data.

OVER-OPTIMISTIC PERFORMANCE RISK AND EVALUATION PRACTICES

A commonly observed practice in ML-based TBM studies is the use of random train–test splitting for model evaluation. While this approach is straightforward and widely used in generic ML applications, it may produce overly optimistic performance estimates when applied to tunneling data, which are often strongly correlated in space and time.

Operational data collected along a tunnel alignment typically exhibit gradual transitions within geological units and abrupt changes at stratigraphic boundaries. This situation is particularly common in studies where datasets are constructed from long, continuous tunnel drives and randomly partitioned without regard to geological boundaries or segment transitions. As a result, reported prediction accuracy may primarily reflect interpolation within a limited range of conditions rather than robust performance under new or changing geological scenarios.

Several recent studies have explored alternative validation strategies, such as segment-based splitting along the tunnel axis or geology-aware partitioning, in which entire geological sections are excluded from the training process.[4] In practice, this involves defining training and testing datasets at the level of geological units or stratigraphic segments, thereby approximating deployment scenarios under previously unseen ground conditions. Such findings suggest that evaluation design plays a critical role in interpreting model accuracy and should be reported transparently alongside performance metrics.

From an engineering standpoint, these observations indicate that validation strategy can exert a greater influence on reported results than the choice of ML algorithm itself. Consequently, algorithm benchmarking without careful consideration of evaluation methodology may produce misleading conclusions regarding model superiority.

EVIDENCE SYNTHESIS DIFFICULTIES IN ML-BASED TBM STUDIES

With the rapid growth of machine-learning applications in TBM parameter prediction, researchers have increasingly attempted to synthesize findings across published studies to obtain transferable insights. In practice, however, directly aggregating performance metrics such as root-mean-square error (RMSE) or R^2 across studies often provides limited interpretive value.[5]

This limitation stems from the pronounced heterogeneity of existing work. Prediction targets differ, input features are defined and preprocessed in project-specific ways, and evaluation strategies are closely tied to individual datasets. Under such conditions, conventional meta-analysis techniques, typically developed for controlled experimental settings, are difficult to apply without introducing strong assumptions.

Rather than pursuing absolute performance rankings, several recent reviews have implicitly adopted a stratified perspective by grouping studies according to prediction objectives, data availability, or geological and operational complexity. Within these groups, emphasis is placed on performance variability and recurring methodological patterns, instead of identifying a universally optimal algorithm.

ENGINEERING APPLICATION IMPLICATIONS

In engineering practice, robustness, interpretability, and consistency across conditions are often more important than optimal performance on a single dataset. Accordingly, ML models that exhibit stable and interpretable behavior may be preferred even if their aggregate prediction accuracy is marginally lower. Recent studies have therefore increasingly employed interpretability tools, such as feature importance analysis and SHAP values, to assess whether learned relationships are physically plausible and engineering-relevant.

Looking ahead, while large-scale foundation models, large language models, and agent-based frameworks have shown promise in other engineering domains, their application to TBM parameter prediction remains limited and is currently oriented toward decision support and knowledge integration rather than direct parameter prediction.

CONCLUSION

The observations presented here do not diminish the potential of machine learning in mechanized tunneling, but rather emphasize the importance of aligning data-driven modeling practices with engineering realities. Based on the above discussion, several methodological considerations may help improve the practical relevance of future ML-based TBM parameter prediction studies. Prediction tasks should be explicitly defined together with their applicable geological and operational contexts. Validation strategies should reflect realistic deployment scenarios, for example through segment-based or geology-aware data partitioning. Model evaluation should emphasize robustness, stability, and interpretability across varying conditions rather than relying solely on aggregated accuracy metrics.

REFERENCES

1. Huang G., Qin C., Wang H., et al. (2025). TBM rock fragmentation classification using an adaptive spot denoising and contour-texture decomposition attention-based method. *Tunn. Undergr. Space Technol.* **161**:106498. DOI:10.1016/j.tust.2025.106498
2. Erharter G.H., Unterlass P., Radončić N., et al. (2025). Challenges and opportunities of data-driven advance classification for hard rock TBM excavations. *Rock Mech. Rock Eng.* **58**:1–14. DOI:10.1007/s00603-025-04542-4
3. Yang M.H., Song M.Y., Yao G.Z., et al. (2024). Research progress and prospects of machine learning for predicting ground surface settlement induced by shield tunneling. *Tunnel Constr.* **44**:2119–2132. DOI:10.3973/j.issn.2096-4498.2024.11.002
4. Man K., Wu L., Liu X., et al. (2024). Prediction of rock mass classification in tunnel boring machine tunneling using a PCA–GRU neural network. *Deep Undergr. Sci. Eng.* **3**:413–425. DOI:10.1002/dug2.12084
5. Liu J., Liu M., Liu J.P., et al. (2024). Towards provably efficient quantum algorithms for large-scale machine-learning models. *Nat. Commun.* **15**:434. DOI:10.1038/s41467-023-43957-x

FUNDING AND ACKNOWLEDGMENTS

L.X.C. and L.W. were supported by the National Natural Science Foundation of China [42502273;52309132], L.W. were supported by Taishan Scholars Program for Young Experts of Shandong Province, China [NO.tsqnz20240822]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Xianzhou Lyu: Conceptualization, Methodology, Funding acquisition, Writing-review & editing. Haoran Wu: Investigation, Data curation, Writing – original draft. Lucia Landaverde Robles: Methodology, Writing-review & editing. Wei Lu: Conceptualization, Funding acquisition, Writing-review & editing. Yayong Li: Investigation, Writing-review & editing. Yaran Wu: Data curation, Writing – original draft. Xukui Fan: Data curation, Investigation, Writing–original draft. Bincheng Liang: Methodology, Supervision. Baocai Zheng: Resources, Bole Sun: Resources. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.