UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Optimisation of Tunnel Boring Machine Performance by Machine Learning

by Feng Shan

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Professor Daichao Sheng and Dr Xuzhen He

University of Technology Sydney
Faculty of Engineering and Information Technology

May 2024

Certificate of Original Authorship

I, Feng Shan, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 30/05/2024

Acknowledgements

Firstly, I would like to express my gratitude to my supervisors Distinguished Professor Daichao Sheng and Dr Xuzhen He. I am specifically grateful to my supervisors for providing me with the opportunity to do research at the University of Technology Sydney and guiding me with great patience, encouragement and a fund of knowledge and enthusiasm. Without their consistent support throughout my period of candidature it would not have been possible for me to conduct and finish this research.

I would like to acknowledge the University of Technology Sydney International Research Scholarship for tuition fees and the China Scholarship Council Scholarship for living expenses.

I would like to thank Dr Danial Jahed Armaghani, Professor Jidong Teng, and Professor Sheng Zhang for their invaluable support throughout machine learning modelling and numerical simulation efforts. I could not have completed this research successfully without their generous and unselfish assistance.

Living and studying in Australia is a wonderful experience. I would also like to thank the friends along the way of my studies: Jinwen Bai, Miao Ge, Chenxi Tong, Jinbiao Wu, Xinyu Ye, Yuting Zhang, and other friends from the University of Newcastle; Dongle Cheng, Shaoheng Dai, Zhizhong Deng, Chaoqun Feng, Feng Gao, Qiang Hao, Wei Huang,

Yang Jiang, Ruxia Liang, Jibin Li, Huan Liu, Hang Luong, Jie Ma, Shudi Mao, Haimin Qian, Fulin Qu, Xingdong Shi, Yihan Shi, Ye Su, Chen Wang, Li Wang, Lijuan Wang, Haoding Xu, Zhengheng Xu, Bing Zhang, Jiaqi Zhang, Leijie Zhang, Zehao Zhang, Dong Zhao, Ting Zhou, and other friends from the University of Technology Sydney.

Last but not least, a particular thank you to my parents and sibling for their support throughout my life and studies. I am grateful to my parents for their love, support, and motivation in all of the things I do.

Publications

Journal papers related to the theme of this thesis:

- **Shan, F.**, He, X., Armaghani, D. J., Zhang, P. & Sheng, D. (2022) Success and Challenges in Predicting TBM Penetration Rate using Recurrent Neural Networks, *Tunnelling and Underground Space Technology* **130**:104728.

- **Shan, F.**, He, X., Armaghani, D. J. & Sheng, D. (2024) Effects of data smoothing and recurrent neural network (RNN) algorithms for real-time forecasting of tunnel boring machine (TBM) performance, *Journal of Rock Mechanics and Geotechnical Engineering* **16(5)**:1538-1551.

- **Shan, F.**, He, X., Armaghani, D. J., Zhang, P. & Sheng, D. (2023) Response to Discussion on "Success and Challenges in Predicting TBM Penetration Rate using Recurrent Neural Networks" by Georg H. Erharter, Thomas Marcher, *Tunnelling and Underground Space Technology* **139**:105064.

- **Shan, F.**, He, X., Xu, H., Armaghani, D. J. & Sheng, D. (2023) Applications of Machine Learning in Mechanised Tunnel Construction: A Systematic Review, *Eng—Advances in Engineering* **4(2)**:1516-1535.

- Xu, H., He, X., **Shan, F.**, Niu, G. & Sheng, D. (2023) 3D Simulation of Debris Flows with the Coupled Eulerian–Lagrangian Method and an Investigation of the Runout, *Mathematics* **11(16)**:3493.

- Xu, H., He, X., **Shan, F.**, Niu, G. & Sheng, D. (2023) Machine Learning in the Stochastic Analysis of Slope Stability: A State-of-the-Art Review, *Modelling* **4(4)**:426-453.

- **Shan, F.**, He, X., Armaghani, D. J., Xu, H., Liu, X. & Sheng, D. (2024) Real-time Forecasting of TBM Cutterhead Torque and Thrust Force Using Aware-context Recurrent Neural Networks, *Tunnelling and Underground Space Technology* (under review)

- **Shan, F.**, He, X., Armaghani, D. J., Xu, H., Liu, X. & Sheng, D. (2024) Semi-supervised Learning for Rock Mass Classification with Imbalanced Data between TBM Operations and Geotechnical Investigations, *Tunnelling and Underground Space Technology* (under review)

Conference paper related to the theme of this thesis:

- **Shan, F.** (2023) Forecasting TBM Performance Using Univariate and Multivariate Models by Deep Learning, *Proceedings of the 18th Conference of the Associated Research Centers for the Urban Underground Space*

I also researched on a very different topic during my first 2 years of PhD candidature, which resulted in the following publications:

- Teng, J., **Shan, F.**, He, Z., Zhang, S., Zhao, G. & Sheng, D. (2019) Experimental Study of Ice Accumulation in Unsaturated Clean Sand, *Géotechnique* **69(3)**:251-259.

- Liang, S., Teng, J., **Shan, F.** & Zhang, S. (2020) A Numerical Model of Vapour Transfer and Phase Change in Unsaturated Freezing Soils, *Advances in Civil Engineering* **2020**.

- Teng, J., **Shan, F.**, He, Z., Zhang, S. & Sheng, D. (2019) Numerical Modelling of Vapour-Ice Desublimation Process in Unsaturated Freezing Soils, *Proceedings of the 8th International Congress on Environmental Geotechnics* Springer, pp. 560-568.

- **Shan, F.**, Teng, J., Yan, X., Zhang, S. & Sheng, D. (2021) Numerical Modeling of Water-Vapour Migration and Phase Transformation in Unsaturated Freezing Soils, *Proceedings of the 16th International Conference of the International Association for Computer Methods and Advances in Geomechanics* Springer, pp. 925-931.

Table of Contents

List of Tables

List of Figures

List of Notations

All variables used in this thesis are defined as they are introduced into the text. For convenience, frequently used variables are described below.

Symbols

a_{10}	Percentage of samples within ±10% deviation
b	Argument of bias
C	Regularisation parameter
c	Cohesion
D_{TBM}	TBM diameter
E	Compressive modulus
F_n	Normal force
F_r	Rolling force
F_t	Total resultant force
F	Average cutter load
J_a	Joint alteration number
J_c	Joint condition
J_n	Joint set number
J_r	Joint roughness number
J_s	Joint spacing
J_w	Joint water parameter
K_1	Groundwater corrective coefficient
K_2	Structural plane occurrence corrective coefficient
K_3	Initial in-situ stress corrective coefficient
K_v	Integrity index
k_{DRI}	Correlation factor for DRI
k_d	Correlation factor for cutter diameter
k_{ekv}	Equivalent fracturing factor
k_p	Coefficient factor for penetration
k_{por}	Correlation factor for porosity
k_s	Rock mass fracturing factor
k_{spa}	Correlation factor for average cutter spacing

k	Sliding window size
M_1	Critical thrust to reach a penetration of 1 mm per revolution
M_{ekv}	Equivalent cutter thrust factor
M_{ga}	Gross average cutter thrust
N_c	Number of cutters
n_{10}	Number of samples within ±10% deviation
P^0	Pressure of the crushed zone
qz	Quartz content
R_c	Cutter radius
S_{Ti}	Total-effect Sobol index
S_c	Cutting spacing
S_i	First-order Sobol index
s_n	Smoothed data point
T	Cutter tip width
V	Linear velocity limit of the cutter
$\boldsymbol{W}$	Argument of the weight matrix
x'	Generated sample
x_{max}	Maximum in the data
x_{min}	Minimum in the data
x_n	Original data point
x_{nor}	Normalised data
$\boldsymbol{X}$	Input vector
$\hat{y}_i$	Predicted values
$\bar{y}$	Mean of actual values
y_i	Actual values
$\boldsymbol{Y}$	Output vector
α	Angle between fractures and tunnel axis
γ	Variance of radial basis function kernel
θ	Water content
λ	Smoothing factor
ξ	Degree of misclassification
σ_c	Uniaxial compressive strength of intact rock
σ_t	Tensile strength of intact rock
σ_θ	Average biaxial stress

φ	Internal friction angle
χ	Random value between 0 and 1
ψ	Constant for the pressure distribution function
ϕ	Angle of contact between rock and cutter
$\mathrm{ReLU}(x)$	Rectified linear unit
$\tanh(x)$	Hyperbolic function
$\mathrm{Var}(x)$	Variance
$K(x_i, x_j)$	Kernel functions
$f(x)$	Machine learning algorithm
$\mathbb{E}(x)$	Expected value
$\sigma(x)$	Sigmoid function
$\|W\|$	Euclidean norm of the weight vector
$\circ$	Element-wise product

Abbreviations

AdaBoost	Adaptive boosting
Adam	Adaptive moment estimation
AI	Artificial intelligence
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average
BBO	Biogeography-based optimisation
BI	Brittleness index
BO	Bayesian optimisation
BTS	Brazilian tensile strength
CART	Classification and regression tree
CatBoost	Categorical boosting
CD	Cover depth
CDAE	Constrained dense convolutional autoencoder
CFF	Core facture frequency
CLI	Cutter life index
CNN	Convolutional neural network

CSM	Colorado school of mines
CUSUM	Cumulative sum
DBN	Deep belief network
DBSCAN	Density-based spatial clustering of applications with the noise
DL	Deep learning
DNN	Deep neural network
DPW	Distance between planes of weakness
DRI	Drilling rate index
DT	Decision tree
DWT	Discrete wavelet transform
EMA	Exponential moving average
EPB	Earth pressure balance
F1	F1-score
FA	Firefly algorithm
FN	False negative
FP	Face pressure/earth pressure
FP	False positive
FPI	Field penetration index
FZ	Fault zone
GBDT	Gradient boosting decision tree
GBM	Gradient boosting machine
GC	Ground condition
GCN	Graph convolutional network
GEO	Geological parameters
GP	Genetic programming
GRU	Gated recurrent unit
GWO	Grey wolf optimiser
HC	Hydropower classification
HDA	Horizontal deviation of the shield articulation
HDH	Horizontal deviation of the shield head
HDT	Horizontal deviation of the shield tail
ICA	Imperialism competitive algorithm
IForest	Isolation forest
IQR	Interquartile range

KNN	k-nearest neighbour
LSTM	Long short-term memory
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MFO	Moth flame optimisation
ML	Machine learning
MSE	Mean squared error
MVO	Multi-verse optimisation
NTNU	Norwegian University of Science and Technology
OP	Operational parameters
PCA	Principal component analysis
PCC	Pearson correlation coefficient
Pev	Penetration per revolution
PLSI	Point loading strength index
POW	Cutterhead power
PR	Penetration rate
PR_20	Penetration rate by SMA within a sliding window of 20
PR_set	Setting values of penetration rate
PR_{t+1}	Penetration rate in the next step
$\{PR_t\}$	Time series of penetration rate
PRC	Precision
PSI	Punch slope index
PSO	Particle swarm optimisation
Q1	First quartile
Q3	Third quartile
Q_{TBM}	Q system
R^2	Coefficient of determination
REC	Recall
RF	Random forest
RMR	Rock mass rating
RMS	Rock mass strength
RMSE	Root mean squared error
RNN	Recurrent neural network
RPM	Revolutions per minute

RPM_set	Setting values of revolutions per minute
RQD	Rock quality designation
RQD_0	Rock quality designation in the tunnel axis
RSR	Rock structure rating
SCA	Sine cosine algorithm
SCS	Screw conveyor speed
SE	Specific energy
SMA	Simple moving average
SMBO	Sequential model-based optimisation
SMOTE	Synthetic minority over-sampling technique
SNN	Shared nearest neighbour
SPT	Standard penetration test
SRF	Stress reduction factor
SSAE	Stacked sparse autoencoder
SSO	Social spider optimisation
SST	Stacked single-target
SVM	Support vector machine
TBM	Tunnel boring machine
TH	Thrust force
TN	True negative
TO	Cutterhead torque
TP	True positive
TPI	Torque penetration index
UCS	Uniaxial compressive strength
VAF	Variance account for
VDA	Vertical deviation of the shield articulation
VDH	Vertical deviation of the shield head
VDT	Vertical deviation of the shield tail
VIM	Variable importance measure
WT	Water table
WZ	Weathering zone
XGBoost	Extreme gradient boosting

Abstract

Tunnel boring machines (TBMs) have been widely utilised in tunnel construction due to their efficiency and reliability. However, tunnel collapse, rock bursting, water inrush, and machine jamming remain severe challenges in complex geotechnical conditions. The main objective of this thesis is to investigate if and to what extent TBM operation can be forecasted in real time by machine learning, and if such forecasts can result in TBM tunnelling optimisation, dissecting it into four aspects: (1) penetration rate regression, (2) penetration rate forecasting, (3) cutterhead torque and thrust force forecasting, and (4) rock mass classification.

Penetration rate regression is to explore the relationship between penetration rate and other related parameters. Traditional theoretical and empirical methods often provide less accurate results, suited for overall project time management before the start of a project. Machine learning models, using data from the Pahang-Selangor water tunnel, demonstrate better results with support vector machine (SVM), random forest (RF), and artificial neural network (ANN). While other models inclusive of operational and geological parameters can have high accuracy, their applicability is questioned because operational parameters are not accessible during training.

Real-time forecasting of penetration rates leverages historical data to predict unknown penetration rates in the future. Accurate forecasts, even for a short distance ahead, can aid

in refining TBM operations, resource allocation, decision-making process, and early warning of unexpected geotechnical conditions. We build recurrent neural network (RNN) and long short-term memory (LSTM) models to predict short-term and long-term penetration rates. These models are trained on data from Changsha but tested on data from Zhengzhou. In addition, a random walk theory is used to explain time lags. It is found that data smoothing enhances accuracy at the potential cost of losing original characteristics.

Centring on real-time forecasting of cutterhead torque and thrust force, the thesis underscores their significance in optimising performance and ensuring operational safety. By integrating the historical operational parameters and current setting values, the aware-context RNN model outperforms other RNN models, as reflected in the low-frequency data from the Yinsong water diversion project.

Lastly, we accentuate the need for accurate rock mass classification, avoiding inappropriate operation and low safety of excavation. Utilising supervised and semi-supervised learning, our results emphasise the superiority of the RF-based self-training classifier over other RF classifiers, especially when leveraging unlabelled data.

Overall, the findings serve as a comprehensive guide for TBM professionals, aiming to streamline and enhance TBM operations.

Keywords: Tunnel boring machine, machine learning, TBM performance regression, time series forecasting, rock mass classification, penetration rate, cutterhead torque, thrust force, random forest, recurrent neural network

Chapter 1. Introduction

1.1 Background

In academic terms, a tunnel is defined as a subterranean passage or conduit constructed under the ground, through a hill, mountain, or under a body of water. The Tunnel of Eupalinos, one of the oldest tunnels, was constructed in Greece in the 6th century BC for transporting water. The onset of the Industrial Revolution in the late 18th century brought about a significant increase in tunnel construction, along with the demand for mining, defensive fortifications, and transportation. Drilling and blasting, a traditional technique, involves the use of explosives to break rock, resulting in unpredictable nature of blasting in different geological conditions. The first shield successfully completed the Thames Tunnel in London in 1843, marking the beginning of mechanised tunnelling. As technology continues to evolve, tunnel boring machines (TBMs) were developed and became prevalent because of higher efficiency, safer workplaces, minimal environmental disturbance, and reduced project costs (Rostami, 1997). TBMs, equipped with a circular rotating cutterhead, are adept at boring through a diverse range of geologies. Specifically, the earth pressure balance (EPB) shield is used for soft soil, the slurry TBM for soft ground with high water pressure, the hard rock TBM for solid rock, and the mixshield TBM for varying ground conditions. Their versatility makes them ideal for various underground projects, including subways, railways, water conveyance systems, gas

transmission pipelines, and underground mines. However, tunnel collapse, rock bursting, water inrush, and machine jamming remain major challenges when TBMs encounter complex geotechnical conditions. For example, 19 water inrush hazards occurred during the construction of the Maluqing Tunnel of Yiwan Railway in China, which disastrously caused 15 fatalities (Li et al., 2017). The Snowy 2.0 pumped storage project in Australia is located in complex alpine geological and hydrogeological conditions, which poses significant geotechnical uncertainties (Gomes et al., 2021), resulting in a sinkhole, toxic gas and a 2 billion dollar mistake.

Predicting TBM performance and rock mass classification are vital in increasing efficiency and avoiding potential risks. On the one hand, TBM performance refers to the effectiveness and efficiency of TBMs in excavating a tunnel and is reflected by various indicators: Penetration Rate (PR) is the speed of boring distance divided by the working time; Field Penetration Index (FPI) evaluates TBM efficiency in the field calculated as the average cutter force divided by penetration per revolution; Cutterhead Torque (TO) is the twisting force applied to the cutterhead; Thrust Force (TH) is the force that TBM exerts on the excavation face. On the other hand, rock mass classification is for understanding the properties and behaviours of rock masses. There are a few well-known rock mass classification systems: Rock Mass Rating (RMR) system uses parameters like uniaxial compressive strength, rock quality designation, joint conditions, water condition, and orientation of discontinuities to classify rock masses (Bieniawski, 1989); Q-system

(Q_{TBM}) considers parameters such as rock quality designation, joint set number, joint roughness, joint alteration, water inflow, and stress reduction factor (Barton et al., 1974); Hydropower Classification (HC), a specialised engineering geological classification system in China, is crucial for determining the feasibility, design, and construction method (Lin, 1999).

Traditionally, TBM operators rely primarily on empiricism based on site geology, operational parameters, and tunnel geometry. Theoretical and empirical methods are widely used in practice but are not accurate and are strictly limited by the TBM type and geological conditions. The abundance of data collected by the data acquisition system provides an opportunity for the application of machine learning in TBM tunnelling. Machine learning techniques are known for their high effectiveness and versatility in capturing complex, non-linear relationships. However, their applicability is limited due to their project-specific nature and lack of generalisation across different TBM types and geological conditions.

A comprehensive review of scholarly publications from the Web of Science Core Collection reveals a progressive growth in research on the integration of machine learning techniques and TBM tunnelling. The review used two query strings: ‘tunnel boring machine or earth pressure balance shield’ and ‘machine learning or deep learning or neural network’. As depicted in Figure 1.1, the body of literature, 254 papers in total,

experienced a rather muted interest before 2018. A marked rise in publications was observed from 2018 to 2022, as reported by Shan et al. (2023c). The rising number of published papers shows the growing interest in and recognition of the benefits of machine learning techniques in TBM tunnelling, such as support vector machine (SVM), random forest (RF), artificial neural network (ANN), recurrent neural network (RNN), and long short-term memory (LSTM). The integration of machine learning with TBM tunnelling has emerged as a promising approach to enhance the efficiency and safety of tunnel construction. However, this integration faces several challenges and gaps that need to be addressed.

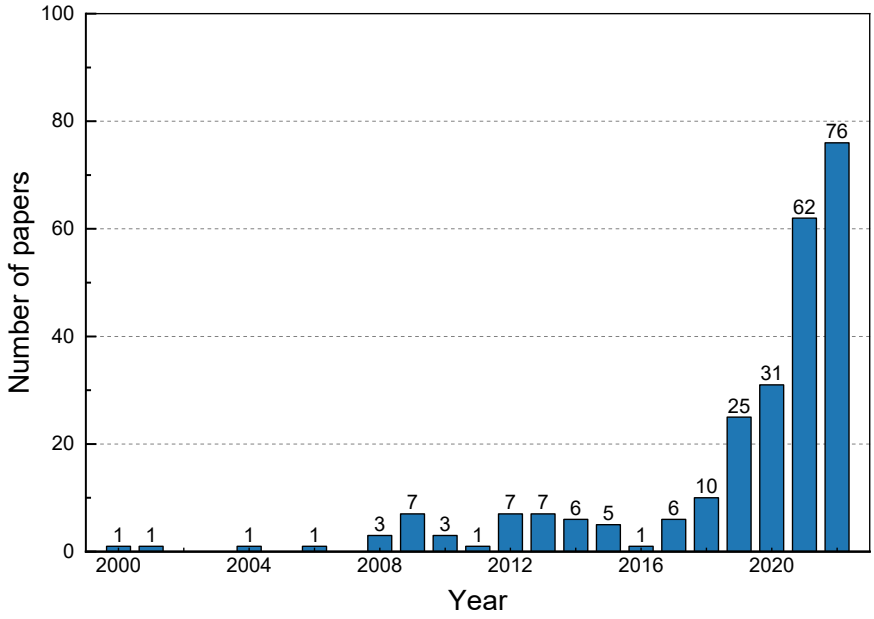


Figure 1.1 Number of papers using machine learning models in TBM tunnelling.

1.2 Knowledge gaps

Despite the growing interest in the integration of machine learning with TBM tunnelling, it is crucial to understand and address critical knowledge gaps.

1. Machine learning models, though powerful, are often criticised for their lack of clear reasoning. The black box problem makes it difficult for engineers and decision-makers to understand and trust the predictions and recommendations of models.

2. Model generalisability refers to the tendency of machine learning models to be too specialised for specific datasets. Models developed and trained on one set of tunnelling data may not perform well when applied to different datasets.

3. Operational data is frequently sourced from the data acquisition system every few seconds or minutes, while geological data is obtained from costly borehole drilling with limited samples spanning several metres, even more than 200 metres, leading to an imbalance between operational and geological data.

4. The field of TBM tunnelling generates vast amounts of data, but much of this data remains underutilised. Leveraging diverse data sources (geological, operational, and machine data) effectively for machine learning can lead to more robust and reliable models, but this requires sophisticated data processing and machine learning algorithms.

1.3 Research objectives

The primary objectives of this thesis are twofold. First, the thesis aims to answer the

question of whether TBM operational efficiency can be optimised through machine learning, thereby reducing time and costs. More specifically, the thesis aims to answer questions such as: Which operational parameters are able to control TBM operation? Is it possible to forecast future penetration rates in real time? How far into the future can we forecast? Regarding the practical questions, this thesis involves penetration rate regression, which is a causal relationship between TBM performance and other related variables, as well as penetration rate forecasting using historical and real-time data to streamline tunnelling processes.

Second, the thesis focuses on enhancing safety in TBM operations and addresses the following questions: What operational parameters are indicative of geological conditions? Can we detect potential risks ahead of the cutterhead by forecasting related operational parameters? Can we classify rock mass classification with accuracy? As a result, cutterhead torque and thrust force forecasting indirectly reflect geological conditions ahead of the cutterhead, while rock mass classifiers directly assess geological conditions to mitigate risks.

1.4 Scope of work

As a result, this thesis proposes a comprehensive work from four aspects:

1. Penetration rate regression in Chapter 5 builds a causal relationship using accessible geological data, including comparisons on machine learning algorithms

and input parameters;

2. Penetration rate forecasting in Chapter 6 builds univariate and multivariate models to predict the near future (0–7.5 m) penetration rate, which are generalised to different geological conditions using training data from Changsha and test data from Zhengzhou;

3. Cutterhead torque and thrust force forecasting in Chapter 7 predict future values in the next boring cycle, approximately 1.14 metres ahead of the cutterhead, integrating the historical operational parameters and aware context of setting values;

4. Rock mass classification in Chapter 8 addresses the imbalance of geological and operational data by semi-supervised learning approaches that generate confident pseudo-labels to enhance training processes.

Regarding the gap of black boxes, we utilise the Sobol method to make a sensitivity analysis in penetration rate forecasting (Chapter 6) and cutterhead torque and thrust force forecasting (Chapter 7), accounting for the main effects and all interaction effects of input parameters. Concerning model generalisability, penetration rate forecasting models in Chapter 6 are trained on training data from Changsha and evaluated on test data from Zhengzhou, where geological conditions are quite different. To address the imbalance between operational and geological data, semi-supervised learning is employed to make full use of the unlabelled geological data in Chapter 8. Although operational and geological data have been considered recently, robust and reliable models are still rather

limited, particularly in TBM types. It requires further studies when more datasets are available.

Concerning the first objectives of operational efficiency, penetration rate regression builds a causal relationship between TBM performance and other related variables in Chapter 5, while penetration rate forecasting uses historical and real-time data to optimise tunnelling processes in Chapter 6. With regard to the second objective, cutterhead torque and thrust force forecasting indirectly reflect geological conditions ahead of the cutterhead in Chapter 7, while rock mass classifiers directly assess geological conditions to mitigate risks in Chapter 8. This thesis has the potential to make a significant contribution to the field of TBM tunnelling, offering insights and tools that can improve current practices.

1.5 Thesis outline

The thesis primarily delves into optimising Tunnel Boring Machine (TBM) tunnelling to augment operational efficiency and safety, including nine Chapters. In Chapter 2, the literature review presents a comprehensive exploration of methods to predict TBM performance, encompassing theoretical, empirical, and machine learning approaches. Chapter 2 also sheds light on time series forecasting, focusing on TBM performance and positional aspects. The intricacies of rock mass classification and rock mass parameters are then examined.

The methodology Chapter 3 outlines the strategies and techniques used in the research. It begins with a deep dive into data processing methods, such as outlier detection and data smoothing, and subsequently transitions into a detailed exposition of various machine learning algorithms, from support vector machines to self-training methodologies. Evaluation metrics are highlighted to evaluate the performance of the applied machine learning algorithms.

Chapter 4 presents the datasets employed in the research, sourced from the Pahang-Selangor raw water tunnel, the Changsha and Zhengzhou metro lines, and the Yinsong water diversion project. Each dataset reviews the respective projects and provides a statistical data analysis.

Subsequent Chapters 5–8 are centred on specialised research aspects. Chapter 5 focuses on regression models for TBM penetration rates. Chapter 6 extends this topic, aiming at forecasting penetration rates through a combination of univariate and multivariate models. Chapter 7 delves into predicting two crucial parameters: cutterhead torque and thrust force, employing regression and time series forecasting models. Chapter 8 centres on rock mass classification, emphasising supervised and semi-supervised learning approaches.

Concluding the research, Chapter 9 encapsulates the findings and offers recommendations for future work and the practical application of the research in TBM tunnelling.

Chapter 2. Literature Review

2.1 TBM performance regression

TBM performance regression is to find the causal relationship between TBM performance and other related variables. The prediction models are normally categorised into theoretical, empirical, and machine learning models. The accuracy of theoretical or empirical models is acceptable but not sufficiently high to meet the demands for safe and efficient construction. The rapid development of machine learning models is attributed to the advancements in high-performance computing and the abundance of collected data. However, machine learning models are limited to the quality and quantity of the data and often suffer from a lack of interpretability.

2.1.1 Theoretical method

The theoretical methods developed force equilibrium equations based on mechanics principles and full-size linear cutting tests, considering intact rock properties, cutterhead load, and TBM geometry (Roxborough and Phillips, 1975; Ozdemir, 1977; Snowdon et al., 1982; Rostami and Ozdemir, 1993; Rostami, 1997; Yagiz, 2002). The force equilibrium equations are flexible with TBM design parameters, allowing design optimisation for cutterhead layout, machine specifications, and check cutterhead balancing.

The Colorado School of Mines (CSM) model, a widely accepted theoretical model, was developed using Robbins company data on granite, quartzite, schist, and shale. Full-size cutting tests simulate intact rock-cutting conditions using linear cutting machines. Rostami (1997) integrated TBM geometry (cutter spacing, cutter tip width, and cutter diameter) and intact rock properties to predict the cutter force for a given penetration, excluding the influence of joints. Subsequently, Yagiz (2002) updated the CSM model by adding the rock brittleness index (BI), distance between planes of weakness (DPW), and angle between fractures and tunnel axis (α) as inputs to the model. In detail, the original formula to estimate the total resultant force is as follows:

$$F_t = \frac{P^0 \phi R_c T}{1 + \psi} \tag{2.1}$$

$$\phi = \cos^{-1}\left(\frac{R_c - \text{Pev}}{R_c}\right) \tag{2.2}$$

$$P^0 = 2.12 * \sqrt[3]{\frac{S_c}{\Phi\sqrt{R_c W_{ct}}}\sigma_c{}^2 \sigma_t} \tag{2.3}$$

where F_t = Total resultant force

R_c = Cutter radius

T = Cutter tip width

ϕ = Angle of contact between rock and cutter

P^0 = Pressure of the crushed zone

ψ = Constant for the pressure distribution function, between -0.2 and 0.2

Pev = Penetration per revolution

S_c = Cutting spacing

σ_c = Uniaxial compressive strength of intact rock

σ_t = Tensile strength of intact rock

The pressure formula of the crushed zone was derived from the regression analysis of available data. The pressure formula in Eq. 2.3 is predicated on the assumption that the resultant force passes through the centre of the contact area. The normal and rolling forces are expressed in Eqs. 2.4 and 2.5.

$$F_n = F_t \cos\left(\frac{\phi}{2}\right) \quad 2.4$$

$$F_r = F_t \sin\left(\frac{\phi}{2}\right) \quad 2.5$$

where F_n = Normal force

F_r = Rolling force

As a result, the overall thrust force, cutterhead torque, and revolutions per minute are calculated as follows:

$$\mathrm{TH} = \sum_1^N F_n \approx N_c * F_n \quad 2.6$$

$$\mathrm{TO} = \sum_1^N F_r * R \approx 0.3 * D_{\mathrm{TBM}} * N_c * F_r \quad 2.7$$

$$\mathrm{RPM} = \frac{V}{\pi * D_{\mathrm{TBM}}} \quad 2.8$$

where TH = Overall thrust force

N_c = Number of cutters

TO = Overall cutterhead torque

D_{TBM} = TBM diameter

RPM = Revolutions per minute

V = Linear velocity limit of the cutter

While these theoretical models enhance a fundamental understanding of TBM cutting

mechanics, they do not accurately reflect the complexities of actual rock mass conditions in situ. In addition, only a few laboratories worldwide are equipped with linear cutting machines to perform linear cutting tests.

2.1.2 Empirical method

Empirical models are based on the historical field performance of TBM in certain rock types. Typically, the empirical models are a set of empirical equations obtained from the regression analysis correlating rock mass properties, operational parameters, and machine specification (Innaurato et al., 1991; Bruland, 1998; Barton, 2000; Yagiz, 2008; Gong and Zhao, 2009; Hassanpour et al., 2011). Though they naturally incorporate geological conditions in the field and TBM type, their predictive abilities are limited by the similarity of historical systems.

Bruland (1998) improved the Norwegian University of Science and Technology (NTNU) model to predict penetration rate, utilising field data collected from Norwegian tunnels. The NTNU prediction model for hard rock estimates the penetration rate by incorporating rock mass properties and machine parameters, as detailed in Table 2.1.

Table 2.1 Rock mass properties and machine parameters influencing penetration rates

Rock mass properties	**Machine parameters**
Drilling rate index	TBM diameter
Porosity	Cutter diameter
Rock mass fracturing factor	Number of cutters
	Cutter spacing
	Installed cutterhead power
	Gross average cutter thrust
	Revolutions per minute

The drilling rate index (DRI) represents the boreability of the intact rock based on the Brittleness Value test and the Sievers' J-Value miniature drill test. The rock porosity is associated with the penetration rate because the pores act as crack initiators and amplifiers of crack propagation. Rock mass fracturing is found to be the most important geological factor in penetration rate, evaluated by a rock mass fracturing factor. The factor for rock mass properties is equivalent to the fracturing factor in Eq. 2.9.

$$k_{ekv} = k_s * k_{DRI} * k_{por} \qquad 2.9$$

where k_{ekv} = Equivalent fracturing factor
k_s = Rock mass fracturing factor
k_{DRI} = Correlation factor for DRI
k_{por} = Correlation factor for porosity

Increased thrust enables the cutter edge to penetrate deeper into the rock surface and effectively convey the energy from the cutterhead to the rock. The gross average cutter thrust, the most important machine parameter, simplifies the thrust distribution over the

cutterhead. A detailed simulation of the thrust distribution would likely include parameters such as cutter diameter, average cutter spacing, and rock mass fracturing, expressed as the equivalent cutter thrust factor in Eq. 2.10.

$$M_{ekv} = M_{ga} * k_d * k_{spa} \quad 2.10$$

where M_{ekv} = Equivalent cutter thrust factor
M_{ga} = Gross average cutter thrust
k_d = Correlation factor for cutter diameter
k_{spa} = Correlation factor for average cutter spacing

The NTNU model relies on the normalised penetration curve in Eq. 2.11 (Bruland, 1998), offering an excellent fit to a wide range of penetration tests concerning variations in rock mass and machine parameters. Penetration rate equals penetration per revolution times revolutions per minute in Eq. 2.12.

$$\text{Pev} = \left(\frac{M_{ekv}}{M_1}\right)^{k_p} \quad 2.11$$

$$\text{PR} = \text{Pev} * \text{RPM} \quad 2.12$$

where M_1 = Critical thrust to reach a penetration of 1 mm per revolution
k_p = Penetration coefficient

Critical cutter thrust and penetration coefficient are related to the equivalent fracturing factor. However, the NTNU model requires specialised experiments in drilling, so these tests are seldom available outside Norway.

Additionally, it is observed that penetration rate is related to rock mass classification, such

as the rock structure rating (RSR) system (Innaurato et al., 1991), RMR system (Sapigni et al., 2002), and Q-system (Barton, 2000). The Q_{TBM} model, in particular, originates from the Q system and includes many parameters for practical application, including rock quality designation (RQD), joint condition, stress condition, intact rock strength, quartz content, and thrust. Yagiz (2008) conducted a statistical analysis on data and proposed an empirical model to predict penetration rates associated with uniaxial compressive strength (UCS), punch slope index (PSI), DPW, and α. Gong and Zhao (2009) performed a nonlinear regression analysis on data and developed an empirical equation to estimate the rock mass boreability. Hassanpour et al. (2011) suggested a boreability classification system based on four tunnel projects and proposed an empirical field penetration index (FPI) model using only RQD and UCS. A summary of empirical equations is presented in Table 2.2.

Table 2.2 Empirical models for TBM performance

Empirical equation	**Reference**
$\mathrm{PR} = {\sigma_c}^{-0.437} - 0.0437 * \mathrm{RSR} + 3.15$	Innaurato et al. (1991)
$\mathrm{PR} = 5 * {\mathrm{Q_{TBM}}}^{-0.2},\ \mathrm{Q_{TBM}} = \frac{\mathrm{RQD}}{J_n} * \frac{J_r}{J_a} * \frac{J_w}{\mathrm{SRF}} * \frac{20^9 \mathrm{RMS}}{F^{10}} * \frac{20}{\mathrm{CLI}} * \frac{qz}{20} * \frac{\sigma_\theta}{5}$	Barton (2000)
$\mathrm{PR} = 1.093 + 0.029 * \mathrm{PSI} - 0.003 * \mathrm{UCS} + 0.437 * \log(\alpha) - 0.219 * \mathrm{DPW}$	Yagiz (2008)
$\mathrm{FPI} = 9.401 + 0.397 * \log(\alpha) + 0.011 * \mathrm{Jc}^2 + 1.14 * 10^{-5} * \mathrm{RQD}^3 + 1.32 * 10^{-8} * \mathrm{UCS}^4$	Hamidi et al. (2010)
$\mathrm{FPI} = e^{0.008*\mathrm{UCS}+0.015*\mathrm{RQD}+1.384}$	Hassanpour et al. (2011)

where $\mathrm{RQD_0}$ = RQD in the tunnel axis

J_n = Joint set number

J_r = Joint roughness number

J_a = Joint alteration number

J_w = Joint water parameter

SRF = Stress reduction factor

RMS = Rock mass strength

F = Average cutter load

CLI = Cutter life index

qz = Quartz content

σ_θ = Average biaxial stress

J_c = Joint condition

2.1.3 Machine learning method

Machine learning techniques are highly effective and versatile in capturing complex, non-linear relationships and have been successfully applied to TBM tunnelling. Table 2.3 provides a summary of articles on TBM performance, involving input and output parameters, machine learning algorithms, tunnel projects (or datasets), and the number of used samples (or data size). Understanding and optimising TBM performance is crucial for project time management and cost control. In machine learning, predicting TBM performance is a function of input parameters in Eq. 2.13.

$$\boldsymbol{Y} = f(\boldsymbol{W}\boldsymbol{X}, b) \qquad 2.13$$

where $\boldsymbol{Y}$ = Output vectors

$\boldsymbol{X}$ = Input vectors or input parameters

$\boldsymbol{W}$ = Argument of the weight matrix

b = Argument of bias

$f(x)$ = Machine learning algorithm

Since machine learning models are data-driven, the quality of data (e.g., availability to the public, number of samples, and input parameters used) is crucial. Table 2.4 displays three typical datasets and their respective input parameters. Six hundred forty tunnel projects include geological conditions, operational parameters, and TBM specification; The Queens water tunnel only includes geological conditions; The Pahang-Selangor includes geological and operational parameters.

Table 2.3 Summary on TBM performance regression

Outputs	**Inputs**	**Algorithms**	**Data size and dataset**	**Reference**
PR, AR	CFF, UCS, RPM, TH/cutter, R_c	ANN, ANFIS	640 tunnel project	Grima et al. (2000)
PR	UCS, BI, DPW, α	ANN	151, Queens water	Yagiz (2009)
PR	UCS, RQD, DPW	ANN	185, three tunnel projects	Javad and Narges (2010)
PR	UCS, BTS, BI, DPW, α, SE, TO, TH, POW	SVM	150, Queens water	Mahdevari et al. (2014)
PR	UCS, BTS, RQD, RMR, WZ, TH, RPM	PSO-ANN, ICA-ANN, ANN	1286, Pahang-Selangor	Armaghani et al. (2017)
PR	UCS, BTS, RQD, RMR, WZ, TH, RPM	FA-ANN, ANN	1200, Pahang-Selangor	Koopialipoor et al. (2020)
PR	UCS, BTS, PSI, DPW, α	ANN, SVM, CART, RF, bagging tree, AdaBoost	151, Queens water	Zhang et al. (2020c)
PR	UCS, BTS, RQD, RMR, WZ, TH, RPM	PSO, GWO, SSO, SCA, MVO, MFO +XGBoost	1286, Pahang-Selangor	Zhou et al. (2021a)
PR	TH, FP, TO, RPM	SNN-DBSCAN, K-	904, Thomson	Fu et al.

		means, Ward, DBSCAN +RF, ANN, SVM	line	(2022)
PR	FP, TH, TO, RPM, ratio of boulder, RQD, UCS, c, φ, E	GWO-SVM, BBO-SVM, SVM	503, Shenzhen metro	Yang et al. (2022b)
PR, FPI	UCS, BI, DPW, α	RF, evolutionary polynomial regression	151, Queens water	Yang et al. (2022c)
AR	RQD, RMR, UCS, CD, WT, weathering degree, overload factor, permeability	ANN	11, Athens metro	Benardos and Kaliampakos (2004)
AR, TO, TH	position and thickness of layers, 53 operational data	RF	22232, Shenzhen metro	Sun et al. (2018)
AR	UCS, BTS, RQD, RMR, qz, WZ, TH, RPM	PSO-ANN, ICA-ANN, ANN	1286, Pahang-Selangor	Armaghani et al. (2019)
AR	53-40-10 operational data	BO-SVM	over 500000, Northgate link	Mokhtari and Mooney (2020)
AR	UCS, BTS, RQD, RMR, TH, RPM	PSO- extreme learning machine	1286, Pahang-Selangor	Zeng et al. (2021)
AR	UCS, BTS, RQD, RMR, WZ, TH, RPM	BO-XGBoost	1286, Pahang-Selangor	Zhou et al. (2021b)
FP, TO, TH	PR, SCS	GBM, DT, SVM	450, Xi'an metro	Bai et al. (2021)
TH	CD, GC, c, φ, E, θ, unit weight, FP, TO, RPM, PR, Pev, SCS, foam volume, foam pressure, grouting volume, grouting pressure	PSO-LSTM	1500, Shenzhen railway	Lin et al. (2022b)
TO	unit weight, porosity, φ, E, θ, TH, FP, PR, Pev, SCS, foam volume, grouting volume, grouting pressure	PSO-GRU	1000, Shenzhen railway	Lin et al. (2022a)
FPI	UCS, Js	SVM, ANFIS	75, Zagros lot	Salimi et al.

				(2016)
FPI	UCS, Js, RQD, Jc, rock type code	CART, GP	580, seven tunnel projects	Salimi et al. (2019)
FPI	UCS, RQD, volumetric joint count	CART	666, eight tunnel projects	Salimi et al. (2022)

CFF, core facture frequency; UCS, uniaxial compressive strength; BTS, Brazilian tensile strength; BI, brittleness index; PSI, punch slope index; DPW, distance between planes of weakness; α, angle between fractures and tunnel axis; RQD, rock quality designation; WZ, weathering zone; WT, water table; Js, joint spacing; Jc, joint condition; RMR, rock mass rating system; R_c, cutter diameter; CD, cover depth; c cohesion; φ, internal friction angle; E, compressive modulus; θ, water content; Pev, penetration per revolution; SE, specific energy; SCS, screw conveyor speed; POW, cutterhead power; GC, ground condition;

ANFIS, adaptive neuro-fuzzy inference system; GP, genetic programming; DT, decision tree; CART, classification and regression tree; GRU, gated recurrent unit; GBM, gradient boosting machine; AdaBoost, adaptive boosting; XGBoost, extreme gradient boosting; SNN, shared nearest neighbour; DBSCAN, density-based spatial clustering of applications with the noise;

PSO, particle swarm optimisation; GWO, grey wolf optimiser; ICA, imperialism competitive algorithm; FA, firefly algorithm; SCA, sine cosine algorithm; SSO, social spider optimisation; MVO, multi-verse optimisation; MFO, moth flame optimisation; BBO, biogeography-based optimisation; BO, Bayesian optimisation;

Table 2.4 Three datasets on input parameters and limitations

Dataset	Data size	Input parameters	Open access	Limitations
640 tunnel projects	-	Geological conditions, Operational parameters, TBM specification	No	no access
Queen water tunnel	151	Geological conditions	Yes	overfitting or lack of generalisation
Pahang-Selangor raw water transfer	1286	Geological conditions, Operational parameters	Yes	not applicable in practice

2.1.3.1 Machine learning for penetration rate

The penetration rate (PR) measures the speed of boring distance divided by the working

time, typically quantified in m/h or mm/min. The penetration rate is crucial in tunnelling operations as it directly affects overall productivity. A higher penetration rate results in faster tunnel excavation, ultimately reducing project time and costs. The ANIFS model (Grima et al., 2000) predicted penetration rate better than multiple regression and empirical methods based on a database of 640 TBM projects in rock. The ANIFS model is adaptable as it takes into account geological conditions CFF and UCS, operational parameters of thrust per cutter, and even cutter diameter. However, most TBM projects are not available for public access.

The ANN and SVM methods (Yagiz, 2009; Mahdevari et al., 2014) outperformed linear and non-linear regression when applied to the publicly available Queen water tunnel dataset with 151 samples. Zhang et al. (2020c) further proposed an AutoML method involving data preprocessing, selection of algorithms, and optimisation of related hyperparameters. The methods use geological parameters (UCS, BTS, BI, DPW, and α) from site investigation to predict penetration rates. In the sensitivity analysis, interestingly, the brittleness index was found to be the least effective parameter in the SVM model (Mahdevari et al., 2014) but the most sensitive parameter in the RF model (Yang et al., 2022c). These contrasting results can be attributed to a limited number of samples for training, which leads to overfitting or a lack of generalisation.

In the project of Pahang-Selangor raw water transfer with 1286 samples, machine learning

models for predicting penetration rate were robust and reliable because of more data and adding operational parameters of thrust force and revolutions per minute (Armaghani et al., 2017; Koopialipoor et al., 2020; Zhou et al., 2021a). However, TBM performance is a real-time operational parameter that cannot be obtained before the start of a project, making it infeasible to apply in practice. For example, although the average thrust force is an effective parameter for predicting penetration rate (Koopialipoor et al., 2020), it is an operational input that is unavailable as it is collected in real-time as well as the output of penetration rates.

In the Singapore Thomson Line project, four operational parameters (TH, TO, FP, and RPM) are assigned to four clusters using the unsupervised learning SNN-DBSCAN method (Mokhtari and Mooney, 2020). These four clusters correspond to four major soil types, including fluvial sand, fluvial sand mixed with marine clay, marine clay, and old alluvium. Based on these clusters, four differentiated RF models are further built to predict the penetration rate, demonstrating greater accuracy than random forest models without clustering. Essentially, operational inputs and penetration rate output are collected, which means little applicable in practice.

When predicting the penetration rate, optimisation techniques can be applied to optimise correlations of weighting in multiple regression, as demonstrated by Yagiz and Karahan (2015). On the other hand, optimisation techniques also serve to fine-tune

hyperparameters in machine learning models, such as the hyperparameters in the XGBoost model (Zhou et al., 2021a). In comparing different optimisation techniques, Figure 2.1(a) shows multiple regression models, while Figure 2.1(b) shows XGBoost models. Optimisation techniques improve accuracy, but the difference between different optimisation techniques remains minimal.

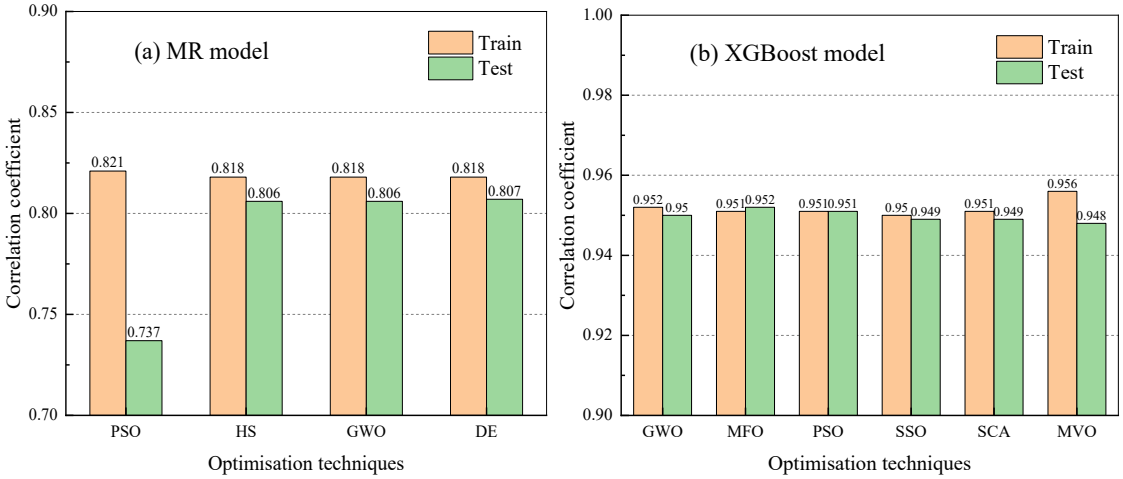


Figure 2.1 Comparison optimisation techniques (a) multiple regression models based on the Queen water tunnel; (b) XGBoost models based on the Pahang-Selangor raw water transfer

2.1.3.2 Machine learning for advance rate

The advance rate (AR) is a crucial indicator in tunnelling operations, calculated as the boring distance divided by the working time and stoppages. Compared with penetration rates, advance rates additionally consider stoppages due to TBM maintenance, cutters change, breakdowns, or tunnel collapses. When predicting advance rate, the ANN model by Benardos and Kaliampakos (2004) was limited to 11 samples in the Athens metro. In

contrast, the Pahang-Selangor raw water transfer allows for the development of more robust and reliable machine learning models (Armaghani et al., 2019; Zeng et al., 2021; Zhou et al., 2021b), but has the same practical issue of inaccessible operational inputs.

The above machine learning models for TBM performance are in low frequency, with each sample representing the average value in one boring segment, approximately 1–2 metres in length. In contrast, real-time predictions for advance rate involve working time and stoppages in high frequency, collecting one sample every few seconds (Sun et al., 2018; Mokhtari and Mooney, 2020). In the Northgate Link tunnelling project in Seattle, more than 500,000 high-frequency data were collected. Mokhtari and Mooney (2020) utilised the feature selection technique of RReliefF to select 10 inputs from 107 operational parameters. They developed six differentiated SVM models to predict real-time advance rates across distinct engineering soil units.

2.1.3.3 Machine learning for cutterhead load

Cutterhead torque (TO) refers to the twisting force applied to the cutterhead, whereas thrust force (TH) refers to the force that TBM exerts on the excavation face. The cutterhead torque or thrust force depends on the hardness and strength of the excavated material and TBM size and type. In addition to predicting real-time advance rates, Sun et al. (2018) built RF models to predict real-time cutterhead torque and thrust force within heterogeneous strata, in which Kriging interpolation was implemented to solve the

imbalance between geological and operational data.

Lin et al. (2022a; 2022b) developed PSO-LSTM and PSO-GRU models to predict cutterhead torque and thrust force, respectively. In the Shenzhen intercity railway, each sample represents the average value of a 1.6-metre-wide segment. They incorporated both geological (such as unit weight, cohesion, internal friction angle, compression modules, and natural water content) and operational parameters (such as face pressure, cutterhead torque, thrust force, penetration rate, penetration per revolution, screw conveyor speed, foam volume, foam pressure, grouting volume, and grouting pressure) within a segment. Using historical data in conjunction with the loop structure of LSTM and GRU models contributed to good model performance.

Bai et al. (2021) utilised only penetration rate and screw conveyor speed to build a machine learning model for predicting cutterhead torque, thrust force, and face pressure. In the optimisation process, seasonal-trend decomposition using Loess removes noise, and an SVM classifier identifies the location of the interbedded clay or stratum interface. Upon comparison of four machine learning algorithms, they discovered that gradient boosting machines exhibited the best performance, while linear regression was the least effective.

2.1.3.4 Machine learning for field penetration index

The field penetration index (FPI) serves as an evaluative measure for TBM efficiency in

the field, combining average cutter force and penetration per revolution in Eq. 2.14.

$$\mathrm{FPI} = \frac{\mathrm{TH}}{N_c * \mathrm{Pev}} \qquad 2.14$$

In a comparative study, Salimi et al. (2016) assessed empirical models, SVM, and ANFIS for predicting FPI, using data from Zagros lots 1B and 2 with 75 observations. Through the employment of principal component analysis (PCA), they identified uniaxial compressive strength (UCS) and joint spacing (Js) as the most influential parameters. Salimi et al. (2019) extended the dataset to 580 observations across seven tunnel projects for model generalisation. The proposed classification and regression tree (CART) and genetic programming models not only incorporated geological parameters of UCS, RQD, Js, and joint condition (Jc), but also innovatively added a rock type code for five distinct rock types. Again, the dataset was extended to 666 observations from eight tunnel projects. Alternatively, they found that the significance of input parameters varied according to the rock type codes, leading to four differentiated CART models (Salimi et al., 2022). By employing a variable importance measure (VIM) in the decision tree analysis, they elucidated the intricate relationship between the FPI and input parameters, thereby enhancing the understanding of how these factors interact to influence TBM performance in various geological contexts.

2.2 Time series forecasting

Time series forecasting is a real-time prediction using current and historical data to forecast future unknown values. Several studies using machine learning techniques for time series forecasting are shown in Table 2.5. It is crucial in TBM tunnelling for predicting TBM performance, cutterhead load, and moving trajectory in real time because operators can make necessary adjustments when potential issues are detected. Since the quality and quantity of data heavily influence model performance, moving average or wavelet transform are employed to eliminate noise and fine-grained variation to reveal the underlying information in time series data (Wang et al., 2021a; Xu et al., 2021; Shan et al., 2022; Shen et al., 2022). In machine learning, time series forecasting is a function of input parameters in Eq. 2.15.

$$\boldsymbol{Y_{n+1}} = f[\boldsymbol{W}(X_1, X_2, \dots, X_n), b] \qquad 2.15$$

where $\boldsymbol{Y}_{n+1}$ = Output vectors, target value in the next step

$X_1, X_2, \dots, X_n$ = Input vectors, historical sequential data

For effective time series forecasting, datasets not only need to be publicly available, possess a substantial number of samples, and have appropriate input parameters, but they must also be sequential. Operational data is frequently sourced from data acquisition systems, which capture details every few seconds or minutes. Conversely, geological data from site investigations is typically procured every few meters. This frequency disparity between operational and geological data is commonly addressed by interpolating between

two labelled rock mass properties. However, this method can obscure the temporal difference in geological data, diminishing the significance of rock mass properties in the dataset (Gao et al., 2021).

Table 2.5 Summary on time series forecasting

Outputs	**Inputs**	**Algorithms**	**Data size and dataset**	**Reference**
TO, TH, AR, FP	44 operational data (TO*, TH*, AR*, FP*)	RNN, LSTM, GRU	3000, Shenzhen metro	Gao et al. (2019)
TO	TO*	SVM, LSTM, RF	200000, Brenner base tunnel	Erharter and Marcher (2021)
PR	PR*, TO, TH, RPM, FP	LSTM, SVM, RF, ARIMA, ANN	12738, Thomson coastline	Fu and Zhang (2021)
PR	TO, TH, POW, PSLI, lithology, rock mass classification	LSTM, ARIMA, RNN	Hangzhou water source	Gao et al. (2021)
TO	1465-51 operational data (TO*)	residual CNN-LSTM, XGBoost, RF, SVM, LSTM, RNN, CNN	150000, Singapore metro	Qin et al. (2021)
TO	TO*	LSTM, SVM, RF, RNN, GRU, CNN	60000, Singapore metro	Shi et al. (2021)
PR, TO	20 operational data	Bi-LSTM, LSTM	25543, Sutong gas line	Wang et al. (2021a)
TO	10 operational data	Bi-LSTM+ incremental learning	1230725, Yinsong water	Huang et al. (2022)
PR, energy consumption	PR*, energy consumption*, net stroke, TO, TH RPM, SCS, FP, soil pressure	attention-GCN, DNN, RF, SVM	20130, Thomson coastline	Pan et al. (2022)
TO	TO*	GRU, SVM, RF, RNN, LSTM, CNN	50000, Singapore metro	Qin et al. (2022)

FP	10 operational data (FP*)+ CD, thickness of layers	genetic algorithm - GRU	1534, Luoyang metro	Gao et al. (2020)
TPI	TPI*	DBN	11816, Yinsong	Chen et al. (2021)
FPI	FPI*	DBN	8915, Yinsong	Feng et al. (2021)
PR	PR*	RNN, LSTM	550, Changsha and Zhengzhou metro	Shan et al. (2022)
PR	PR*, TO, TH, FP, RPM, CD, GC, WT, modified SPT,	RNN variant, RNN, ARIMA	550, Changsha and Zhengzhou metro	Shan et al. (2023a)
TO, TH	191-154-128-51-12 operational data+ RPM_set, PR_set	LSTM	4650, Yinsong	Li et al. (2021)
PR, TO, TH	PR, TO, TH, Pev, POW, gripper pad pressure, lithology, HC	SVM, ANN, RF, GBDT	4802, Yinsong	Wang et al. (2021b)
PR, TO, TH, RPM,	30 operational data (PR*, RPM*, TO*, TH*)	LSTM, CNN	7000, Yinsong	Xu et al. (2021)
HDH, VDH, HDT, VDT, roll, pitch	32 operational data	CNN-LSTM, ARIMA, LSTM	5005, Sanyang Road Tunnel	Zhou et al. (2019)
HDH, VDH, HDT, VDT, roll, pitch	14 operational data+ WT, thickness of layers, UCS, qz	LSTM, SVM, RNN	1200, Shenzhen railway	Shen et al. (2022)
HDH, VDH, HDT, VDT	33 operational data (HDH*, VDH*, HDT*, VDT*)+ CD, thickness of layers	GRU. RNN, SVM	18471, Guangfo railway	Zhang et al. (2022)
VDT, VDT, HDA, VDA	18 operational data (VDH*, VDT*, HAD*, VDA*)	GCN-LSTM, PCA-GRU, LSTM, XGBoost	12591, Thomson coastline	Fu et al. (2023)

* Time series forecasting includes historical targets

TPI, torque penetration index; PLSI, point loading strength index; HDH, horizontal deviation of the shield head; VDH, vertical deviation of the shield head; HDT, horizontal deviation of the shield tail; VDT, vertical deviation of the shield tail; HDA, horizontal deviation of the shield articulation; VDA, vertical deviation of the shield articulation; STP, standard penetration test; RPM_set, setting values of revolutions per minute; PR_set, setting values of penetration rate;

CNN, convolutional neural network; DNN, deep neural network; DBN, deep belief network; ARIMA, autoregressive integrated moving average; GBDT, gradient boosting decision tree; GCN, graph convolutional network;

2.2.1 Forecasting TBM performance

Time series forecasting of TBM performance is a real-time prediction using historical data to predict unknown TBM performance in the future. Real-time prediction of penetration rate can be a reference for effective scheduling, which helps logistics personnel and material for segment lining and grouting. It is also beneficial to make necessary adjustments when potential risks are detected based on predicted TBM performance ahead of the cutterhead. Moreover, an alarm can be triggered when the predicted cutterhead torque or thrust force exceeds the maximum values, further safeguarding the tunnelling operation.

On the one hand, high-frequency forecasts rely solely on historical TBM performance (Erharter and Marcher, 2021; Shi et al., 2021) or historical TBM performance and operational parameters (Gao et al., 2019; Qin et al., 2021; Wang et al., 2021a). Such forecasts overlook geological conditions ahead of the cutterhead and have limited success in predicting TBM performance a few seconds or centimetres ahead. On the other hand, low-frequency forecasts can predict a few metres in advance but have a limited number

of samples after preprocessing (Shan et al., 2022), which in turn affects model robustness. In addition, reduced accuracy is problematic for multi-step forecasts in both high-frequency and low-frequency data.

2.2.1.1 High-frequency forecast

High-frequency one-step forecasting predicts the next-step TBM performance using real-time operational data, which is frequently collected at intervals ranging from seconds to minutes via a data acquisition system. They can be achieved with high accuracy using RNN, LSTM, and GRU, which outperform others by incorporating current and historical data.

Univariate models leverage historical data by machine learning, e.g. the next-step cutterhead torque is predicted based on the trend of historical cutterhead torque itself (Erharter and Marcher, 2021). Neglecting the temporal effect of targets, Wang et al. (2021a) selected 20 operational parameters to real-time predict penetration rate and cutterhead torque, while Huang et al. (2022) selected 10 operational parameters by the SelectKBest algorithm and predicted the next-step cutterhead torque. It is more desirable to integrate both historical targets and related operational parameters into input features (Gao et al., 2019; Qin et al., 2021). For instance, Pan et al. (2022) demonstrated that models incorporating historical penetration rates outperformed their counterparts that did not. Fu and Zhang (2021) proposed a predictive framework for the next-step penetration

rate, shaping the operational data as given in Eqs. 2.16 and 2.17.

$$\boldsymbol{X} = \begin{bmatrix} x_{1,t-n} & x_{1,t-n+1} & \cdots & x_{1,t-1} & x_{1,t} \\ x_{2,t-n} & x_{2,t-n+1} & & x_{2,t-1} & x_{2,t} \\ \vdots & \vdots & \ddots & \vdots & \\ x_{k-1,t-n} & x_{k-1,t-n+1} & & x_{k-1,t-1} & x_{k-1,t} \\ x_{k,t-n} & x_{k,t-n+1} & \cdots & x_{k,t-1} & x_{k,t-1} \end{bmatrix} \quad 2.16$$

$$\boldsymbol{Y} = x_{k,t} \quad 2.17$$

The input matrix $\boldsymbol{X}$ represents operational parameters from $x_1$to x_k and from *t*-*n* to *t*, noting that the bottom right entry is replaced by the last-step target $x_{k,t-1}$. The objective of output matrix $\boldsymbol{Y}$ is to predict the operational parameter x_k at time *t*. A key limitation, however, is that operational parameters $x_1$to x_{k-1} at time *t* are not known in advance, making it infeasible in practice.

Moreover, Gao et al. (2021) incorporated operational and geological data, including point loading strength index (PLSI), lithology, and rock mass classification, to predict the next penetration rate. PLSI illustrates the linear correlation between UCS and BTS, whereas lithology and rock mass classification are indicated by one-hot encoding. Notably, the method involved mapping 127 geological entries onto 2570 operational data points, diluting the significance of the geological data in time series forecasting.

However, it is less meaningful to predict TBM performance mere seconds or millimetres ahead, as shown in Table 2.6. Therefore, multi-step forecasts for cutterhead torque were explored in the dataset from Singapore metro lines (Shi et al., 2021; Qin et al., 2022).

These univariate models employed the last ten steps of cutterhead torque to predict the next five steps. It was crucial to decompose the original cutterhead torque sequence into multiple sub-sequences by variational mode decomposition and empirical wavelet transform, but errors increase with an increasing forecast horizon. Erharter and Marcher (2021) failed to predict the cutterhead torque 100 steps ahead using data from the historical 50 steps because it is too far away from the cutterhead.

Table 2.6 Historical data and forecast horizon of time series forecasting

Outputs	**Category**	**Historical data**		**Forecast horizon**		**Reference**
		Step behind	**Distance behind**	**Step ahead**	**Distance ahead**	
TO, TH, AR, FP	High-frequency	5 steps	1.25 mm*	1 step	0.25 mm*	Gao et al. (2019)
TO		6 steps	22.4 mm*	1 step	3.73 mm*	Huang et al. (2022)
HDH, VDH, HDT, VDT		5 steps	47.6 mm*	1 step	9.52 mm*	Zhang et al. (2022)
PR		7 steps	140 mm	1 step	20 mm	Fu and Zhang (2021)
HDH, VDH, HDT, VDT, roll, pitch		60 steps	425 mm*	1 – 6 steps	7.08 – 42.5 mm*	Zhou et al. (2019)
TO		50 steps	2.75 m	1 or 100 steps	0.055 or 5.5 m	Erharter and Marcher (2021)
TPI	Low-frequency	5 steps	5 m	1 step	1 m	Chen et al. (2021)
FPI		7 steps	7 m	1 step	1 m	Feng et al. (2021)
FP		5 steps	7.5 m	1 step	1.5 m	Gao et al. (2020)
HDH, VDH, HDT, VDT, roll, pitch		4 steps	6.4 m	1 step	1.6 m	Shen et al. (2022)

PR	5 steps	7.5 m	1 – 5 steps	1.5 – 7.5 m	Shan et al. (2022)

* Distance is estimated based on the time step, sampling period, and average penetration rate

2.2.1.2 Low-frequency forecast

High-frequency data is preprocessed into low-frequency data where each data point represents a fixed segment or a boring cycle, typically 1–2 metres. A typical boring cycle is categorised into four stages, including start-up (S1), ascending (S2), steady-state (S3), and end stages (S4). In Figure 2.2, penetration rate (PR), revolutions per minute (RPM), cutterhead torque (TO), and thrust force (TH) are scaled to the same magnitude for comparison.

- Start-up stage (S1): the TBM initiates to ramp up. Revolutions per minute (RPM) attain stability, but the TBM does not commence cutting rock.
- Ascending stage (S2): the TBM begins to cut rock. Operational parameters of penetration rate (PR), thrust force (TH), and cutterhead torque (TO) increase before achieving dynamic stability.
- Steady-state stage (S3): the TBM cuts rock at a constant speed, and four operational parameters remain dynamic stability.
- End stage (S4): the TBM gradually halts with four operational parameters decreasing to zero.

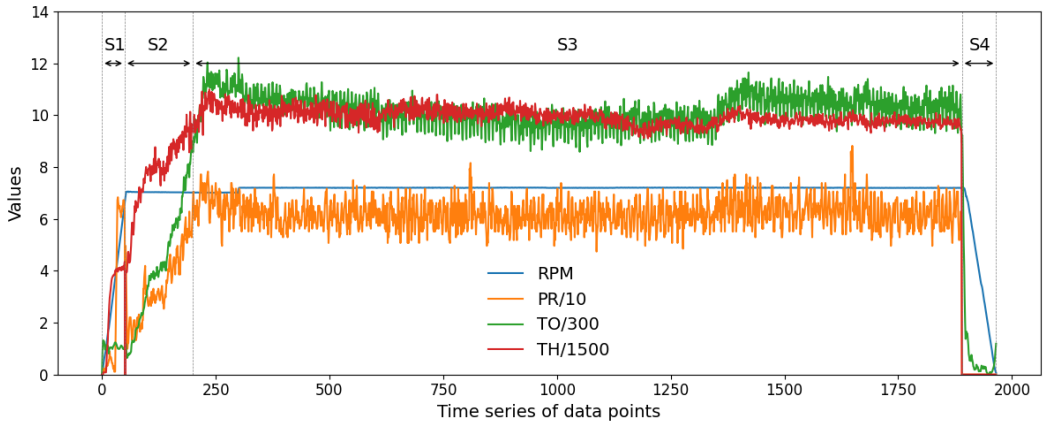


Figure 2.2 Four stages in a typical boring cycle

According to different rock types (such as diorite, granite, and limestone), differentiated deep belief network models are built to predict the toque penetration index (TPI) by Chen et al. (2021), and to predict field penetration index (FPI) by Feng et al. (2021). These univariate models used historical data (5 or 7 m behind) to predict next-step TBM performance (1 m ahead), as shown in Table 2.6. The next-step face pressure (FP) is subsequently investigated by GRU, combining 10 related operational and geological parameters of cover depth and thickness of rock types, namely multivariate models (Gao et al., 2020). An advanced RNN model that integrates historical operational parameters and setting values of penetration rate (PR_set) and revolutions per minute (RPM_set) was proposed, leading to a significant improvement in forecasting one-step cutterhead torque (TO) and thrust force (TH).

In contrast, Shan et al. (2022) employed RNN and LSTM to predict near-future TBM performance (1.5–7.5 metres ahead), focusing on the difference in geological conditions

between training data and test data. However, as the models predicted further into the future, the accuracy again decreased due to the reduced impact farther away from the TBM cutterhead (Shan et al., 2022, 2023b).

Regarding the number of steps back required to predict future TBM performance, Table 2.6 demonstrates that the number of steps used for training ranges from 5 to 10, except for those who used data from the last 50 steps. High-frequency prediction normally uses data just a few millimetres behind for training, while low-frequency prediction uses data up to seven metres behind. Nevertheless, these data are collected a few millimetres to a few meters away from the current cutterhead location and essentially reflect the current operation of the TBM (Shan et al., 2023b).

Li et al. (2021); Wang et al. (2021b) and Xu et al. (2021) used the time series from the first two-minute ascending stage to forecast the average TBM performance at the steady-state stage. Separating the start-up, ascending, and steady-state stages is important by identifying two change points. The cumulative sum (CUSUM) is used to identify the second change point, and the maximum gradient is used to recognise the first change point (Wang et al., 2021b), similar to Figure 2.2. Li et al. (2021) focused on feature selection from 191 operational parameters on the dataset from the Yinsong water diversion project. They successively filtered 37 parameters with constant, 26 low-variance parameters, and 77 highly correlated parameters, and they finally selected 12 operational parameters by

variable importance measure (VIM) in the decision tree. It is noted that setting values of penetration rate and revolutions per minute are also considered in their LSTM models.

2.2.2 Forecasting attitude and position

Shield tunnelling misalignment is a critical factor compromising tunnel quality, representing a deviation between the designed tunnel axis and movement trajectory. The occurrence of misalignment complicates the assembly process and causes tunnel quality problems, such as segment cracks, segment dislocations, and even water leakage. Additionally, misalignment can exacerbate ground disturbances and increase the risk of tunnelling-induced geo-hazards, especially in complex ground conditions (Zhang et al., 2022).

Shield tunnelling misalignment primarily results from challenges in controlling the attitude and position. This misalignment is due to the uneven shield quality, varying frictional resistances, and complex shield operations. Moreover, adverse ground conditions, like tunnelling in mixed strata, lead to irregular vibrations during excavation. The shield trajectory regulation is currently executed using a feedback-based method that relies on real-time attitude and position measurements from an automatic navigation system. However, this approach, termed the hysteresis effect, has a time and distance lag, leading to proactive operations to prevent significant deviations in the shield movement trajectory.

Predicting the shield attitude and position in advance can aid the operators in making timely adjustments, thereby addressing the limitations of feedback-based methods. Time series forecasting models predict the attitude and position at a future period by feeding the historical data of the relevant parameters. The parameters of attitude and position include horizontal deviation of the shield head (HDH), vertical deviation of the shield head (VDH), horizontal deviation of the shield tail (HDT), vertical deviation of the shield tail (HDT), roll, and pitch. Consequently, various research efforts introduced hybrid deep learning models to predict the subsequent parameters in high-frequency (Zhang et al., 2022; Fu et al., 2023) and low-frequency data (Shen et al., 2022) in real time. Multi-step forecasting (6 steps ahead) was investigated using historical operational parameters (60 steps behind) by Zhou et al. (2019). Their CNN-LSTM models demonstrated comparable accuracy between one-step and six-step forecasts. For example, accuracies in one-step forecasts are better than six-step forecasts for pitch, HDT, VDT, and HDH but worse than six-step forecasts for roll and VDH. This is because TBM operations and geological conditions are almost unchanged over a forecasting span of 7.08–42.5 mm, leading to a slight difference in attitude and position.

2.3 Rock mass classification

Selecting the appropriate TBM type and adjusting operational parameters according to rock mass parameters is pivotal for maintaining the efficiency and safety of TBM

excavation. High rock strength zones often lead to low penetration rates due to the limited maximum thrust force. In addition, the joint conditions greatly influence the degree of difficulty of rock breakage.

Although geological surveys are usually available before excavation, borehole drilling is conducted at limited locations and offers a coarse description of a geological profile. Advanced non-destructive geophysical prospecting techniques, such as tunnel seismic prediction and ground penetrating radar, allow visualisation of upcoming geological challenges like adverse structures or water bodies. Although these forward prospecting techniques can yield detailed insights into geological conditions, they are often expensive, time-consuming, and necessitate a halt in TBM operations during detection. In other words, they cannot keep pace with the TBM tunnelling process, resulting in a notable lag in detecting rock information (Zhang et al., 2019b). Due to the abundant data collected by the data acquisition system, machine learning applications can address these geological issues in real time, as outlined in Table 2.7.

Table 2.7 Summary on rock mass classification

Outputs	**Inputs**	**Algorithms**	**Data size and dataset**	**Reference**
GC	GC, SPT, clay content	Neural- hidden Markov model	over 200, Zhonghe water	Leu and Adi (2011)
UCS, BI, DPW, α	TH, TO, RPM, PR, chamber pressure	SST-SVM	180, Yinsong	Liu et al. (2019)
UCS, BI,	TH, TO, RPM, PR,	SA-ANN	320, Yinsong	Liu et al.

DPW, α	chamber pressure			(2020a)
GC	TH, TO, PR	ANN	6414, four projects	Jung et al. (2019)
HC	Fn, Fr, Pev, RPM, FPI, TPI	AdaBoost-CART, ANN, KNN, SVM	3166, Yinsong	Liu et al. (2020b)
Rock type	9 operational data	attention-LSTM, LSTM, KNN, CART, logistic regression	7639, Yinsong	Liu et al. (2021b)
HC	195-152-66-10 operational data	SMBO-CatBoost, XGBoost, RF, SVM, KNN	4464, Yinsong	Bo et al. (2022)
HC	10 operational data	Stacking (SVM, KNN, RF, GBDT +GBDT)	7538, Yinsong	Hou et al. (2022)
Rock type	TH, TO, RPM, PR	K-means + SVM, RF, KNN	12038636, Yinsong	Zhang et al. (2019b)
Rock mass class	10 operational data, FPI, TPI	spectral clustering+ DNN, RF, KNN, AdaBoost	10807, Yinsong	Wu et al. (2021)
GC	c, φ, E; PR, RPM, TH, TO, TH/TO, chamber pressure	canopy algorithm, K-means+ GBDT, RF, SVM	491, Shenzhen metro	(Yang et al., 2022a)
GC	20 operational data	SSAE-DNN, SVM, DT, KNN, RF	548915, 4348, Mumbai metro	Yu et al. (2021)
GC	177 operational data	CDAE-DNN, SVM, DT, KNN, RF	15851, Thomson coastline	Yu et al. (2022)

SST, stacked single-target; KNN, k-nearest neighbour; CatBoost, categorical boosting; SMBO, sequential model-based optimisation; SSAE, stacked sparse autoencoder; CDAE, constrained dense convolutional autoencoder

2.3.1 Predicting rock mass parameter

During excavation, TBMs are highly susceptible to variations in rock mass conditions. Unknown rock mass information may lead to inappropriate operation and even low safety

and efficiency of excavation. As a result, an accurate and reliable prediction method of rock mass parameters is required in TBM tunnelling. It is a regression analysis where TBM operational parameters serve as input variables to predict rock mass parameters.

Liu et al. (2019) developed stacking SVM models to predict rock mass parameters. They established models by training and used initial prediction results as the new inputs to refine and finalise the model. The rock mass parameters involved UCS, BI, DPW, and α with 180 samples from site investigation. The selected input variables, aligned with the position of rock mass parameters, comprised cutterhead torque, thrust force, revolutions per minute, penetration rate, and cutterhead power. Liu et al. (2020a) extended the dataset to 320 samples and integrated ANN with simulated annealing—a nonlinear optimisation technique—to overcome the local optimum issues in gradient descent. The prediction models for rock mass parameters were established using limited geological information and corresponding operational data (labelled data). Meanwhile, a large amount of operational data that lacked geological information (unlabelled data) remained unused.

2.3.2 Rock mass classification and clustering

Different lithology affects the wear of the cutter, which in turn affects the efficiency of TBM tunnelling. Liu et al. (2021b) integrated LSTM and the attention mechanism to predict the lithology in the steady-state stage, and the inputs are nine operational parameters in the first 15 seconds in the ascending stage. However, traditional methods

of lithology identification, such as geological surveys and geological analysis, are generally reliable and known in advance.

During TBM excavation, the rock mass classification forms an intuitive basis for adjusting operational parameters by operators. Based on the dataset from the Songhua water diversion project, several studies have predicted the hydropower classification (HC), a type of rock mass classification, using advanced classifiers (Liu et al., 2020b; Bo et al., 2022; Hou et al., 2022). Table 2.8 shows their average accuracy and F1-score.

Table 2.8 Comparison of rock mass classification in the Songhua water diversion project

Algorithm	**Accuracy**	**F1-score**	**Reference**
AdaBoost-CART	0.865	0.770	Liu et al. (2020b)
SMBO-CatBoost	0.938	0.937	Bo et al. (2022)
Stacking	0.931	0.928	Hou et al. (2022)

In the AdaBoot-CART model, the mean value and standard deviation of operational parameters were used as input features, including normal force, rolling force, penetration per revolution, revolutions per minute, field penetration index, and torque penetration index (Liu et al., 2020b). The model attained an accuracy of 0.865 and an F1-score of 0.770. In contrast, the SMBO-CatBoost and stacking models utilised the mean values of 10 operational parameters as input features. These models achieved impressive results: 0.938 accuracy and 0.937 F1-score for SMBO-CatBoost (Bo et al., 2022), and 0.931 accuracy and 0.928 F1-score for the stacking model (Hou et al., 2022).

Since different data sizes and data preprocessing can influence the predicted results, SMBO-CatBoost and stacking models performed equivalently and much better than the AdaBoost-CART model. The poor performance is attributed to input redundancy. For example, thrust force is linearly related to normal force in Eq. 2.6 when the number of cutters and TBM diameter are constant in a tunnel project, and thrust force is also highly related to the field penetration index in Eq. 2.14 when the number of cutters is constant and penetration per revolution is also known. As a result, the normal force is highly related to the field penetration index, and similarly, the rolling force is highly related to the torque penetration index. These highly correlated input features can be redundant in machine learning.

The geological data are labelled in rock mass classification and rough obtained from actual site investigation. The vital defect of the rock mass classification lies in the approximate fitting of the HC of other segments (Liu et al., 2020b; Bo et al., 2022; Hou et al., 2022). This means that the fitted label might not accurately represent the actual geological conditions, which may reduce the prediction reliability. For example, a segment having rock fragments shows a weaker rock mass classification of four, but adjacent segments determined by borehole drilling show a stronger rock mass classification of two. The fitted label indicates a wrong rock mass classification of two, but the actual label of four is more likely to be reflected by operational parameters.

Essentially, it is an imbalance between the operational and geological data. Using the operational data, Zhang et al. (2019b) developed the K-means model to identify possible rock mass types and then established a model to predict labelled rock mass types based on a support vector machine classifier. Since the K-means model is an unsupervised clustering algorithm, the labels it produces may not necessarily represent specific types of rock mass. The combination of clustering and classification can make full use of unlabelled and labelled operational data to increase the accuracy and robustness of the rock mass classification. However, it is a high-frequency prediction compromising 12,038,636 samples where one predicted result represents rock mass type in one second. It means a computational model and predicted results are meaningless because rock mass types in a few millimetres should be unchanged. Similarly, Wu et al. (2021) combined spectral clustering and deep neural networks to predict rock mass types in low-frequency data, in which a sample represents the average values in the steady-state stage in one boring cycle.

Regarding limited rock mass types, related operational parameters are used to extract geological features via a stacked autoencoder. Extracted geological features are added into deep neural networks to predict rock mass types (Yu et al., 2021; Yu et al., 2022). It is noted that the training set used the segments corresponding to the boreholes, while the test set used the first 40 samples of the corresponding adjacent segments. Although recognising the geological formation of the tunnel face, the semi-supervised models are

less meaningful in high frequency, like the studies mentioned above. In addition, the training and test sets are not random and are specifically selected from the segments that assume consistent rock mass type in the borehole so that the models can lead to low generalisation and overfitting.

2.4 Summary

Chapter 2 presents a comprehensive literature review, exploring various methodologies and techniques for optimising TBM performance. It provides a holistic understanding of the literature related to TBM performance regression, time series forecasting, and rock mass classification.

Regarding TBM performance regression, theoretical methods, with a spotlight on the CSM model, offer insights into mechanics principles and full-size linear cutting tests. Empirical methods, particularly the NTNU model, emphasise field performance and hands-on experiences in predicting TBM performance. Machine learning models make significant advancements, showcasing how data-driven algorithms transform the landscape of TBM performance regression. However, the challenge is not merely about prediction but about achieving practical and actionable insights.

Transitioning to time series forecasting, Section 2.2 discerns TBM performance based on data granularity—high-frequency or low-frequency datasets. It elucidates the prowess of

machine learning in harnessing historical trends and data to forecast future TBM performance. Additionally, it sheds light on forecasting attitude and position, emphasising the importance of spatial parameters in tunnelling operations. Low-frequency forecasting is a reference for operators to adjust TBM operations when risks are predicted in advance.

Lastly, Section 2.3 commences with exploring techniques and methodologies to predict rock mass parameters, which are pivotal in tunnel construction. The discussion culminates with an examination of rock mass classification and clustering, highlighting the significance of categorising rock masses for optimised tunnelling processes. In the thesis, semi-supervised learning offers a strategy to harness unlabelled data by generating confident pseudo-labels, facilitating enhanced training processes.

Chapter 3. Methodology

3.1 Data processing method

The data generated during tunnel construction is extensive and diverse, encompassing geological and geotechnical surveys, operational parameters, and monitoring statistics related to surface settlement and structure deformation. In the Yinsong water diversion project, the data acquisition system recorded 199 operational parameters per second, accounting for 86,400 data points per day. Data processing becomes indispensable since the quality and quantity of data heavily influence the performance of machine learning models. This involves a series of steps: deleting outliers to maintain data integrity, removing noise to improve clarity, selecting features to streamline the dataset, and augmenting minority class samples to ensure a balanced model training. The data can be refined and made more conducive for machine learning applications through these methods, thereby improving model accuracy and reliability.

3.1.1 Outlier detection

Outliers are data points that significantly differ from other observations in a dataset, which are considered errors and should be removed. Two quantitative characteristics of outliers are few and different in a dataset. There are several methods available to detect and remove outliers. Assuming the dataset follows a normal distribution, 99.7% of data

points lie within three standard deviations of the mean. As a result, data points that fall outside the range of the mean plus or minus three standard deviations can be removed according to the 3-sigma rule (Feng et al., 2021; Wang et al., 2021b; Xu et al., 2021; Bo et al., 2022; Yang et al., 2022a). Another method is the interquartile range (IQR) method, which sets up a minimum and maximum fence based on the first quartile (Q1) and the third quartile (Q3), respectively (Hou et al., 2022). Any observations that exceed 1.5 times the IQR below Q1 or above Q3 are considered outliers and should be removed.

In addition to statistical methods, the isolation forest (IForest) is an unsupervised decision-tree-based algorithm that employs binary trees to isolate outliers without the process of normal instance profiling explicitly (Liu et al., 2008; Wu et al., 2021; Yu et al., 2021; Yu et al., 2022). Consider a two-dimensional data set in Figure 3.1: points **A** and **B** are isolated by randomly partitioning along two coordinate axes. The IForest randomly selectively picks an attribute and a split value (between its minimum and maximum values). This procedure recurs, subdividing the data until only one unique point remains or all remaining points are the same. The number of partitions to isolate a point can be interpreted as path length. Therefore, entities with shorter path lengths, like point **B**, are deemed more anomalous than point **A**.

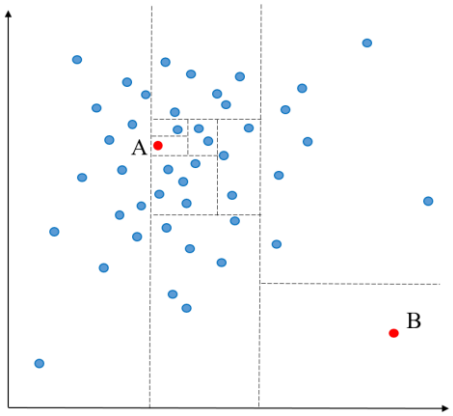


Figure 3.1 Schematic diagram of the isolation forest in two-dimensional data

3.1.2 Data smoothing

Data smoothing is a technique commonly used to eliminate fine-grained variation in time series to remove noise and reveal the signal of underlying information. The simple moving average (SMA) method creates a smoothed version by averaging observations within a specific period, assigning equal weight to each observation (Erharter and Marcher, 2021; Xu et al., 2021; Yu et al., 2021; Shan et al., 2022). In contrast, the exponential moving average (EMA) method places greater weight and significance on the most recent data points while gradually reducing the weight on older data points. These moving averages can be expressed as follows.

$$s_n = (x_n + x_{n-1} + \cdots + x_{n-k+1})/k \qquad 3.1$$

$$s_n = \lambda x_n + (1 - \lambda) s_{n-1} \qquad 3.2$$

where s_n = Smoothed data point

x_n = Original data point

k = Sliding window size

λ = Smoothing factor between 0 and 1.

In Eq. 3.1, the SMA is the unweighted mean of the past *k* data points. In Eq. 3.2, the EMA is determined by a smoothing factor λ. The smoothed statistic s_n is a weighted average of the recent observation x_n and the previous smoothed statistic s_{n-1}. The values of λ closer to 1 correspond to less smoothing and are more sensitive to recent changes in the data. Conversely, the values of λ closer to 0 correspond to a higher degree of smoothing and are less responsive to recent changes.

3.1.3 Feature selection

Feature selection is a crucial process for handling high-dimensional data, where the primary objective is to identify the most relevant features that can offer valuable insights into the underlying patterns and relationships within the data. However, many selected features are based on prior experience of laboratory tests and field studies, ignoring the effects of uncertain factors. The variance threshold method removes features that do not meet a specified threshold (Li et al., 2021), including zero-variance features that have the same value across all samples.

Pearson correlation coefficient (PCC) measures the linear relationship between two or more variables and can quantify the strength and direction of the linear relationship (Yagiz, 2009; Mokhtari and Mooney, 2020; Li et al., 2021; Liu et al., 2021b; Xu et al., 2021; Hou et al., 2022; Lin et al., 2022b). Given two variables, the PCC is calculated in Eq. 3.3 on a

scale between -1 and 1.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad 3.3$$

where r_{xy} = Pearson correlation coefficient

x_i, y_i = Sample points

$\bar{x}, \bar{y}$ = Mean values of sample points

For supervised learning, features with a higher absolute value between each feature and the target variable indicate a stronger relationship with the target variable. Features with very low correlation are usually considered for removal. However, a low correlation does not necessarily mean the two variables are independent; they might just have a non-linear relationship.

3.1.4 Oversampling

For the imbalanced sample set, oversampling is to increase the number of minority class samples. Through oversampling, the number of different classes can become relatively balanced. The synthetic minority over-sampling technique (SMOTE) is an oversampling technique that can balance the categories of the dataset in rock mass classification (Liu et al., 2020b; Hou et al., 2022; Yang et al., 2022a).

SMOTE is based on the idea of *k* nearest neighbours (KNN) and interpolation to generate samples of minority classes (Chawla et al., 2002). For the minority class sample x_i, the Euclidean distance between x_i and other minority class samples are calculated to obtain

k nearest neighbours (generally, k = 5). A minority class sample x_j is selected randomly from the k neighbours, and the new sample x' is generated in Eq. 3.4.

$$x' = x_i + \chi(x_i - x_j) \qquad 3.4$$

where x' = Generated sample

χ = random value between 0 and 1

Until all minority classes are equally represented, this process is repeated for each of them.

3.2 Machine learning method

The terms 'Artificial intelligence (AI)', 'Machine learning (ML)' and 'Deep learning (DL)' are commonly are often bandied about in discussions of smart software systems, and while they are frequently used interchangeably, each represents a distinct level of computational intelligence.

At the broadest level, 'Artificial intelligence' deals with the creation of machines capable of emulating human intelligence. It builds a system that can perform tasks, from problem-solving and decision-making to voice recognition and natural language processing. Nested within AI is 'Machine Learning', which can be visualised as the brain behind the operation. Machine learning is predicated on the idea of feeding machines vast amounts of data and then using statistical methods to enable them to learn from this data. Instead of being explicitly programmed to perform a task, a machine learning system utilises patterns and inference to make decisions and predictions. 'Deep Learning' is an even

more specialised subset of ML. Deep learning seeks to replicate this complex structure through 'deep' neural networks with multiple layers to learn from large amounts of data. Deep learning is particularly powerful for tasks that involve massive datasets, like recognising objects in billions of images or human speech.

3.2.1 Support vector machines

Support vector machines (SVM) is a machine learning algorithm used for both classification and regression tasks. SVM has gained popularity because of the ability to handle high-dimensional data and produce accurate results with limited data (Hearst et al., 1998; Mahdevari et al., 2014; Salimi et al., 2016; Liu et al., 2019; Zhang et al., 2019b; Mokhtari and Mooney, 2020). Given a dataset $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where x_i is a data point and $y_i \in \{-1, +1\}$ is the class label, the objective of SVM is to find the optimal hyperplane that separates the data into different classes. For a two-class problem in Figure 3.2, linear SVM aims to maximise the margin between the two hyperplanes, leading to the optimisation problem in Eqs. 3.5–3.6. The data points closest to the separating hyperplane are termed support vectors.

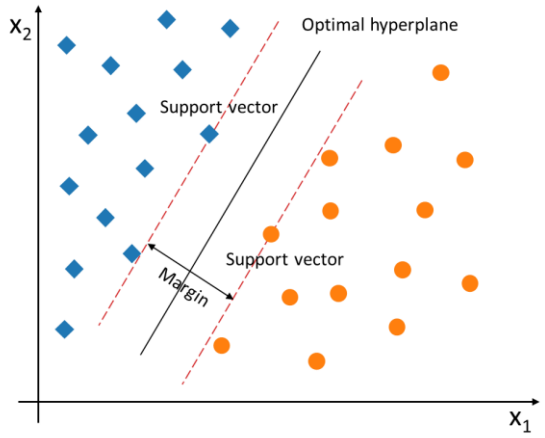


Figure 3.2 Schematic diagram of support vector machine with linearly separable data

Minimise $$\frac{1}{2}\|W\|^2 \quad (3.5)$$

Subject to $$y_i(Wx_i + b) \geq 1 \quad (3.6)$$

where W = Weight vector

$\|W\|$= Euclidean norm of the weight vector

b = Bias (constant)

When faced with non-linearly separable data, SVM employs slack variables for inevitable misclassification, determining the width of the tube around the hyperplane. This results in a trade-off between maximising the margin and minimising misclassification. The optimisation problem becomes Eqs. 3.7–3.8.

Minimise $$\frac{1}{2}\|W\|^2 + C\sum_{i=1}^{n}\xi_i \quad (3.7)$$

Subject to $$y_i(Wx_i + b) \geq 1 - \xi_i \quad (3.8)$$

where C = Regularisation parameter

ξ = Degree of misclassification of the point x_i

In addition, kernel functions $K(x_i, x_j)$ are selected to map input data into a high-

dimensional space by computing the dot products between the images of all pairs of data. Among the spectrum of available kernels, the radial basis function kernel is widely used and is illustrated in Eq.3.9.

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \qquad 3.9$$

where $K(x_i, x_j)$ = Kernel function

γ = Variance of radial basis function kernel

However, SVM can be computationally expensive for large datasets and sensitive to kernel functions and hyperparameter tuning.

3.2.2 Random forest

Decision tree (DT) is a prevalent machine learning algorithm for classification and regression analysis because of its simplicity and interpretability. It is based on a hierarchical structure where each node represents a decision or test of a specific feature. The tree is built by recursively splitting the data into smaller subsets based on the feature that provides the most information gain or reduction in entropy. After constructing the tree, new data can be classified by following the path from the root to the relevant leaf node.

Building on the foundation of decision trees, random forest (RF) introduces an ensemble approach, leveraging multiple trees to enhance stability and diminish the susceptibility to

overfitting (Breiman, 2001; Sun et al., 2018; Fu et al., 2022; Kannangara et al., 2022; Yang et al., 2022c). This technique is particularly advantageous with large datasets, producing highly accurate predictions. As illustrated in Figure 3.3, RF employs bootstrapping to select samples with replacements to form a new dataset randomly. For each of bootstrap sample, a decision tree is constructed by randomly selecting features and finding the best split. In the context of regression, final predictions are derived as the average outputs from all trees, while for classification tasks, a majority vote determines the outcome. Fine-tuning parameters, such as the number of trees, maximum number of features, and minimum number of leaves, can optimise the model performance.

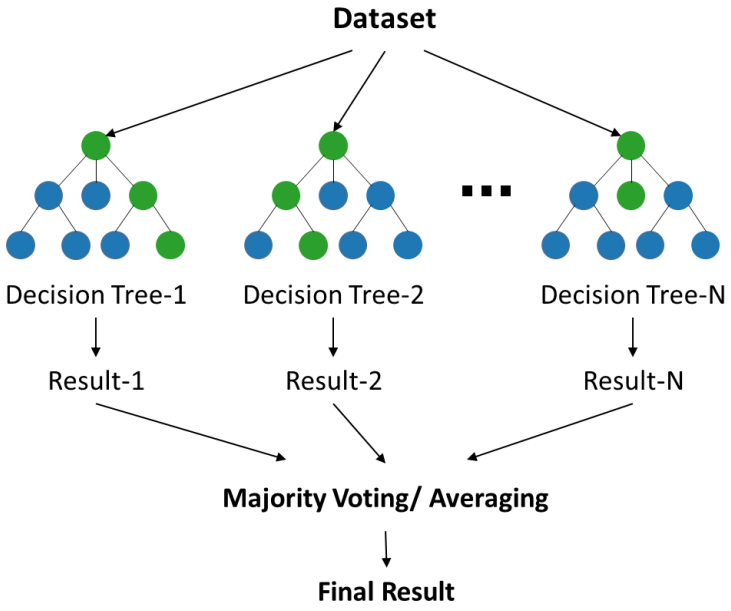


Figure 3.3 Schematic diagram of the random forest algorithm

3.2.3 Artificial neural network

Artificial neural networks (ANNs) are deep learning algorithms inspired by the biological neural networks in the brains of animals (Jain et al., 1996). ANN consists of an input layer, one or more hidden layers, and an output layer, and are fully connected between layers in Figure 3.4. Neurons are the basic components of computation that receive input from other neurons and compute an output. The output of one neuron becomes the input to other neurons, and this process continues until the final output is produced. The equations between layers are expressed in Eqs. 3.10–3.11.

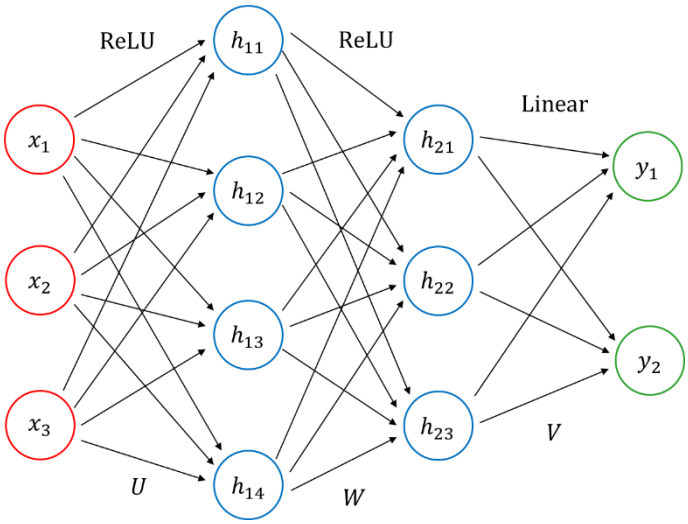


Figure 3.4 Schematic diagram of the artificial neural network

$$h_1 = \text{ReLU}(Ux + b_i) \quad 3.10$$

$$h_2 = \text{ReLU}(Wh_1 + b_h) \quad 3.11$$

$$y = Vh_2 + b_o \quad 3.12$$

where U, W, V = Weights of input, hidden, and output layers

b_i, b_h, b_o = Bias of input, hidden, and output layers

h_1, h_2 = Hidden layers

$\mathrm{ReLU}(x) = \max(0, x)$ = Rectified linear unit

$\mathrm{ReLU}(x)$ behaves a linear function when x is greater than 0 and becomes 0 otherwise. Notably, linear activation function is applied in the last hidden layer of Eq. 3.12, also known as a dense layer with extracted features. Overfitting is a main issue in ANN, often resulting from excessive feature dimensions, overly complex structures, an abundance of noise, or insufficient training data (Srivastava et al., 2014).

ANNs are widely used in applications like TBM performance regression in geotechnical problems (Benardos and Kaliampakos, 2004; Suwansawat and Einstein, 2006; Yagiz, 2009; Armaghani et al., 2017). Various ANN variants have also been developed to improve model accuracy, including the wavelet neural network (Pourtaghi and Lotfollahi-Yaghin, 2012), radial basis function network (Chen et al., 2019), general regression neural network (Zhang et al., 2020a), and extreme learning machine (Zeng et al., 2021).

3.2.4 Recurrent neural network

Recurrent neural networks (RNNs) are deep learning algorithms specially designed to recognise patterns in sequential data, such as speech recognition, natural language processing, or time series forecasting (Schuster and Paliwal, 1997). RNNs are characterised by recurrent connections to maintain an internal state or memory, which enables them to capture temporal dependencies, illustrated in Figure 3.5. The last hidden

state h_{t-1} and recent inputs x_t contribute to the recent hidden state h_t in Eq. 3.13. The recent output y_t extracts temporal features from h_t in a linear transformation in Eq. 3.14.

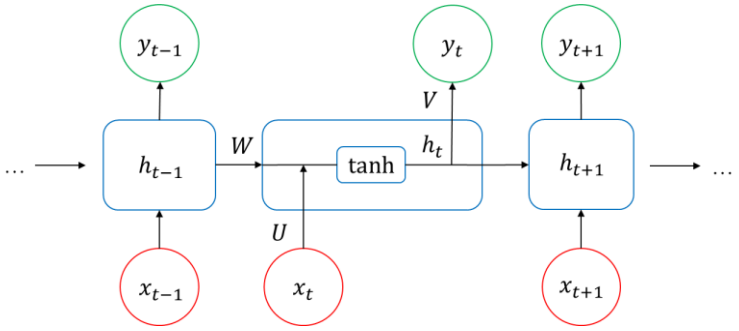


Figure 3.5 Schematic diagram of the recurrent neural network

$$h_t = \tanh(Ux_t + Wh_{t-1} + b_i) \quad 3.13$$

$$y_t = Vh_t + b_o \quad 3.14$$

where $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ = Hyperbolic function

RNNs offer a robust framework for time series forecasting because of inherent memory capability (Gao et al., 2019; Shan et al., 2022). However, RNNs find it challenging to capture long-term dependencies in time series due to the vanishing and exploding gradient problems (Bengio et al., 1994).

3.2.5 Advanced recurrent neural network

In addition to the time series data, some parameters are not time-dependent but have an impact on future values. These inputs can be categorised into two types: sequential and non-sequential data. While RNN is adept at handling sequential data owing to its loop

architecture, it cannot process non-sequential data. Therefore, an advanced RNN with a fully connected layer for non-sequential data is proposed, as shown in Figure 3.6.

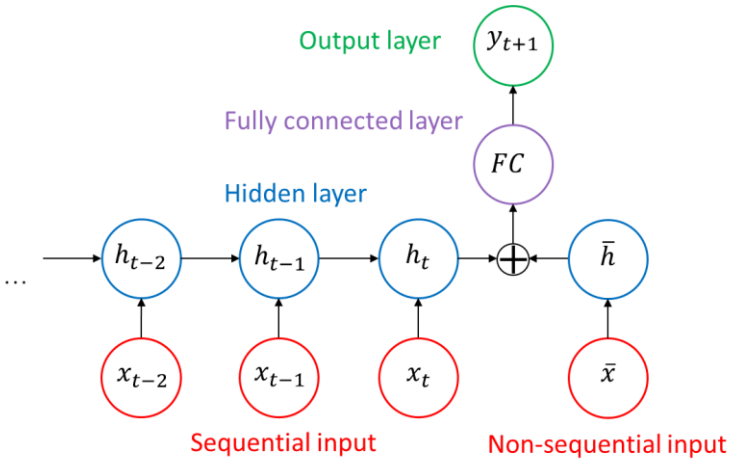


Figure 3.6 Schematic diagram of the advanced recurrent neural network

The inherent of algorithms is to extract relevant features from input data. RNN extracts temporal features from the time series x_{t-2}, x_{t-1} and x_t through Eq. 3.13 with a result of the recent hidden state h_t. Non-sequential inputs $\bar{x}_t$ are fully connected to extract hidden features $\bar{h}_t$ followed by a ReLU function. After feature extraction, h_t and $\bar{h}_t$ are obtained, each with a length corresponding to their respective hidden size. A fully connected layer FC is concatenated by copying the elements of two hidden arrays, with a total length of h_t and $\bar{h}_t$. Finally, y_{t+1} is fully connected with FC as a final output layer. By leveraging sequential and non-sequential inputs, the proposed method deeply exploited and utilised the available information to predict future values.

3.2.6 Long short-term memory

Long short-term memory (LSTM) is a type of RNN capable of learning long-term dependencies, adding gating mechanisms and memory cells that carry the relevant information from earlier steps to later steps (Hochreiter and Schmidhuber, 1997). In the architecture of LSTM in Figure 3.7, the input gate controls how much new information is stored in the cell, the forget gate controls how much old information is discarded from the cell, and the output gate controls how much information is passed to the next time step. These gates weigh the previous hidden state, the current input, inherent biases, and the current state of the memory cell, as described in Eqs. 3.15–3.18. Following these decisions, the cell state and hidden outputs are updated as per Eqs. 3.19–3.20. This system of gates and memory cells enables LSTMs to capture and retain long-term dependencies in sequences, making them particularly effective for tasks, like TBM tunnelling (Zhou et al., 2019; Qin et al., 2021; Lin et al., 2022b; Shen et al., 2022), landslides (Yang et al., 2019; Zhang et al., 2021c; Wang et al., 2024), and constitutive modelling (Qu et al., 2021; Zhang et al., 2021b; Xiong et al., 2023).

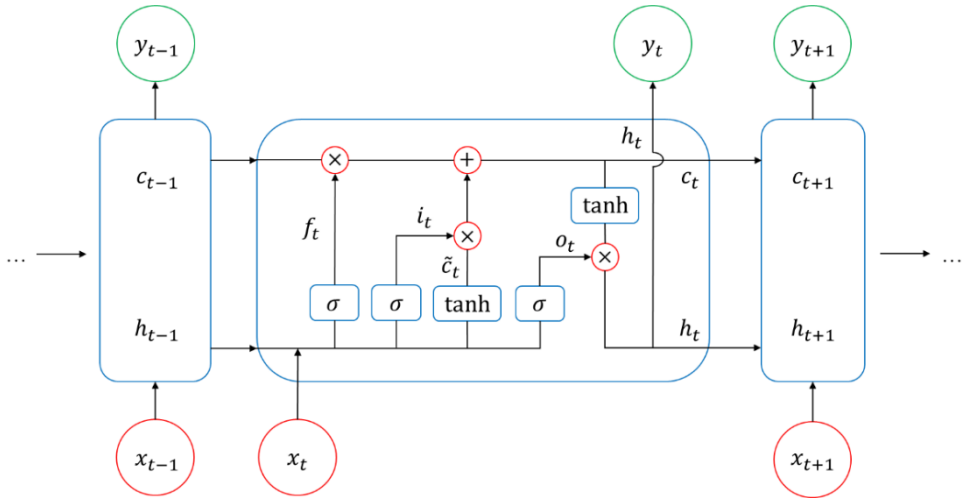


Figure 3.7 Schematic diagram of the long short-term memory

$$f_t = \sigma\left(U_f x_t + W_f h_{t-1} + b_f\right) \quad 3.15$$

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i) \quad 3.16$$

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o) \quad 3.17$$

$$\tilde{c}_t = \text{thah}(U_c x_t + W_c h_{t-1} + b_c) \quad 3.18$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad 3.19$$

$$h_t = \text{thah}(c_t) \circ o_t \quad 3.20$$

where f, i, o, c = Subscripts for forget gate, input gate, output gate, and memory cell

$\sigma(x) = \frac{1}{1+e^{-x}}$ = Sigmoid function

$\circ$ = Element-wise product

3.2.7 Self-training

Semi-supervised learning combines both labelled and unlabelled data during the training phase, aiming to leverage unlabelled data to improve model performance beyond what could be achieved with labelled data alone. The self-training method provides a

framework (Figure 3.8) that allows the adaptation of any supervised classification or regression approach, leveraging both labelled and unlabelled data.

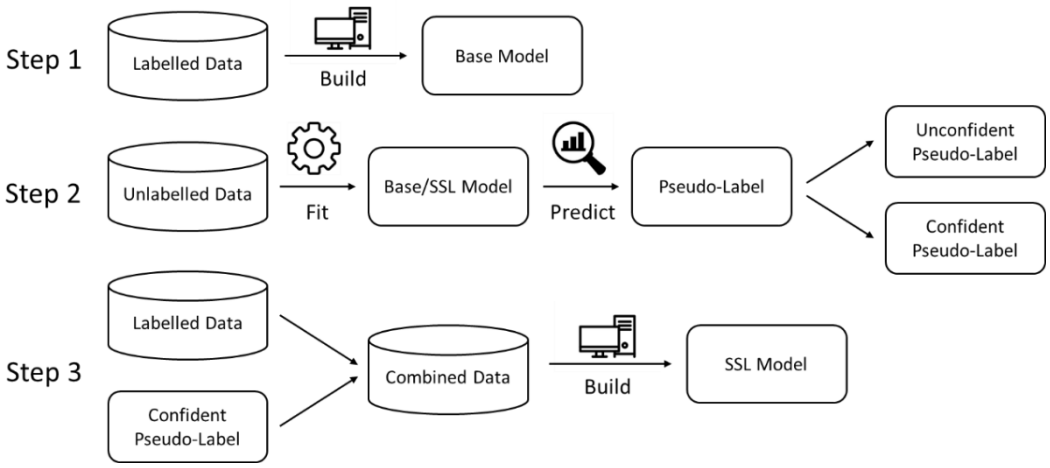


Figure 3.8 Workflow of semi-supervised self-training method

In the self-training, a base classifier is initially trained with a small amount of labelled data. The process commences by training a base classifier—be it SVM, RF, or ANN, contingent on the problem at hand—on a small amount of labelled data. Subsequently, pseudo-labels are generated on the unlabelled data. The confident predictions are combined with the labelled training data. With the combined dataset, an improved classifier undergoes retraining. The process can be iterated multiple times. In evaluation, the performance of the self-training classifier is assessed using labelled test data.

3.3 Evaluation metric

3.3.1 Regression metric

Model performance in testing is an indicator of the quality of the trained model. Eqs. 3.21–3.27 shows various evaluation metrics that quantitatively evaluate prediction errors. Mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) are dimensional and assess the errors between measured and predicted values, while mean absolute percentage error (MAPE) is non-dimensional and expressed as a percentage. The coefficient of determination (R^2) and variance account for (VAF) represent the proportion of the variance in the dependent values between 0 and 1, where a larger value indicates a higher accuracy between predicted and measured values, and vice versa. a_{10} index is an engineering index that shows the percentage of samples that fit the predicted values with a deviation of ±10% compared to measured values.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.21}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.22}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.23}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{3.24}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.25}$$

$$\mathrm{VAF} = 1 - \frac{\mathrm{Var}(y_i - \hat{y}_i)}{\mathrm{Var}(y_i)} \tag{3.26}$$

$$\mathrm{a}_{10} = \frac{n_{10}}{n} \tag{3.27}$$

where y_i = Actual values

$\hat{y}_i$ = Predicted values

$\bar{y}$ = Mean of actual values

n_{10} = Number of samples within ±10% deviation

3.3.2 Classification metric

Evaluation of the performance of a classification model is based on the counts of test records correctly and incorrectly predicted by the model. Confusion Matrix is a performance measurement for the machine learning classification problems in Figure 3.9. True positive (TP): We predicted positive, and it is true; True negative (TN): We predicted negative, and it is true; False positive (FP): We predicted positive, and it is false; False negative (FN): We predicted negative, and it is false.

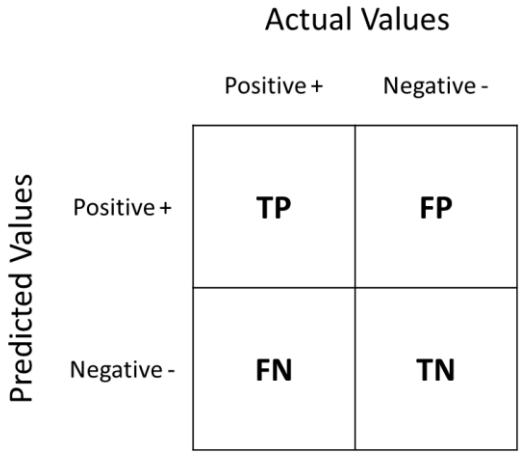


Figure 3.9 Confusion matrix table in classification

Precision (PRC) assesses how many correctly predicted cases are positive in Eq. 3.28. Recall (REC), or sensitivity, explains how many of the actual positive cases we could predict correctly in Eq. 3.29. F1-score (F1) is the harmonic mean of PRC and REC and balances them in Eq. 3.30.

$$\mathrm{PRC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{3.28}$$

$$\mathrm{REC} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3.29}$$

$$\mathrm{F1} = \frac{2 * \mathrm{PRC} * \mathrm{REC}}{\mathrm{PRC} + \mathrm{REC}} \tag{3.30}$$

3.4 Sensitivity analysis method

Machine learning models are highly nonlinear, so the relationship between the input and output is usually poorly understood. Sensitivity analysis in machine learning refers to determining how different input features influence the output of a trained model, such as partial dependence plots, Shapley additive explanations, and variance-based sensitivity analysis.

To analyse how much the variance of each feature affects the output variance, Sobol (1990) introduced the variance-based sensitivity analysis, also known as the Sobol method. The Sobol method decomposes the variance of the model output into fractions attributable to different input features, providing a comprehensive sensitivity measure accounting for main effects and all interaction effects.

A model can be decomposed using Eq. 3.31, and the variance of the output $\mathrm{Var}(y)$ can be decomposed using Eq. 3.32 under orthogonality constraints (Hoeffding, 1992).

$$y = f_0 + \sum_{i=1}^{p} f_i(x_i) + \sum_{i<j}^{p} f_{ij}(x_i, x_j) + \cdots + f_{1,2,\cdots,p}(x_1, x_2, \ldots, x_p) \quad 3.31$$

$$\mathrm{Var}(y) = \sum_{i=1}^{p} V_i + \sum_{i<j}^{p} V_{ij} + \cdots + V_{1,2,\cdots,p} \quad 3.32$$

where f_0 is a constant, f_i is a function of x_i, f_{ij} is a function of x_i and x_j, $V_i = \mathrm{Var}_{x_i}\left(\mathbb{E}_{x_{\sim i}}(y|x_i)\right)$ and $V_{ij} = \mathrm{Var}_{x_{ij}}\left(\mathbb{E}_{x_{\sim ij}}(y|x_i, x_j)\right) - V_i - V_j$ are expanded forms of expected values, and $x_{\sim i}$ indicates all parameters except x_i.

The direct effect of each feature is measured by the first-order Sobol index S_i in Eq. 3.33. The total-effect Sobol index S_{Ti} takes into account the sensitivity of the first-order effect and the sensitivity due to interactions between a given feature and all other features in Eq. 3.34. If a feature has a low Sobol index, then variations in the feature lead to comparatively slight variations in the model output, and vice versa (Fu and Zhang, 2021; Shan et al., 2023a).

$$S_i = \frac{V_i}{\mathrm{Var}(y)} \quad 3.33$$

$$S_{Ti} = 1 - \frac{\mathrm{Var}_{x_{\sim i}}\left(\mathbb{E}_{x_i}(y|x_{\sim i})\right)}{\mathrm{Var}(y)} \quad 3.34$$

where S_i = First-order Sobol index

S_{Ti} = Total-effect Sobol index

3.5 Summary

Chapter 3 delineates a systematic research methodology. Data processing becomes indispensable and involves a series of steps: deleting outliers to maintain data integrity, removing noise to improve clarity, selecting features to streamline the dataset, and augmenting minority class samples to ensure a balanced model training.

Regarding machine learning methods, models like SVM, RF, and neural networks (including ANN, RNN, Advanced RNN, and LSTM) are introduced, shedding light on their utility in the study. Self-training further augments the learning process, utilising both labelled and unlabelled data.

The research employs specific regression and classification metrics to gauge model performance, ensuring robust and reliable results. Sensitivity analysis methods determine the impact of individual input variables on the outcomes, affirming thoroughness.

Chapter 4. Datasets

Since machine learning models are data-driven, data quality (e.g., availability to the public, number of samples, and input parameters used) is crucial. This chapter summarizes four datasets in Table 4.1, accounting for data size and geological and operational parameters.

Table 4.1 Four tunnel projects and corresponding parameters

Project	Data size	Geological parameter	Operational parameter
Pahang-Selangor raw water tunnel	1,286	RQD, UCS, BTS, RMR, WZ	PR, TH, RPM
Changsha metro line	463	CD, WT, GC	PR, TH, TO, FP, RPM
Zhengzhou metro line	87	CD, WT, GC	PR, TH, TO, FP, RPM
Yinsong water diversion tunnel	12,926	HC, FZ	199 parameters (PR, TH, TO, RPM, PR_set, RPM_set)

4.1 Pahang-Selangor raw water tunnel

4.1.1 Project review

The raw water tunnel is located between the states of Pahang and Selangor in Malaysia. It facilitates water transfer from the Semantan River to the Selangor/Kuala Lumpur region for industrial and domestic demands. Abundant water traverses through the tunnel from the connecting basin to an outlet connecting the basin with gravity flow. The tunnel was excavated to cross the Main Range with an overburden of 100–1400 m. Figure 4.1 shows

a detailed geological profile across the tunnel from the inlet to the outlet.

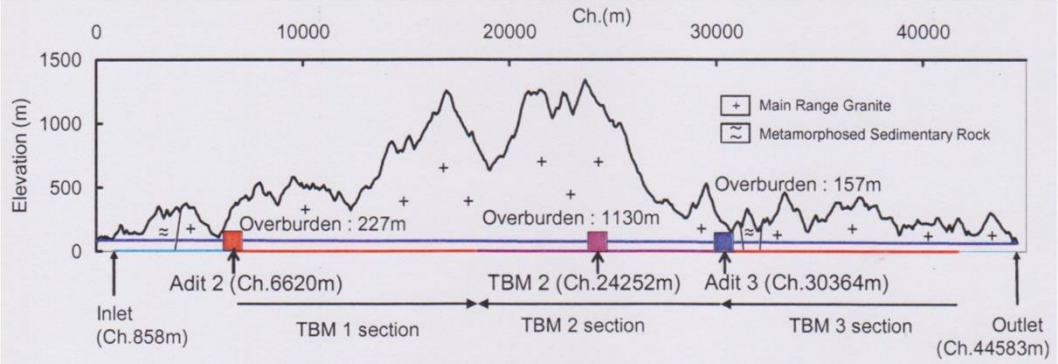


Figure 4.1 Geological profile of the Pahang-Selangor raw water tunnel (Armaghani et al., 2021)

The tunnel route can be dissected into four sections based on rock types. Section 1 (0.8–3.8 km) predominantly encompasses slightly-metamorphosed sedimentary rocks. It chiefly consists of black shale transitioning to schist, strongly folded due to the intrusion of granitic rocks. Coarse-grained granitic rocks dominate section 2 (3.8–12.5 km). Section 3 (12.5–27.0 km), forming a significant portion of the Main Range, is delineated by coarse to medium-grained granitic rock formations. Section 4 (27.0–44.6 km) is characterised by a deeply weathered continuum of the Main Range granite.

Regarding excavation methods, TBMs were slated for approximately 34.74 km of the main tunnel, as illustrated in Figure 4.1. Traditional drill and blast methods excavated the remaining segments. Three TBMs and four traditional drilling and blasting are incorporated to excavate in the project. Within the scope of TBM excavation, distinct ground conditions were addressed: TBM 1 excavated mixed ground conditions, TBM 2

excavated very hard ground, and TBM 3 excavated blocky terrains. The diameter of all TBMs is 5.23 metres, and their specifications are shown in Table 4.2.

Table 4.2 TBM specifications of the Pahang-Selangor raw water tunnel

TBM parameters	Value
TBM diameter	5.23 m
Weight of TBM	250 t
Number of cutters	35
Thrust force	14 MN
Maximum cutterhead torque	3.50 MN·m
Revolutions per minute	0–13.2 rev/min

4.1.2 Statistical analysis

The Pahang-Selangor raw water tunnel was investigated based on both site investigation and laboratory tests. A total of 1286 samples were extracted from a stretch of 12,649 m using three TBMs, each spanning approximately 10 m. In the samples, several rock mass properties were observed, such as strength, degree of weathering, joint conditions, and groundwater conditions. In addition, 154 block samples were collected from the cutterhead face and were conducted laboratory tests such as uniaxial compressive strength (UCS), Brazilian tensile strength (BTS), point load strength, Schmidt hammer, and P-wave velocity. Furthermore, the TBM acquisition system recorded operational parameters, including penetration rate (PR), thrust force (TH), cutterhead torque (TO), and revolutions per minute (RPM). It is noted that the mean values of these parameters were

computed for each sample.

Based on site investigation and laboratory tests, eight parameters were earmarked for analysis, including rock material properties (UCS, BTS), rock quality designation (RQD), rock mass rating (RMR), weathering zone (WZ), and operational parameters (PR, TH, and RPM). For the categorisation of the weathering zones, a rating system was adopted (Benardos and Kaliampakos, 2004). Zones were rated on a scale of 1 for fresh, 2 for slightly weathered, and 3 for moderately weathered zones. Basic statistical details of these parameters are presented in Table 4.3. Notably, TH represents the total thrust force in the cutterhead rather than the thrust force per cutter in the study of Armaghani et al. (2017).

Table 4.3 Basic statistical details in the Pahang-Selangor raw water tunnel

Parameters	**Unit**	**Min.**	**Max.**	**Ave.**
PR	mm/min	2.00	96.94	43.57
PRM	rev/min	4.08	11.95	8.84
TH	MN	2.82	19.80	11.25
UCS	MPa	40	185	107.45
BTS	MPa	4.69	15.68	8.43
RQD	%	6.25	95	44.15
RMR	-	44	95	64.73
WZ	-	1	3	1.70

4.2 Changsha and Zhengzhou metro lines

4.2.1 Project review

Changsha metro line 4, a rapid transit line located in Changsha, China, was excavated and opened in May 2019. The line spans approximately 33.5 km in the northwest-southeast direction, between Guanziling and Dujiaping. Five sections between six stations were investigated for this study: Liugoulong, Wangyuehu, Yingwanzhen, Hunan Normal University, Hunan University, and Fubuhe. The tunnel was excavated using an earth pressure balance (EPB) shield with a cutterhead diameter and length of 6.28 m and 8.74 m, respectively. The segmental lining was used to form a tube along the tunnel alignment to provide structural support. The segments were prefabricated in manufacturing plants with outer and inner diameters of 6 m and 5.4 m and a width of 1.5 m. The data acquisition system recorded operational parameters at one-minute intervals. Site investigation and experimental tests obtained geological conditions along the five sections. The machine excavated the tunnel in rocks such as slate, limestone, mudstone, and sandstone, and in soils such as silty clay, gravel, and marlite (Zhang et al., 2019a; Zhang et al., 2021a). Figure 4.2 shows the geological profile of Changsha metro lines.

Zhengzhou metro line 3 is a rapid transit line that runs from northwest to southeast in Zhengzhou, covering 24.5 km and 23 stations. The tunnel project was extracted by an EPB shield from December 2016 to December 2020. We investigated a section between

Jinshuilu and Taikanglu stations. The segment width in Zhengzhou is the same as in Changsha, but the outer and inner diameters are 3.1 m and 2.75 m, respectively. Unlike the geological conditions of the Changsha section, the Zhengzhou section was excavated in soil strata consisting of fine sand and silty clay, with top layers of backfills and silt in Figure 4.2.

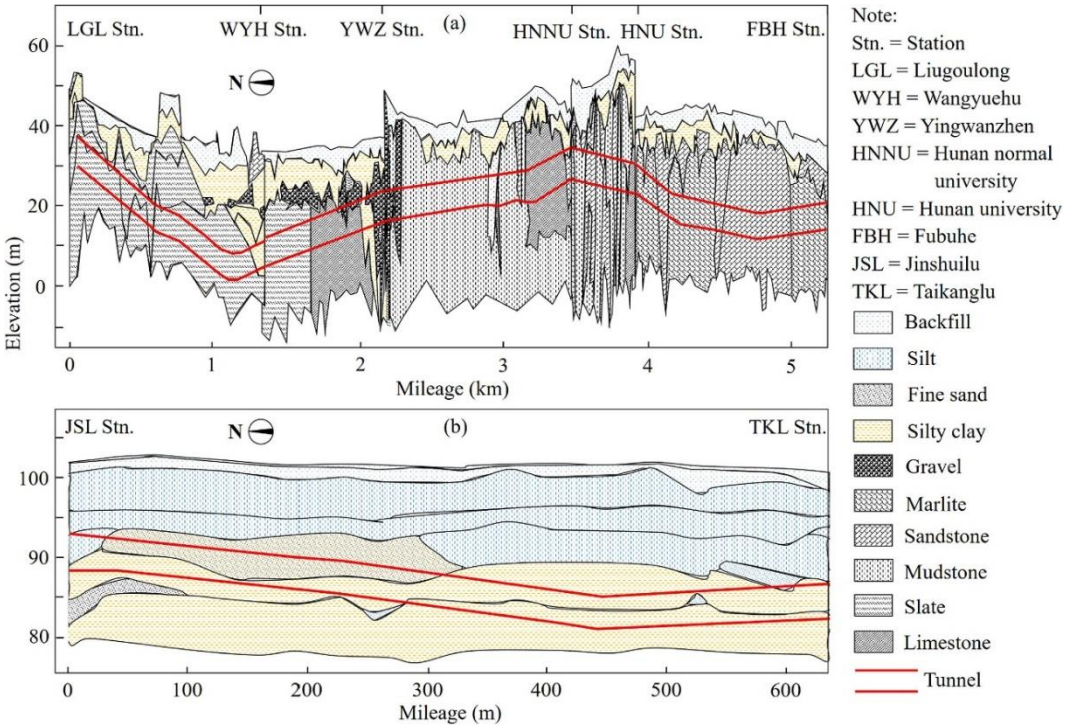


Figure 4.2 Geological profiles of the Changsha and Zhengzhou metro lines (Zhang et al., 2020b)

4.2.2 Statistical analysis

Operational parameters are collected by the data acquisition system at a high frequency every second or minute, while geological parameters are measured per segment at a low frequency. Zhang et al. (2020b) preprocessed the operational data to match the geological

data in five steps: (1) removal of empty data due to TBM maintenance, cutters change, breakdowns, or tunnel collapses; (2) removal of the first and last 2.5% of operational data at one segment; (3) detection and deletion of outliers based on the Mahalanobis distance; (4) the use of wavelet transform to eliminate noise in time series data; (5) averaging operational data at one segment as one sample.

Penetration rate (PR), thrust force (TH), cutterhead torque (TO), face pressure (FP), and revolutions per minute (RPM) are operational parameters reflecting TBM performance in tunnel construction. Regarding geological parameters, cover depth (CD) is the depth over the tunnel crown, water table (WT) is the tunnel depth below the water table, and ground condition (GC) at the tunnel face can be classified into four types: soil, gravel, rock and mixed-face ground, which are marked as 1, 2, 3 and 4, respectively. The basic statistical details of these parameters are presented in Table 4.4.

Table 4.4 Basic statistical details in Changsha and Zhengzhou metro lines

Parameters	**Unit**	**Min.**	**Max.**	**Ave.**
PR	m/h	0.24	4.56	2.1
TH	MN	4.08	23.8	11.56
TO	MN·m	0.63	4.82	2.53
FP	bar	0	2.1	0.95
RPM	rev/min	0.46	2.16	1.41
CD	m	9.09	31.65	17.86
WT	m	0	25.11	8.7
GC	-	1	4	2.69

In Figure 4.3, a time series of geological and operational parameters along the tunnel alignment is displayed, with blue lines presenting data from Changsha metro and orange lines presenting data from Zhengzhou metro. The horizontal coordinate includes 550 sequential data points, with each representing a segment of length 1.5 m. We assume that the two datasets are equally spaced in the sequence despite some samples being removed or missing. GC from Changsha is mainly rock at the tunnel face with a value of 3, while that from Zhengzhou is soil with a value of 1.

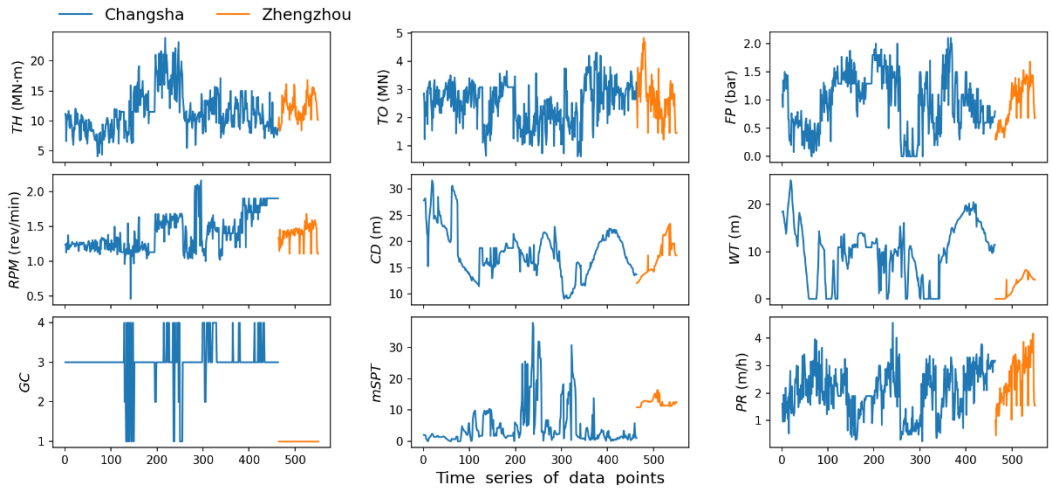


Figure 4.3 Time series of parameters from Changsha and Zhengzhou metro lines

4.3 Yinsong water diversion tunnel

4.3.1 Project review

The Yinsong water diversion project involved the excavation of a 69.86 km long tunnel to transfer water from the Fengman Reservoir to Changchun and its surrounding areas. The excavation was mainly carried out using TBMs, with supplemental drilling and

blasting. In the study, the TBM3 section was excavated using an open-type TBM from July 2015 to February 2018, as specified in Table 4.5.

The TBM excavated a total distance of 17.50 km, with an overburden thickness ranging from 85 to 260 meters. Chen et al. (2021) provided a longitudinal geological profile in Figure 4.4, including limestone, diorite, tuff, and granite. The rock mass class is categorised into five classes of hydropower classification (HC) based on the Standard for Engineering Classification of Rock Mass (GBT 50218-2014). Classes II, III, IV, and V are involved in the TBM section 3. A total of 33 faults were investigated, with exposure length ranging from 1 to 19 metres. In addition, 18 tunnel collapses happened during TBM tunnelling.

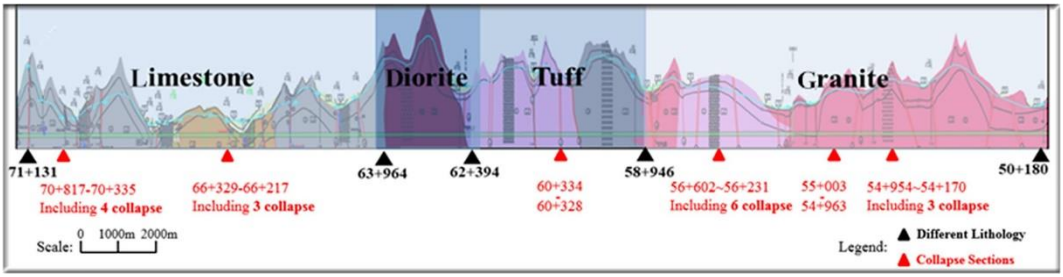


Figure 4.4 Geological profile of the Yinsong water diversion tunnel (Chen et al., 2021)

Table 4.5 TBM specifications of the Yinsong water diversion project

TBM parameters	**Value**
TBM diameter	7.93 m
Weight of TBM	180 t
Number of cutters	56
Maximum thrust force	23.26 MN

Cutterhead torque	8.41 MN·m @ 3.97 rev/min
Revolutions per minute	0–7.6 rev/min

4.3.2 Data preprocessing

In machine learning, the extraction of condensed and clean feature vectors as inputs and outputs holds significant importance. The acquisition system automatically recorded 199 parameters, archiving them daily in an Excel spreadsheet over 728 days in this study. The chronological and geometrical aspects of the data are characterised by time and chainage, respectively, with time recorded at one-second intervals and chainage measured in metres. The preferred method of data processing hinges on boring cycles, which are categorised into four stages, as delineated in Figure 4.5(a).

- Start-up stage (S1): the TBM initiates to ramp up. Revolutions per minute (RPM) attain stability, but the TBM does not commence cutting rock.
- Ascending stage (S2): the TBM begins to cut rock. Operational parameters of penetration rate (PR), thrust force (TH), and cutterhead torque (TO) increase before achieving dynamic stability.
- Steady-state stage (S3): the TBM cuts rock at a constant speed, and four operational parameters remain dynamic stability.
- End stage (S4): the TBM gradually halts with four operational parameters decreasing to zero.

TBM operators can adjust the resistance of potentiometers from 0 to 10,000 Ω, regulating the penetration rate (PR_set) and revolutions per minute (RPM_set), as illustrated in

Figure 4.5(b). PR_set and RPM_set are predefined prior to the subsequent boring cycle based on sense experience and sparse geological conditions. It is crucial to note that these are the values the system displays and not the target values on the control panel. We assume that the setting values align with the target values during the steady-state stage, suggesting that the target setting values roughly equal the average values during this stage.

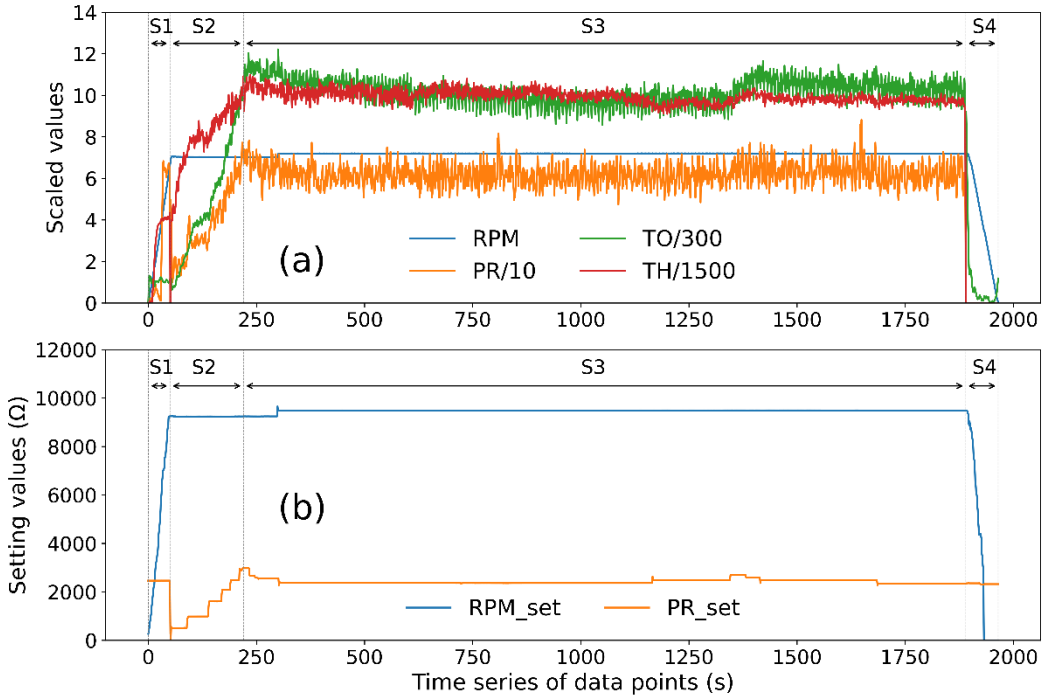


Figure 4.5 Typical boring cycle in (a) operational parameters and (b) setting values

However, errors that suddenly exceed one or more orders of magnitude can be caused by machine failure or operator mistakes. An example of such errors is evident in the working cycle in Figure 4.6(a). An error is an unusual data point outside of the data used to build models. In a standard normal distribution, the 3-sigma rule has been extensively used to remove errors (Gao et al., 2020; Feng et al., 2021; Xu et al., 2021), stating that 99.73%

of data points fall within the range of the mean plus or minus three standard deviations. However, errors can distort the mean and standard deviation but have minimal impact on the median. We adopted an empirical approach and removed any errors exceeding 200 mm/min. As a result, Figure 4.6(b) displays the boring cycle without errors.

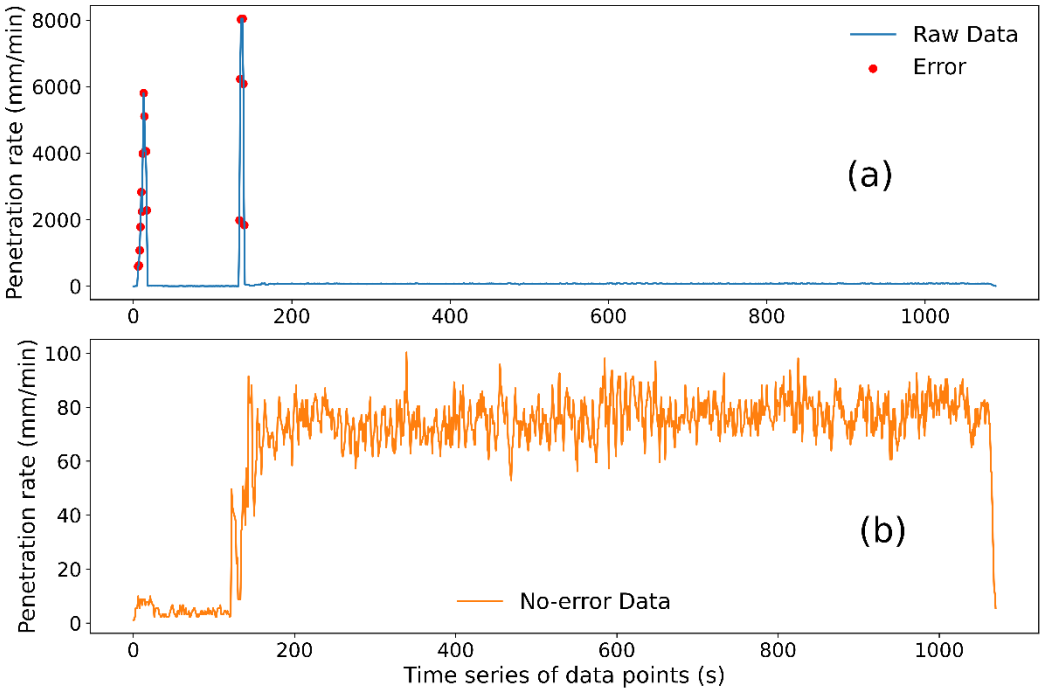


Figure 4.6 Specific boring cycle (a) with errors and (b) removing errors

This study focuses on TBM performance in the steady-state stage, as it most accurately represents the actual excavation efficiency. This stage entails the bulk of the excavation work and is unaffected by the duration of the boring process. However, identifying the turning point between the ascending and steady-state stages poses a challenge. Xu et al. (2021) offered an empirical approximation of the ascending stage's duration at around 2 minutes. To remove fine vibrations, a moving average method creates a smoothed time

series by averaging observations within a sliding window of 20 (Shan et al., 2022), denoted as PR_20 in Figure 4.7. A threshold is then established equal to 0.8 times the median penetration rate. The start and end points are subsequently identified at the intersections between the smoothed time series and the threshold. As a result, we selected 12962 boring cycles with more than 300 data points during the steady-state stage, averaging a span of 1.14 metres. Finally, we computed the average values within the steady-state stage to generate a sample, thus converting high-frequency data into low-frequency data.

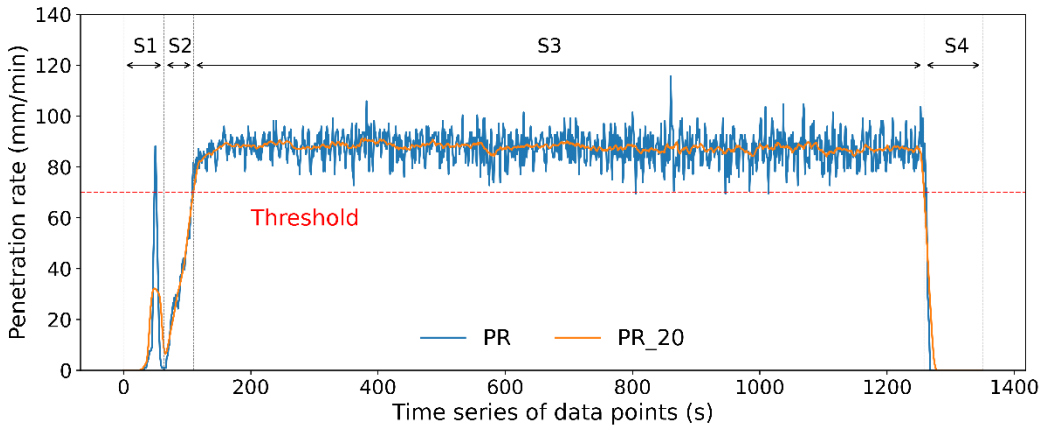


Figure 4.7 Splitting the steady-state stage in statistics

4.3.3 Statistical analysis

Many parameters are irrelevant to TBM performance, such as temperature, concentration of O_2, warning and alarm values. Li et al. (2021) divided them into parameters with constant, low-variance, and high-correlated values, and the accuracy increased as the

number of input features dropped from 154 to 51. Typically, in TBM performance modelling, PR, RPM, TO, and TH are regarded as feature data. These four parameters were the only ones used as feature vectors in part of the studies (Chen et al., 2021; Feng et al., 2021; Liu et al., 2021a). Apart from these four parameters, many techniques were employed to find possibly applicable parameters, such as the Pearson correlation coefficient (Liu et al., 2021b), variance important measure in the random forest (Li et al., 2021; Guo et al., 2022; Hou et al., 2022). Among the 199 parameters, there are no generally recognised standards for selecting input parameters.

During TBM construction, TBM operators drive the machine by setting PR_set and RPM_set on the control panel. Figure 4.8(a) reveals a strong correlation of 0.832 between PR and PR_set, indicating that the actual penetration rates are influenced by setting values and geological conditions. Figure 4.8(b) demonstrates a linear relationship between RPM and RPM_set, indicating that the setting values of revolutions are promptly reflected in the actual revolutions.

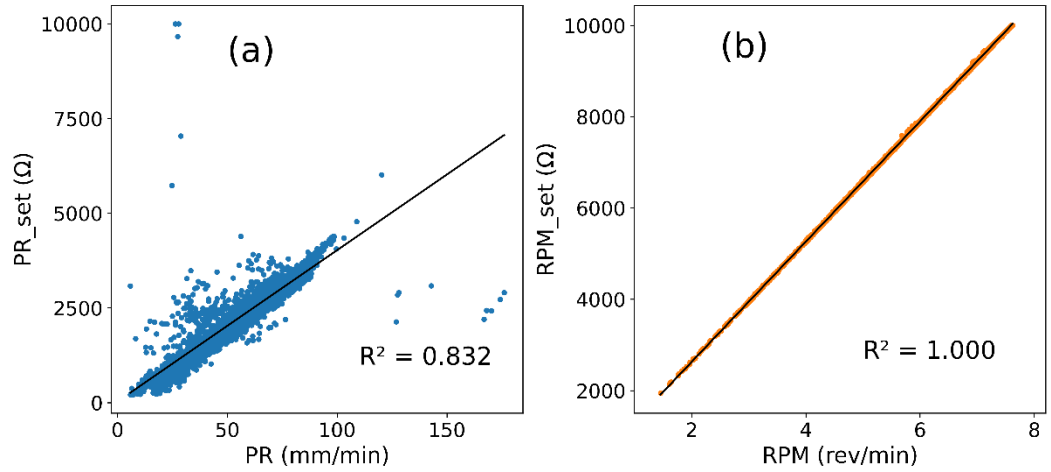


Figure 4.8 Correlations between (a) PR and PR_set, (b) RPM and RPM_set

Regarding geological conditions, hydropower classification (HC) and fault zone (FZ) are measured during the site investigation. Table 4.6 describes HC for classes I through V based on rock solidness and rock mass integrity. A quantitative HC index is devised by evaluating the uniaxial compressive strength and integrity index. In underground engineering, this assessment is further refined based on groundwater, structural plane occurrence, and initial in-situ stress in Eq. 4.1. The fault zone (FZ) was categorized using a binary classification method: a value of 1 was assigned to indicate the presence of a fault zone, and a value of 0 to indicate its absence, without considering the nature of the fault fillings. Geological data is fitted according to the chainage along the tunnel alignment to balance the sampling of operational data.

$$\text{HC index} = 100 + 3\text{UCS} + 250K_v - 100(K_1 + K_2 + K_3) \qquad 4.1$$

where UCS = Uniaxial compressive strength

K_v = Integrity index

K_1 = Groundwater corrective coefficient

K_2 = Structural plane occurrence corrective coefficient

K_3 = Initial in-situ stress corrective coefficient

Table 4.6 Qualitative and quantitative description of hydropower classification

Class	Quantitative description	Qualitative description	
	HC index	Rock solidness	Rock mass integrity
I	>550	Hard	Intact
II	451–550	Hard	Less intact
		Less hard	Intact
III	351–450	Hard	Less fractured
		Less hard	Less intact
		Less soft	Intact
IV	251–350	Hard	Fractured
		Less hard	Less fractured, fractured
		Less soft	Less intact, less fractured
		Soft	Intact, less intact
V	≤250	Less soft	Fractured
		Soft	less fractured, fractured
		Very soft	Very fractured

The feature vectors consist of four operational parameters of PR, RPM, TO, and TH, two geological parameters of HC and FZ, and two setting values of PR_set and RPM_set, summarizing statistical details in Table 4.7.

Table 4.7 Basic statistical details in the Yinsong water diversion tunnel

Parameter	Unit	Min.	Max.	Ave.
PR	mm/min	5.80	176.08	62.95
RPM	rev/min	1.456	7.621	6.287

TO	MN·m	0.118	3.876	2.288
TH	MN	2.763	19.81	12.406
HC	-	2	5	3.32
FZ	-	0	1	0.010
PR_set	Ω	212	10000	2547
PRM_set	Ω	1944	10000	8278

4.4 Summary

This chapter summarizes four TBM datasets—the Pahang-Selangor raw water tunnel, Changsha metro line, Zhengzhou metro line, and Yinsong water diversion tunnel. The Pahang-Selangor raw water tunnel in Malaysia facilitates water transfer, excavated primarily using TBMs. A thorough statistical study was performed on 1286 samples from the tunnel, revealing insights into rock mass properties and operational parameters. Eight significant parameters are analysed: UCS, BTS, RQD, RMR, WZ, PR, TH, and RPM.

Changsha and Zhengzhou metro lines are rapid transit lines in China, each excavated using an EPB shield, differing in geological conditions and tunnel dimensions. The Changsha line traverses rock and soil formations, whereas the Zhengzhou line mainly cuts through soil strata. For statistical analysis, operational parameters were frequently collected, while geological metrics were taken per segment. Critical parameters like PR,

TH, and GC were highlighted, with distinct geological attributes observed between the two metro lines.

The Yinsong water diversion project entailed a 69.86 km excavation to transfer water, with the TBM3 section, spanning 17.50 km, excavated by an open-type TBM. In the acquisition system, 199 parameters were automatically recorded daily over 728 days. The focus of the study lies in the steady-state stage of the boring cycle, with noise and errors removed, converting high-frequency data into low-frequency for analysis. Among many parameters, PR, RPM, TO, and TH are the most relevant for TBM performance. PR_set and RPM_set predetermined by operators match target values during the steady-state stage. Geological data, like HC and FZ, were combined with operational data to create feature vectors.

Chapter 5. Penetration Rate Regression

5.1 Introduction

The Pahang-Selangor raw water tunnel transfers water between the states of Pahang and Selangor in Malaysia. The main tunnel, appropriately 34.74 km, was extracted by three TBMs. A total of 1286 samples are collected, each spanning a length of approximately 10 m. The penetration rate regression is crucial in tunnelling operations as it directly affects overall productivity. A higher penetration rate results in faster tunnel excavation, ultimately reducing project time and costs.

Operational parameters, such as RPM and TH, are not accessible during training. While models reliant on operational and geological parameters might achieve high accuracy, they are impractical for real-world application. Given that geological conditions are determined via site investigations and are available before tunnelling commences, this Chapter proposes three machine learning models based on geological parameters (UCS, BTS, RQD, RMR, and WZ), expressed as Eq. 5.1.

$$\mathrm{PR} = f(\mathrm{UCS}, \mathrm{BTS}, \mathrm{RQD}, \mathrm{RMR}, \mathrm{WZ}) \qquad 5.1$$

Table 5.1 presents a comparison of the SVM, RF, and ANN models (Models 5.1–5.3) and analyses the difference in input parameters between Model 5.2 and Models 5.4–5.5.

Table 5.1 Model performance for penetration rate regression

Model	Output	Input	Method	RMSE	R^2
5.1	PR	UCS, BTS, RQD, RMR, WZ	SVM	4.03	0.876
5.2	PR		RF	3.49	0.905
5.3	PR		ANN	4.03	0.878
5.4	PR	BTS, RMR, WZ	RF	4.61	0.831
5.5	PR	UCS, BTS, RQD, WZ	RF	4.29	0.854

5.2 Data Processing

Outliers are data points that significantly differ from other observations in a dataset. Often perceived as errors, these outliers are generally recommended for removal. The need for outlier detection should arise from input and output data, depending on the nature of the data and the domain of problems. Outliers in the input data can emerge from discrepancies like data entry or measurement errors. Similarly, unusual target values can also appear as outliers.

Isolation forest (IForest) is an unsupervised algorithm rooted in decision tree techniques. Distinct from traditional outlier detection methods, IForest employs binary trees to isolate outliers directly without normal instance profiling. Within the Pahang-Selangor raw water tunnel, a matrix comprising 1286 rows and 6 columns—which donates 1286 samples, five geological parameters, and one target value of PR—is processed through the IForest model to detect outliers. Figure 5.1 shows a 3D scatter plot that distinguishes outliers from normal data. The Figure is anchored on three axes—PR, BTS, and RMR—with

RMR subsuming the information of RQD and UCS. Within the configured model, 1% of all samples are designated as outliers located significantly from other normal data. As a result, 1273 normal data are leveraged to train regression models for PR prediction.

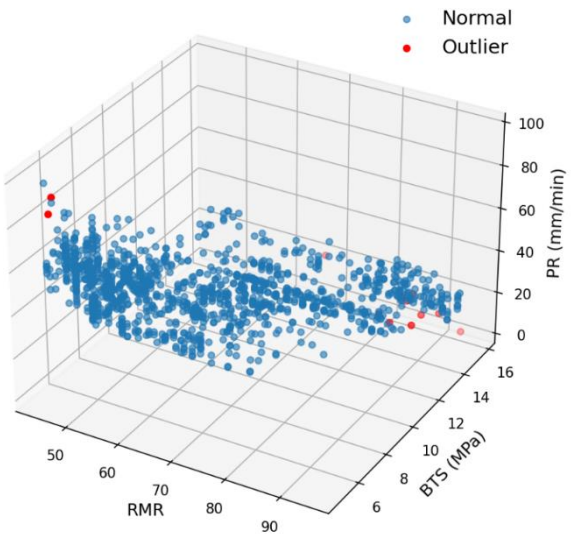
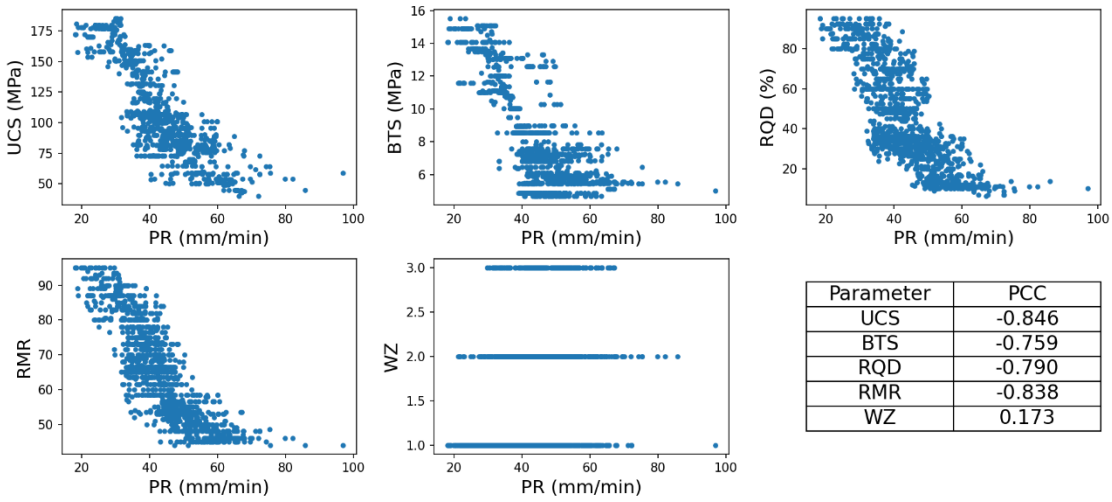


Figure 5.1 Scatter plot with normal data and outliers

The Pearson correlation coefficient (PCC) serves as a statistical measure to quantify the strength and direction of the linear relationship between variables. As illustrated in Figure 5.2, a notable negative correlation is evident between PR and the variables UCS (-0.846), BTS (-0.759), RQD (-0.790), and RMR (-0.838). A PCC value between -0.85 and -0.75 indicates a robust linear relationship. Conversely, the PCC for WZ is relatively low at 0.173, suggesting a poor linear relationship. This low correlation is likely due to the coarse categorisation employed for the weathering zones.



Parameter	PCC
UCS	-0.846
BTS	-0.759
RQD	-0.790
RMR	-0.838
WZ	0.173

Figure 5.2 Scatter plots between PR and operational parameters and their PCC

It is imperative to scale input and output data to a standard, dimensionless scale because they have different units and magnitudes. It is achieved through min-max normalisation, which linearly transforms the data in Eq. 5.2, ensuring values fall within the 0 to 1 range.

$$x_{nor} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad 5.2$$

where x_{nor} = Normalised data

x_{max} = Maximum in the data

x_{min} = Minimum in the data

During the training process, min-max normalisation is applied to the training data, and subsequently, identical normalisation parameters are applied to the test data. The normalisation not only ensures consistent scaling of features but also prevents dominance of particular features. It facilitates the efficiency of optimisation algorithms and improves the accuracy of machine learning models. Importantly, an inverse transformation is

executed to predictions on the test data, reverting them to their original magnitude for practical interpretability.

5.3 Modelling process

Figure 5.3 depicts a basic flowchart for building a near-optimal model using machine learning. Prior to modelling, a dataset is processed to select relevant data by outlier detection, interpolation, data smoothing, and feature selection. The processed data are randomly split into training, validation, and test sets. The training and validation set is used to train the model, while the test set is used to evaluate its performance. The choice of algorithm is crucial as it can significantly impact the model's accuracy and reliability. Hyperparameter tuning aims to find the best hyperparameters, which helps to fine-tune the model performance. As a result, a near-optimal model is built and evaluated in the test set. Notably, the modelling process is also employed in Chapters 6–8.

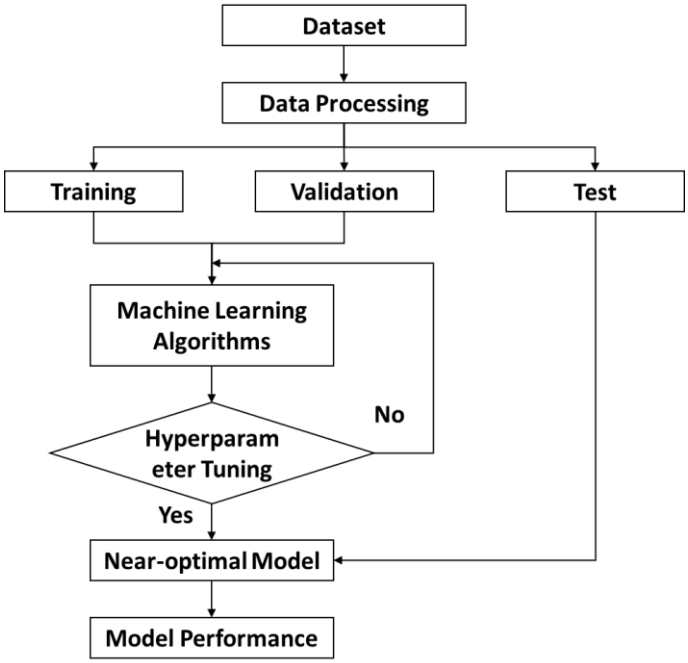


Figure 5.3 Typical flowchart of the modelling process in machine learning

5.4 Machine learning model

In machine learning, a common practice is to allocate 80% of the dataset for training, reserving the remaining 20% for testing on unseen data. Cross-validation is employed as a splitting strategy to mitigate the risk of overfitting. Specifically, a 5-fold cross-validation is utilised: the training dataset is split into five subsets. The model undergoes training on four subsets and validation on the fifth. This process is carried out five times, guaranteeing that each subset is served once for validation.

Taking an ANN model as an example, several hyperparameters influence its performance.

- Number of hidden layers [1, 2, 3];

- Learning rate [0.001, 0.005, 0.01, 0.05];
- Hidden size [8, 16, 32, 64, 128, 256];
- L2 regularisation [0.0001, 0.001, 0.01];
- Batch size [32, 64, 128].

Grid search exhaustively searches the predefined space to find the best combination of hyperparameters, leading to 648 trials (calculated as 3*4*6*3*3). Alternatively, random search is computationally efficient and focuses on random sampling combinations. In this thesis, 60 random trials are conducted, providing a 95% probability of finding a combination within the top 5% of optimal configurations (Bergstra and Bengio, 2012).

5.4.1 Support vector machine model

Support vector machines (SVM) represent a class of machine learning algorithms adept at managing high-dimensional datasets and yielding accurate results, even when data availability is limited. For achieving a near-optimal SVM model, hyperparameter tunning encompasses several domains:

- Regularisation parameter [0.1, 1, 10, 100];
- Degree of misclassification [0.01, 0.1, 1];
- Kernel function [‘linear kernel’, ‘radial basis function kernel’, ‘polynomial kernel’];
- Degree of the polynomial kernel function [2, 3, 4] (Note: This is exclusive to the polynomial kernel and is not considered for the other kernels).

The near-optimal SVM model is built based on the best hyperparameters, involving a regularisation parameter of 0.1, a misclassification degree of 0.01, and a radial basis function kernel. Predicted results are visually represented in Figures 5.4(a) and 5.4(b) for training and test data, respectively. The model performance is further assessed by seven regression metrics—MAE, MSE, RMSE, MAPE, R^2, VAF, and a_{10}—detailed in Table 5.2.

Within the 80% training data, measured and predicted PR yields metrics as follows: MAE of 2.14, MSE of 9.81, RMSE of 3.13, MAPE of 4.91%, R^2 of 0.909, VAF of 0.908, and a_{10} of 0.864. For the left 20% test data, measured and predicted PR are evaluated by MAE of 2.52, MSE of 16.27, RMSE of 4.03, MAPE of 5.69%, R^2 of 0.876, VAF of 0.872, and a_{10} of 0.827. Commonly, training results are slightly better than test results but the ultimate goal is to build models with robust generalisation capabilities for new, unseen data.

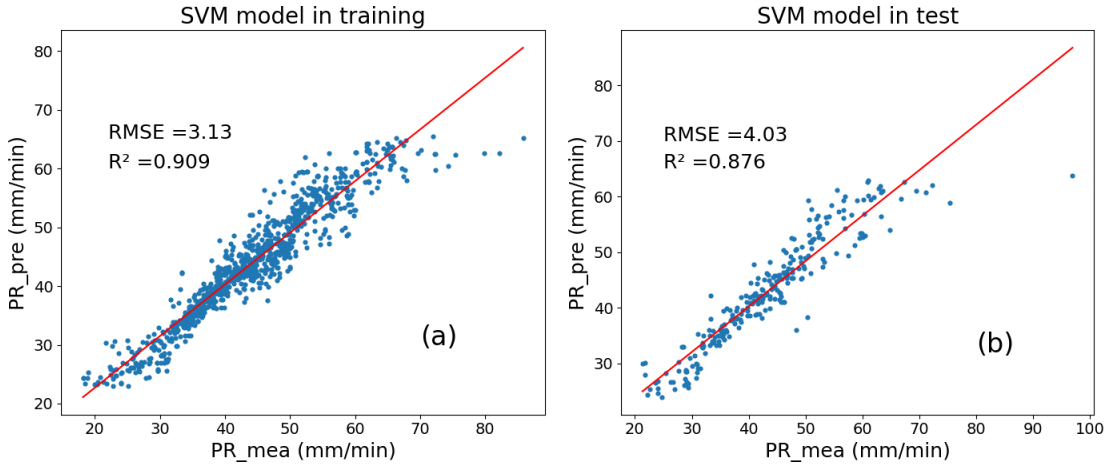


Figure 5.4 SVM model performance (a) in training and (b) in test

Table 5.2 Evaluation metrics in the SVM model for training and test data

	MAE	**MSE**	**RMSE**	**MAPE**	**R^2**	**VAF**	**a_{10}**
Training	2.14	9.81	3.13	4.91%	0.909	0.908	0.864
Test	2.52	16.27	4.03	5.69%	0.876	0.872	0.827

5.4.2 Random forest model

Random forest (RF) is an ensemble learning algorithm combining predictions from multiple decision trees trained on various subsamples of the dataset. The criteria are the mean squared error (MSE) for variance reduction. The optimisation of the RF model entails hyperparameter tuning within predefined domains:

- Number of trees [10, 20, 50, 100, 200];
- Maximum depth of the tree [None, 10, 20, 30, 50];
- Minimum number of samples required to split an internal node [2, 5];
- Minimum number of samples required to be at a leaf node [1, 2].

The near-optimal RF model, predicated on the above hyperparameters, constitutes 200 trees, a maximum tree depth of 20, a bifurcation criterion of at least two samples for internal nodes, and a single sample minimum at leaf nodes. The predictions are represented in 2-D scatter figures between measured PR and predicted PR in Figures 5.5(a) and 5.5(b) for training and test data, respectively. An ensemble of seven regression metrics, namely MAE, MSE, RMSE, MAPE, R^2, VAF, and a_{10}, is used to evaluate the model performance in Table 5.3.

In training data, the alignment between measured and predicted PR is perfectly fitted, yielding metrics as follows: MAE (0.78), MSE (1.80), RMSE (1.34), MAPE (1.76%), R^2 (0.983), VAF (0.983) and a_{10} (0.991). In contrast to test data, these metrics are delineated as MAE (2.15), MSE (12.18), RMSE (3.49), MAPE (4.84%), R^2 (0.905), VAF (0.903) and a_{10} (0.859).

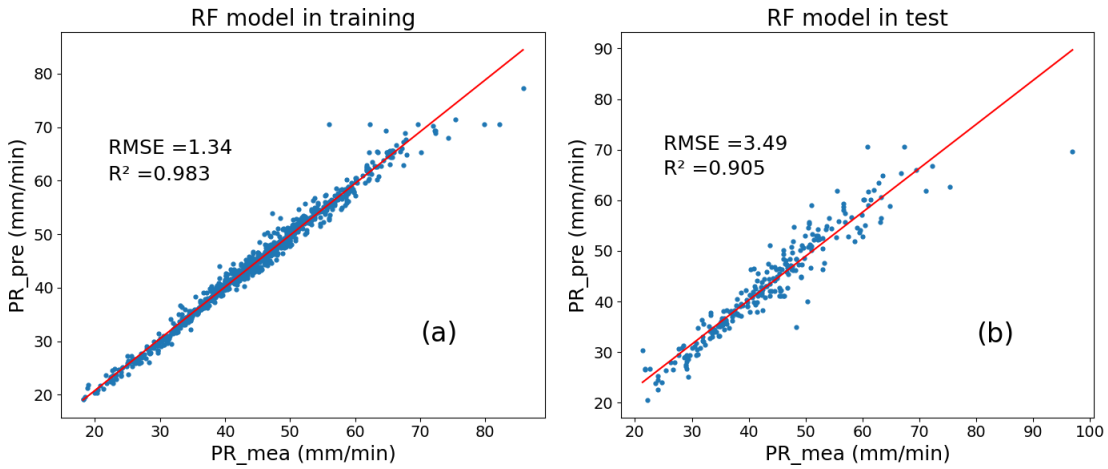


Figure 5.5 RF model performance (a) in training and (b) in test

Table 5.3 Evaluation metrics in RF model for training and test data

	MAE	**MSE**	**RMSE**	**MAPE**	**R^2**	**VAF**	**a_{10}**
Training	0.78	1.80	1.34	1.76%	0.983	0.983	0.991
Test	2.15	12.18	3.49	4.84%	0.905	0.903	0.859

5.4.3 Artificial neural network model

Artificial neural networks (ANNs) represent a class of deep learning algorithms structurally characterised by an input layer, multiple hidden layers, and an output layer. The adaptive moment estimation (Adam) optimiser, renowned for its capability to handle

large datasets and adaptively adjust learning rates, stands out as a predominant optimisation algorithm in deep learning models. The MSE loss function, expressed in Eq. 3.22, quantifies the difference between the measured and predicted results.

A comprehensive discussion on the scope of hyperparameters is provided in the front of Section 5.4. Employing the optimal hyperparameters—comprising two hidden layers with sizes of 256 and 16, respectively, a learning rate set at 0.005, an L2 regularisation parameter of 0.001, and a batch size of 128—a near-optimal ANN model is constructed. The model performance and predictions are illustrated in Figure 5.6(a) for training data and Figure 5.6(b) for test data, delineating 2-D scatter plots between the measured PR and its predicted counterpart. A combination of regression metrics—specifically, MAE, MSE, RMSE, and MAPE (for error evaluations), and R^2, VAF, and a_{10} (for variance evaluations)—facilitates a rigorous model performance evaluation, as tabulated in Table 5.4.

Within the 80% of data designated for training, there is a good congruence between measured and predicted PR, manifesting in metrics such as MAE of 2.48, MSE of 11.27, RMSE of 3.36, MAPE of 6.06%, R^2 of 0.906, VAF of 0.905, and a_{10} of 0.849. Conversely, for the remaining 20% of test data, measured and predicted PR are evaluated by MAE of 2.81, MSE of 16.27, RMSE of 4.03, MAPE of 6.66%, R^2 of 0.878, VAF of 0.876, and a_{10} of 0.827.

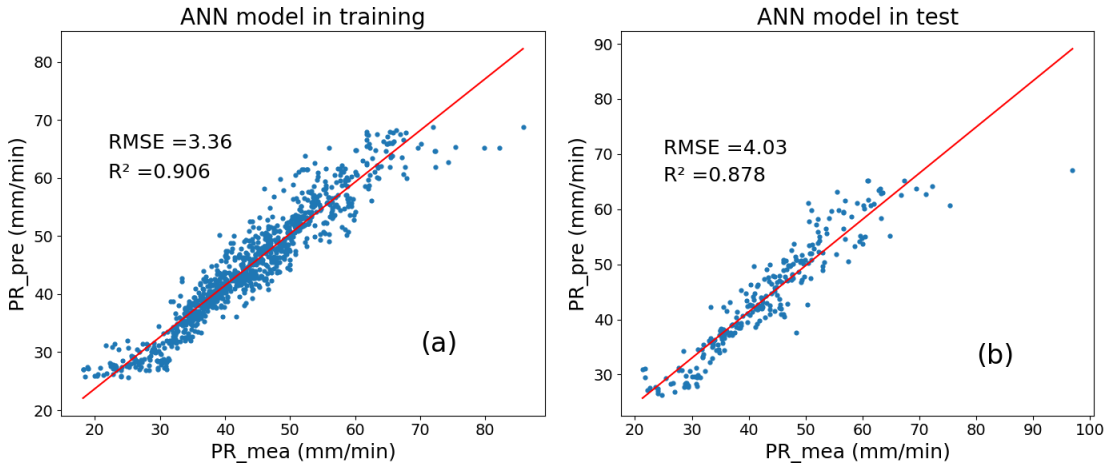


Figure 5.6 ANN model performance (a) in training and (b) in test

Table 5.4 Evaluation metrics in ANN model for training and test data

	MAE	**MSE**	**RMSE**	**MAPE**	**R^2**	**VAF**	**a_{10}**
Training	2.48	11.27	3.36	6.06%	0.906	0.905	0.849
Test	2.81	16.27	4.03	6.66%	0.878	0.876	0.827

5.5 Comparison between proposed models

5.5.1 Models on machine learning algorithms

The near-optimal SVM, RF, and ANN models are built after hyperparameter tuning to predict penetration rates in TBM tunnelling. In comparison, Figure 5.7 shows their model performance in terms of seven regression metrics for training and test data.

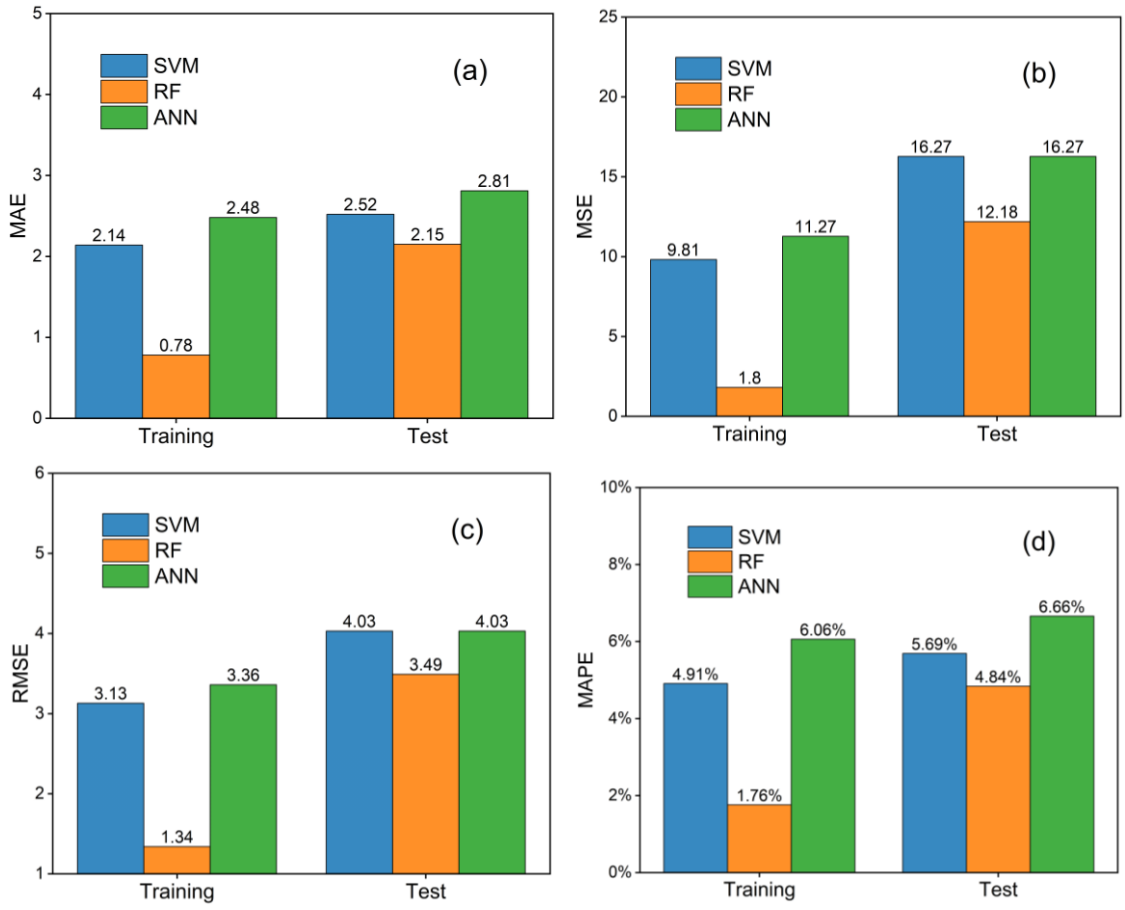
Metrics derived from training data indicate the adeptness with which machine learning models align with training data. Metrics derived from test data indicate the model generalisation to new, unseen data. The training results for the SVM, RF, and ANN

models are slightly better than the test results. Overfitting occurs when training results are perfectly good but test results perform poorly. Overfitting often results from excessive feature dimensions, overly complex structures, an abundance of noise, or insufficient training data. Consequently, the emphasis in model assessment normally focuses on test metrics, a perspective employed in Chapters 6–8.

Metrics like MAE, MSE, RMSE, and MAPE offer insights into model performance from an error perspective, while R^2 and VAF are pivotal for variance evaluations. Among the models, the RF model has the most minimal error metrics in testing (MAE = 2.15, MSE = 12.18, RMSE = 3.49, MAPE = 4.84%) and has the highest variance metrics (R^2 = 0.905, VAF = 0.903, a_{10} = 0.859). Comparatively, the SVM model manifests lower errors than the ANN model in MAE (2.52 < 2.81) and MAPE (5.69% < 6.66%) but demonstrates parity in MSE of 16.27, RMSE of 4.03, and a_{10} of 0.827. Furthermore, the SVM model exhibits lower metrics of R^2 (0.876 < 0.878) and VAF (0.872 < 0.876) than the ANN model. Given the nuanced disparities across these seven regression metrics, the SVM and ANN models perform equivalently. Navigating through seven metrics for model comparison can be time-consuming. Hence, in forthcoming Chapters, we will predominantly employ RMSE or MAPE to represent errors and R^2 for variance characteristics.

It is commonly believed that the ANN model can capture more underlying information

than the SVM and RF models. However, the RF model outperforms SVM and ANN models in penetration rate regression. Penetration rate regression is not a very complex problem. The ANN model is likely too complex to find the causal relationship in penetration rate regression, while the SVM model is probably too simple to capture information. Alternatively, the RF model is relatively suitable for accurately predicting penetration rate.



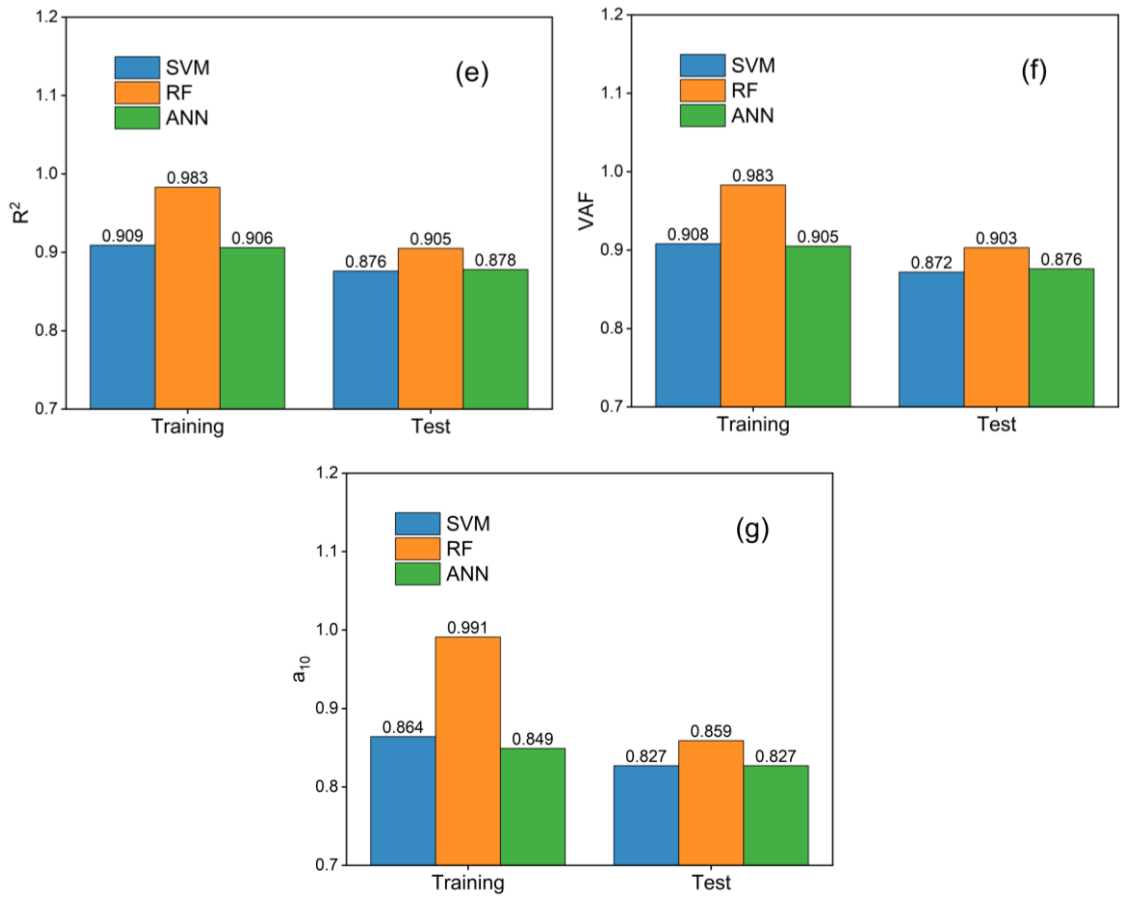


Figure 5.7 Comparing model performance on machine learning algorithms in terms of (a) MAE, (b) MSE, (c) RMSE, (d) MAPE, (e) R^2, (f) VAF, and (g) a_{10}

5.5.2 Models on input parameters

Six geological parameters are used to classify a rock mass in the RMR system, including UCS, RQD, spacing of discontinuities, condition of discontinuities, orientation of discontinuities, and groundwater conditions. It is found an overlap between RMR and its internal parameters (UCS and RQD) in Eq. 5.1. Models 5.2 and 5.4–5.5 use the same RF model and compare different input parameters. In Figure 5.8, Model 5.4 removes the internal parameters of UCS and RQD, leading to RMSE of 4.59 and R^2 of 0.833. Model

5.5 shows RMSE of 4.28 and R^2 of 0.854 when removing RMR. As a result, Model 5.2, with 5 geological parameters, performs the best in predicting penetration rate with the lowest RMSE of 3.49 and the highest R^2 of 0.905. In other words, weights of removed parameters in Models 5.4–5.5 are manually set as 0, while weights in Model 5.2 are automatically learned by the RF model. Although UCS and RQD are sub-parameters of RMR, Model 5.2 with all geological parameters is the best.

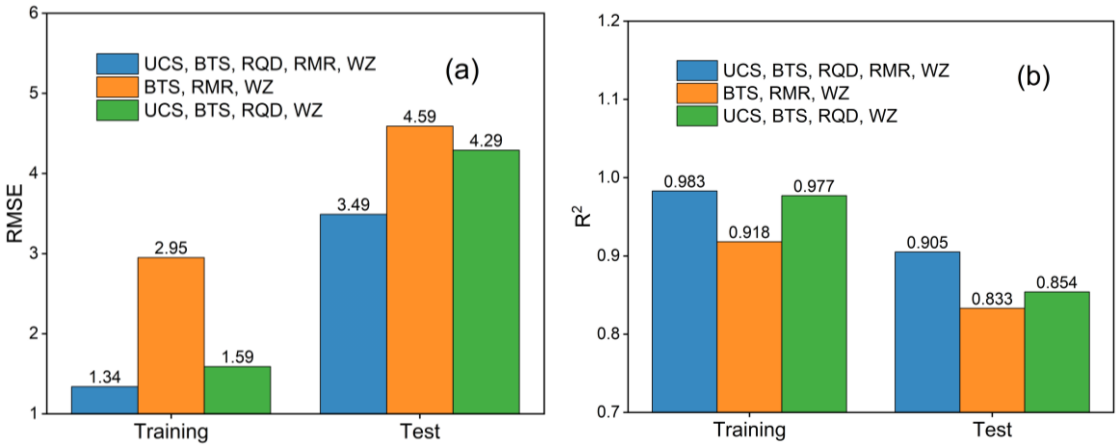


Figure 5.8 Comparing model performance on input parameters in terms of (a) RMSE and (b) R^2

5.6 Comparison with other studies

The Pahang-Selangor raw water tunnel dataset has emerged as a reference in various studies on predicting penetration rates (Armaghani et al., 2017; Koopialipoor et al., 2020; Zhou et al., 2021a). Notably, input features involve both geological and operational data, as delineated in Eq. 5.3. When predicting penetration rates, the operational parameters of TH and RPM are not accessible in practice.

$$\mathrm{PR} = f(\mathrm{UCS}, \mathrm{BTS}, \mathrm{RQD}, \mathrm{RMR}, \mathrm{WZ}, \mathrm{TH}, \mathrm{RPM}) \quad 5.3$$

Accounting to algorithms, Armaghani et al. (2017) integrated the imperialism competitive algorithm (ICA) into ANN, and Koopialipoor et al. (2020) similarly integrated the firefly algorithm (FA) into ANN to improve predictive accuracy. In a more comprehensive endeavour, Zhou et al. (2021a) combined extreme gradient boosting (XGBoost) with six different optimisation algorithms.

The unit of penetration rate is 'mm/min' in this Chapter, and the unit of other literature is 'm/h' or a normalised scale ranging from 0 to 1. Given that the coefficient of determination, R^2, remains impervious to magnitude fluctuations, it serves as a variance metric for comparative analysis. Thus, the SVM, RF, and ANN models proposed herein are juxtaposed against their counterparts from the literature based on R^2, as delineated in Figure 5.9. Although the literature models seem to outperform these proposed models by adding operational parameters, the practicality of SVM, RF, and ANN models in tunnelling remain indisputably robust.

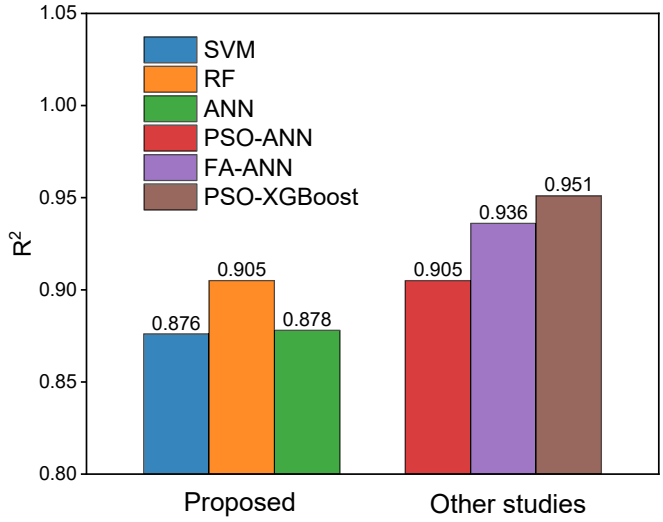


Figure 5.9 Comparing model performance between proposed and other literature models by R^2

5.7 Summary

This Chapter attempts to predict penetration rate using geological parameters as it directly affects overall productivity in tunnelling. During data processing, the IForest technique is employed to remove outliers. The penetration rate shows strong negative correlations with UCS, BTS, RQD, and RMR variables, evidenced by the respective PCC. Conversely, the correlation between PR and WZ appears rather low, attributable to its broad categorisation.

It is expected that training metrics are slightly better than test metrics. Training metrics indicate the adeptness of machine learning models in congruence with training data. Overfitting, a prevalent pitfall, occurs when the training loss diminishes, but the test loss commences an upward trajectory. So, test metrics are regarded as criteria of model

performance, reflecting the generalisation to new, unseen data.

In terms of evaluation metrics, the model performance is evaluated by seven regression metrics. Given a little discrepancy, RMSE and MAPE are primarily employed to represent errors, while R^2 is harnessed for variance characterisation.

Regarding model optimisation, the hyperparameters for SVM, RF, and ANN models are fine-tuned via random search. Comparative analysis reveals that the RF model performs better than the SVM and ANN models because penetration rate regression is not a very complex problem. The ANN model is likely too complex, and the SVM model is probably too simple to capture information. Alternatively, the RF model is the best suitable to predict penetration rate with accuracy.

It is found an overlap between RMR and its internal parameters (UCS and RQD). Using the same RF model, Model 5.2, which includes all five geological parameters, is better than Models 5.4–5.5, which exclude either UCS and RQD or RMR.

Although other literature models seem to outperform the proposed models by adding operational parameters, the input features of operational parameters (RPM and TH) are inaccessible during the training process. While the models integrating both operational and geological parameters might achieve high accuracy, they are impractical for real-world applications.

Chapter 6. Penetration Rate Forecasting

6.1 Introduction

Changsha metro line No.4 was excavated by EPB shield in slate, silty clay, limestone, gravel, mudstone, and sandstone, while Zhengzhou metro line No.3 was also excavated by EPB shield in fine sand and silty clay. Changsha and Zhengzhou metro lines collected 550 samples in the sequence, and the length of the samples was kept at 1.5 m. Time series forecasting of penetration rate (PR) is a real-time prediction using known historical data to predict the unknown future PR. Such a real-time prediction is beneficial for project time management, early warning of possible accidents and making necessary adjustments.

This chapter firstly builds univariate models to predict the near future (0–7.5 m) TBM penetration rate, which only uses its historical development in Eq. 6.1. Multivariate models further study one-step forecast (0–1.5 m) and combine other geological and operational parameters as input features, expressed in Eq. 6.2.

$$\mathrm{PR}_{t+1}, \mathrm{PR}_{t+2}, \mathrm{PR}_{t+3}, \mathrm{PR}_{t+4}, \mathrm{PR}_{t+5} = f(\{\mathrm{PR}_t\}) \quad 6.1$$

$$\mathrm{PR}_{t+1} = f(\{\mathrm{PR}_t, \mathrm{TO}_t\ , \mathrm{TH}_t, \mathrm{RPM}_t, \mathrm{FP}_t, \mathrm{CD}_t, \mathrm{WT}_t, \mathrm{GC}_t\}) \quad 6.2$$

where PR_{t+1} = Penetration rate in the next step

$\{\mathrm{PR}_t\}$ = Time series of penetration rate

6.2 Data processing

Noise can affect the intrinsic characteristics of a time series, introduced by sensor errors or document digitalisation. The simple moving average (SMA) is a data preparation technique that removes fine-grained variations in the original data by calculating averages across different subsets (Box and Pierce, 1970). In contrast, the exponential moving average (EMA) creates smoothed data by placing a greater weight and significance on recent data points (Roberts, 2000).

Figure 6.1 illustrates the time series of original and smoothed data. The SMA data is delineated by an orange line, accounting for subset lengths of 5 segments (equivalent to 7.5 m of tunnel length). Meanwhile, the green line represents the EMA data with a smoothing factor of 0.67. It is observed that the SMA data is smoother than the EMA data. However, the correlation between original data and SMA stands at 0.668, notably lower than that of EMA (0.946). After data smoothing, sharp changes in the time series are substantially attenuated.

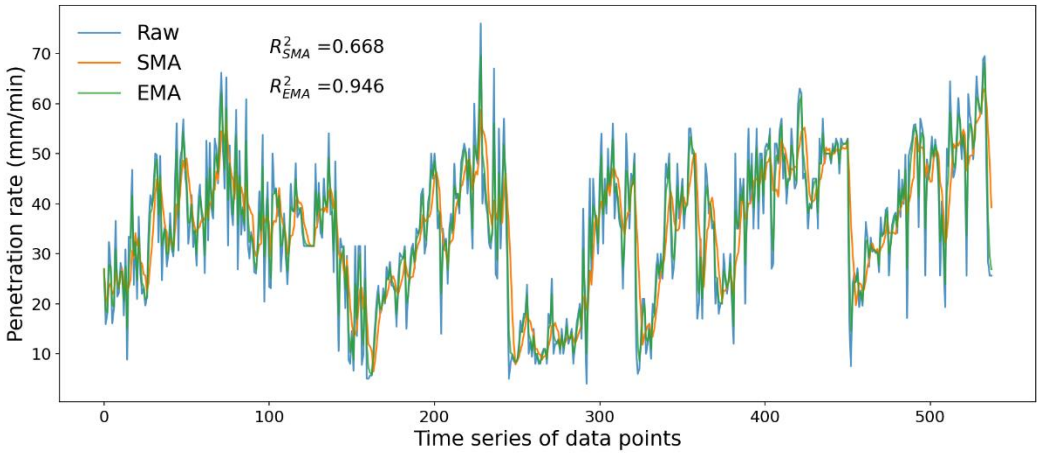


Figure 6.1 Time series of original and smoothed penetration rate

Regarding residuals between the original and smoothed penetration rates, the SMA residuals supersede those of EMA, as presented in Figure 6.2(a). The probability density estimations are normally distributed with means near zero in Figure 6.2(b) and 6.2(c), respectively. The SMA removes more noise and has a larger standard deviation of 8.330 relative to EMA's 3.485. In practical engineering, the residuals between original and smoothed data are sufficiently in a normal distribution, or wavelet transform assumes that noise is high-frequency components after sequential decomposition (Shi et al., 2021).

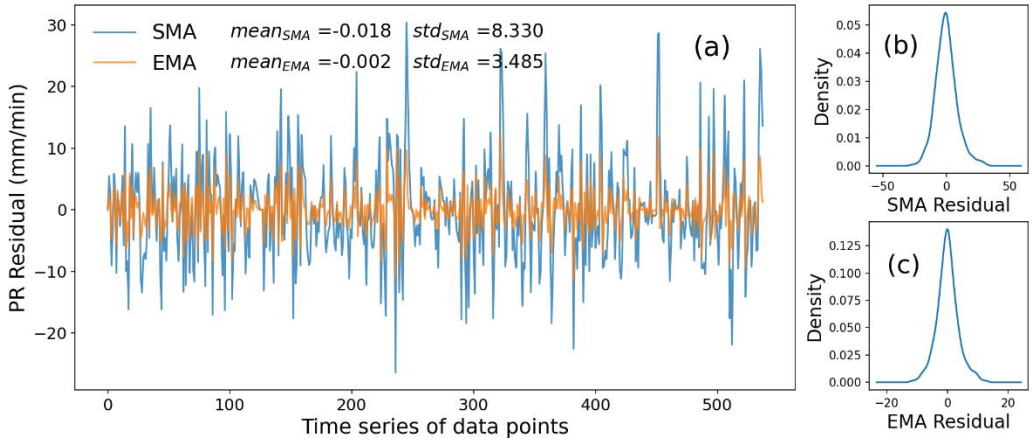


Figure 6.2 Time series of (a) residuals and probability density estimations of (b) SMA and (c) EMA

Although the differences between original and smoothed data are noticeable at some point, integrations of penetration rate, i.e. the total distance, are almost the same. That means there is no influence on the estimation of excavation distance after the moving average. In addition, moving average methods smooth data rather than remove sharp points, and the information from sharp points is averaged and stored in smoothed data. It is worthwhile noting that subsequent sections apply SMA in Section 6.3 and EMA in Section 6.4. Likewise, geological and operational parameters are scaled by min-max normalisation in Eq. 5.2, ensuring consistent feature scaling.

6.3 Univariate model

Univariate models are built to predict penetration rates in the future based on historical data, excluding other related parameters. The performance of 22 models is summarized in Table 6.1. The N^{th} step penetration rate in the future is predicted based on RNN, LSTM,

and recursive RNN. The effects of forecast horizons are studied from the first step forecast (0–1.5 m ahead of cutterhead) to the fifth step forecast (6–7.5 m). In addition, RNN models are also used to predict the average penetration rate of the next *N* steps in the future, with *N* being one (0–1.5 m) to five (0–7.5 m).

Table 6.1 Model performance in univariate models

Model	**Forecast horizon**	**Algorithm**	**Evaluation metric**		**Baseline**	
			RMSE	**R^2**	**$RMSE_{bs}$**	**R^2_{bs}**
6.1	0 – 1.5 m	RNN	0.147	0.943	0.178	0.917
6.2		LSTM	0.152	0.938		
6.3	1.5 – 3 m	RNN	0.241	0.856	0.283	0.804
6.4		LSTM	0.248	0.852		
6.5		RNN_re	0.249	0.844		
6.6	3 – 4.5 m	RNN	0.307	0.752	0.364	0.685
6.7		LSTM	0.317	0.748		
6.8		RNN_re	0.317	0.746		
6.9	4.5 – 6 m	RNN	0.370	0.672	0.419	0.606
6.10		LSTM	0.372	0.651		
6.11		RNN_re	0.374	0.671		
6.12	6 – 7.5 m	RNN	0.410	0.609	0.458	0.552
6.13		LSTM	0.418	0.604		
6.14		RNN_re	0.412	0.625		

6.15	0 – 3 m	RNN_ li	0.205	0.876	0.253	0.831
6.16		RNN_si	0.193	0.908	0.214	0.896
6.17	0 – 4.5 m	RNN_ li	0.262	0.799	0.302	0.756
6.18		RNN_si	0.213	0.881	0.247	0.850
6.19	0 – 6 m	RNN_ li	0.301	0.790	0.324	0.784
6.20		RNN_si	0.237	0.856	0.268	0.823
6.21	0 – 7.5 m	RNN_ li	0.317	0.767	0.341	0.758
6.22		RNN_si	0.266	0.831	0.283	0.820

Figure 6.3(a) is an example of the RNN structure, using the last three steps to predict the next step (0–1.5 m), called the first step forecast or one-step forecast. The N^{th} step forecast attempts to predict a one-step value in a longer forecast horizon. For example, we predict the second step penetration rate (1.5–3 m) using the last three steps. The direct method intuitively feeds the second step into outputs like Figure 6.3(b), called direct RNN. Although inputs are time series with time sequence and values of penetration rate, outputs only contain values of penetration rate. The forecast horizon of $t+2$ is manually fed into the output, so there is no difference between the one-step forecast and the second step forecast. Recursive RNN attempting multiple one-step forecasts tries to address the lack of the forecast horizon in the output, as shown in Figure 6.3(c). Every predicted value is fed back as a new input to predict the next step value. In this way, it indirectly adds the forecast horizon according to the number of iterations in the N^{th} step forecast. However, errors between measured and predicted results may accumulate as the number of

iterations increases.

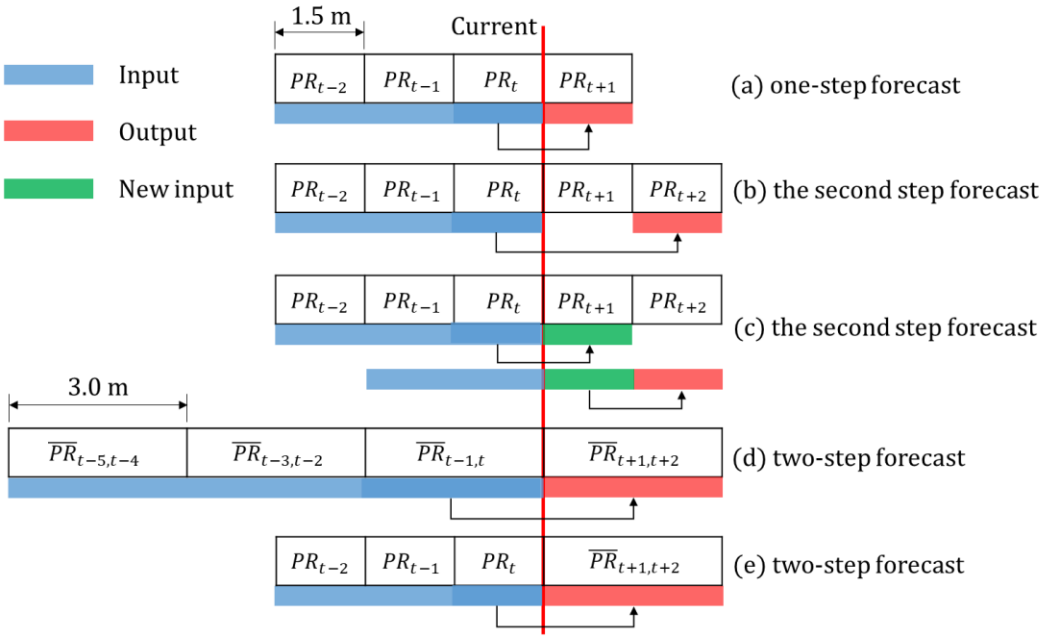


Figure 6.3 Time series forecasting on (a) one-step forecast, (b) the second step forecast of RNN, (c) the second step forecast of recursive RNN, (d) two-step forecast using long input, and (e) two-step forecast using short input

Alternatively, the two-step forecast attempts to predict the average penetration rate of the next two steps in the future (0–3 m ahead of the cutterhead), for example, using the last three steps. According to the forecast distance of two segments length, we split data every two segments in the sequence and then calculate the average of every two segments as a new sample, shown in Figure 6.3(d). Similarly, the last input $\overline{PR}_{t-1,t}$ represents the average penetration rate of the last two steps (-3–0 m) equal to the length of the output $\overline{PR}_{t+1,t+2}$. The structure of the two-step forecast is the same as that of RNN in Figure 6.3(a), but the length of every sample changes from one segment (1.5 m) to two segments

(3 m). To make good use of data, Figure 6.3(e) inputs the last three steps to predict the average penetration rate of the next two steps (0–3 m) in the future. In this structure, the length of every input is 1.5 m, while that of every output is 3 m.

6.3.1 One-step forecast

Model 6.1 predicts the next one-step penetration rate using RNN, called the one-step forecast or the first step forecast. Their optimal hyperparameters are a batch size of 32, a hidden layer size of 10, a window size of 7 time steps, and a learning rate of 0.005. In Figure 6.4, the training process saves a model at the epoch of 65 and is stopped at the epoch of 85 by early stopping. Validation loss in blue within the last twenty epochs (65–85) starts to increase, showing that the saved trained model is convergent without over-fitting.

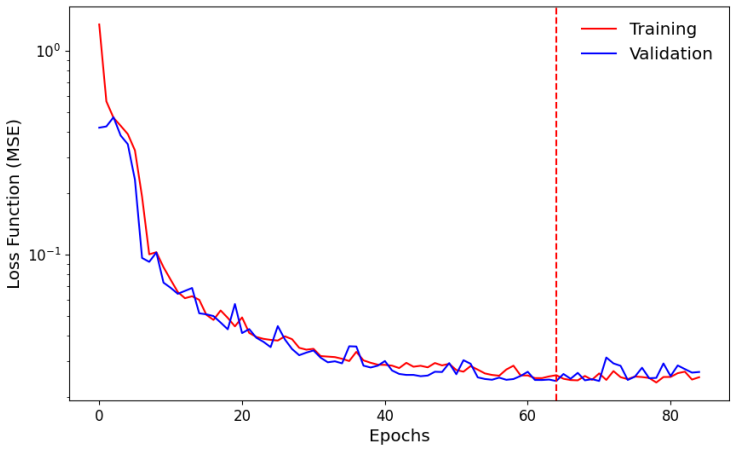


Figure 6.4 Loss function against epochs in training and validation

The embedded figure in Figure 6.5(a) shows the whole training results from Changsha,

where the blue line is measured data, and the orange and green lines are predicted results by RNN (Model 6.2) and LSTM (Model 6.3), respectively. The first eighty training results, i.e. the same number of samples as in test data, are highlighted in Figure 6.5(a) to facilitate the comparison between the training and evaluation results. RNN and LSTM perfectly predict the next step penetration rate in the training data. The performance of trained models is then evaluated by new test data from Zhengzhou, shown in Figure 6.5(b). The test results follow the trend of measured data very closely, even for sharp increases or decreases. The RMSE equals 0.147 for RNN and 0.152 for LSTM, which are relatively low compared to the average PR of 2.10. The high R^2 of 0.943 for RNN and 0.938 for LSTM indicate a strong correlation between the measured and predicted results. Models have seemingly transcended the boundary, given the different geological conditions involved in the training data from Changsha and test data from Zhengzhou. Many researchers, therefore, believed that one-step forecasts of TBM performance by machine learning were fully acceptable with high confidence (Lau et al., 2010; Gao et al., 2019; Feng et al., 2021). In addition, RNN and LSTM perform well in one-step forecast, with the former slightly better. This is because the input length is only the last seven time steps.

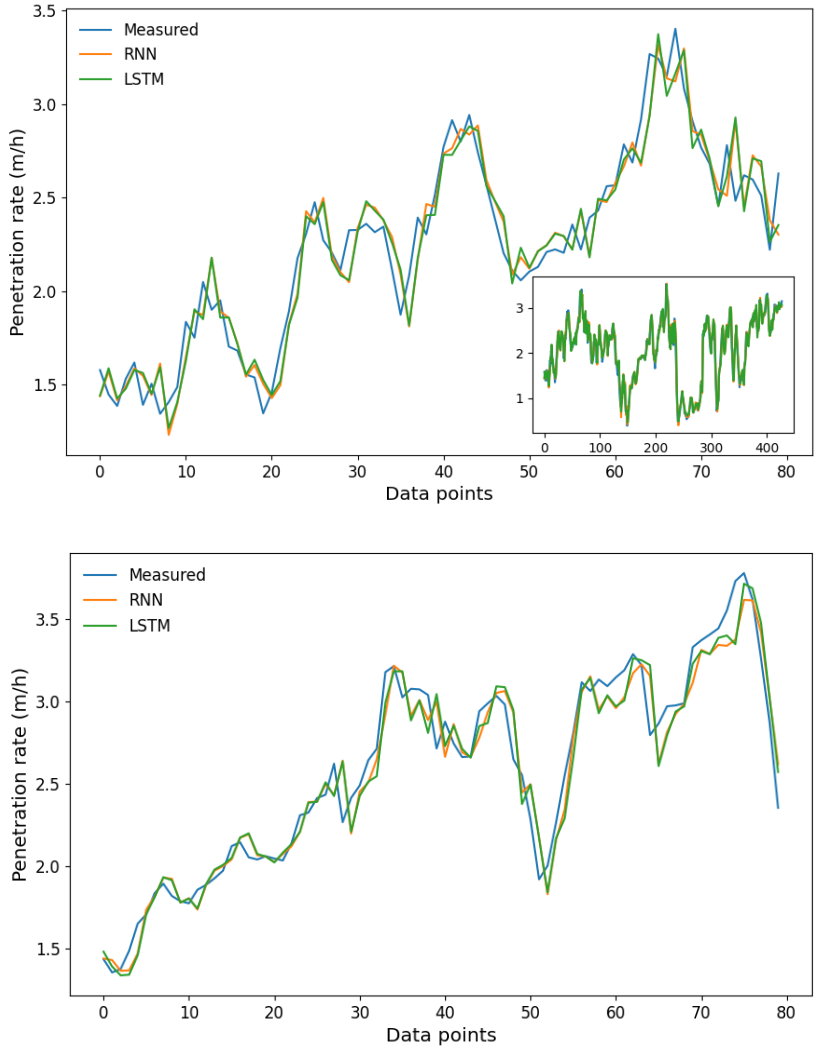


Figure 6.5 Measured and predicted penetration rate by RNN and LSTM in one-step forecast in the (a) training set and (b) test set

6.3.2 The N^{th} step forecast

The N^{th} step forecast of penetration rate studies the effects of forecast horizon using the RNN, LSTM, and recursive RNN from Model 6.3 to Model 6.14. Figure 6.6 shows the measured and predicted penetration rates at various forecast horizons, including Model 6.3 (the second step or 1.5–3 m ahead of the cutterhead), Model 6.6 (the third step or 3–

4.5 m ahead of the cutterhead), Model 6.9 (the fourth step or 4.5–6 m ahead of the cutterhead), and Model 6.12 (the fifth step or 6–7.5 m ahead of the cutterhead). It is notable that the prediction of penetration rate generally becomes less reliable for more steps into the future.

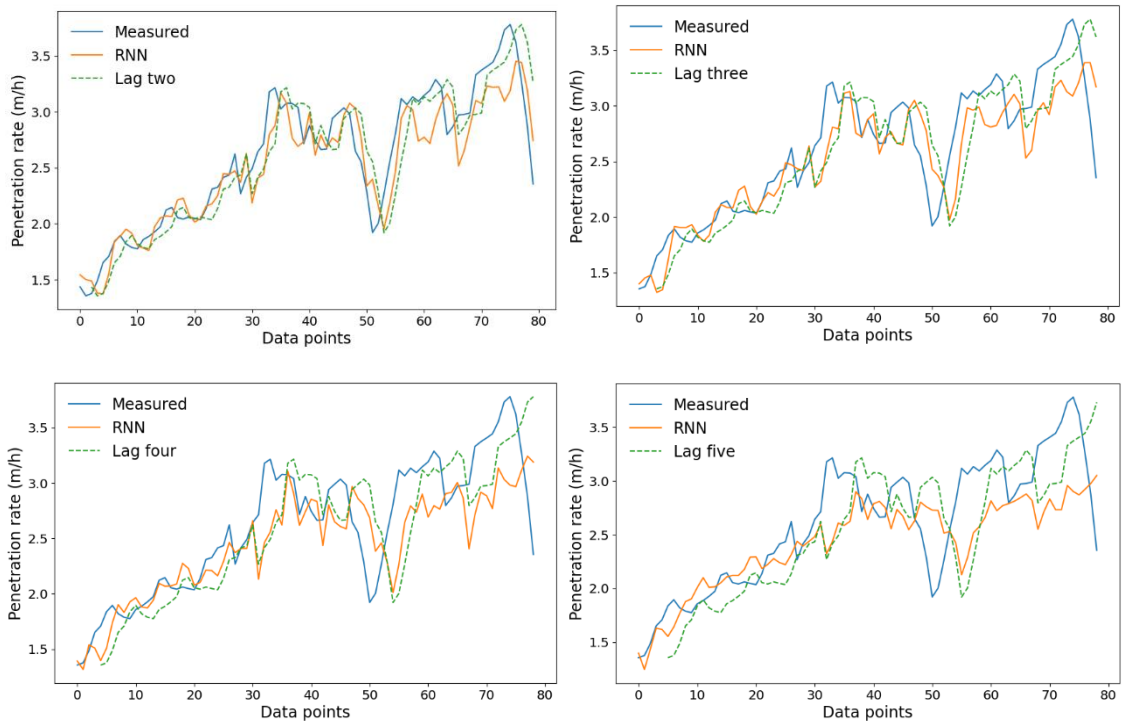


Figure 6.6 Comparisons of measured, predicted and lagged results in (a) the second step forecast, (b) the third step forecast, (c) the fourth step forecast, and (d) the fifth step forecast

Figure 6.7 shows the performance of RNN, LSTM, and recursive RNN models in terms of RMSE, R^2 and forecast horizon. With regard to the RNN (red circles), the RMSE increases from 0.147 to 0.410 as the forecast horizon increases from the first step to the fifth step. R^2 decreases from 0.943 to 0.856, 0.752, 0.672, and 0.609 for the forecast

horizon of one, two, three, four, and five. Results indicate that the accuracy decreases with an increasing forecast horizon. Despite many research attempts in the literature (Shi et al., 2021), this finding is not surprising because the geological conditions further away from a TBM cutterhead should have less effect on its current operational parameters. The LSTM in blue triangles connected by blue lines in Figure 6.7 displays similar trends as the RNN.

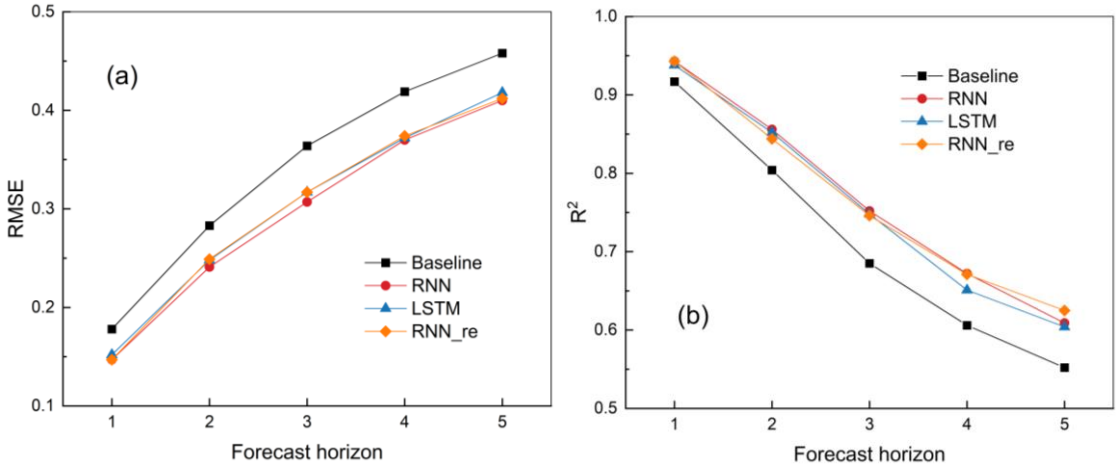


Figure 6.7 Effects of forecast horizons in the N^{th} forecast of (a) RMSE and (b) R^2

Looking back at Figure 6.6, we find that time lags exist between the measured and predicted results. If measured data (blue lines) is shifted *N* steps to the right (future), they (green dot lines) will align well with the predicted results (orange lines). In other words, when an observation peaks at the data point of *t*, the corresponding prediction peaks at the data point of *t*+*N*. Interestingly, the number of time lags (*N*) corresponds to the forecast period of *N* steps, which is like time-delayed predictions instead of time series

predictions. The training process lacks information on the forecast horizon, probably leading to *N* lags in the N^{th} step forecast.

Time lags seem to affect the accuracy of the N^{th} step forecast significantly. Alternatively, the recursive RNN (RNN_re) makes multiple one-step forecasts where each output can be fed back into itself for the next prediction, as shown in Figure 6.3(c). The method indirectly adds a forecast horizon according to the number of iterations. However, errors between measured and predicted results accumulate with the number of iterations. Somewhat surprisingly, the recursive RNN is similar to the RNN and does not offer much improvement in the evaluation criteria of RMSE or R^2 in Figure 6.7.

6.3.3 *N*-step forecast

N-step forecast attempts to predict the average penetration rate of *N* steps in the future using RNN from Model 6.15 to Model 6.22. On the one hand, the input length equals the length of the forecast period. Referring to the structure in Figure 6.3(d), the sample inputs $\overline{PR}_{t-5,t-4}, \overline{PR}_{t-3,t-2}, \overline{PR}_{t-1,t}$ and outputs $\overline{PR}_{t+1,t+2}$, and the next sample inputs $\overline{PR}_{t-4,t-3}, \overline{PR}_{t-2,t-1}, \overline{PR}_{t,t+1}$ and outputs $\overline{PR}_{t+2,t+3}$, and so on. On the other hand, we use short inputs equivalent to one segment of 1.5 m, and the forecast period is *N* steps in the future in Figure 6.3(e). Figure 6.8 shows the measured and predicted penetration rates at various forecast horizons. Like the N^{th} step forecast, the reliability of the *N*-step forecast becomes worse as the forecast horizon increases.

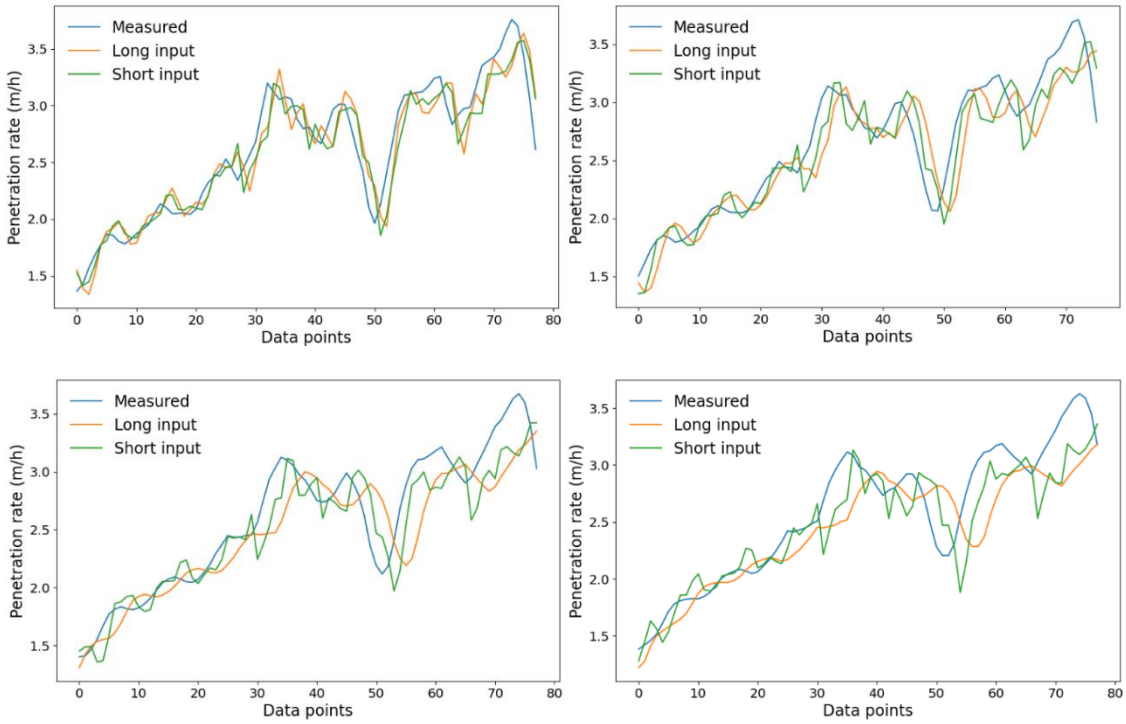


Figure 6.8 Comparisons of measured, predicted and lagged results in (a) two-step forecast, (b) three-step forecast, (c) four-step forecast, and (d) five-step forecast

Figure 6.9 compares the two RNN methods regarding RMSE, R^2 and forecast horizon. Again, the error increases with an increasing forecast horizon. The RMSE for long inputs (RNN_li) increases from 0.147 to 0.205, 0.262, 0.301 and 0.317 for forecast distances of 1.5, 3, 4.5, 6, and 7.5 m. The RMSE for short inputs (RNN_si) increases slightly from 0.147 in the one-step forecast to 0.266 in the five-step forecast. The R^2 decreases with an increasing forecast horizon, from 0.943 to 0.767 for RNN_li and from 0.943 to 0.831 for RNN_si. Using short inputs in *N*-step forecasts is much better than using long inputs. A possible reason is that helpful information in the long input may get lost when averaging

several segments into one sample.

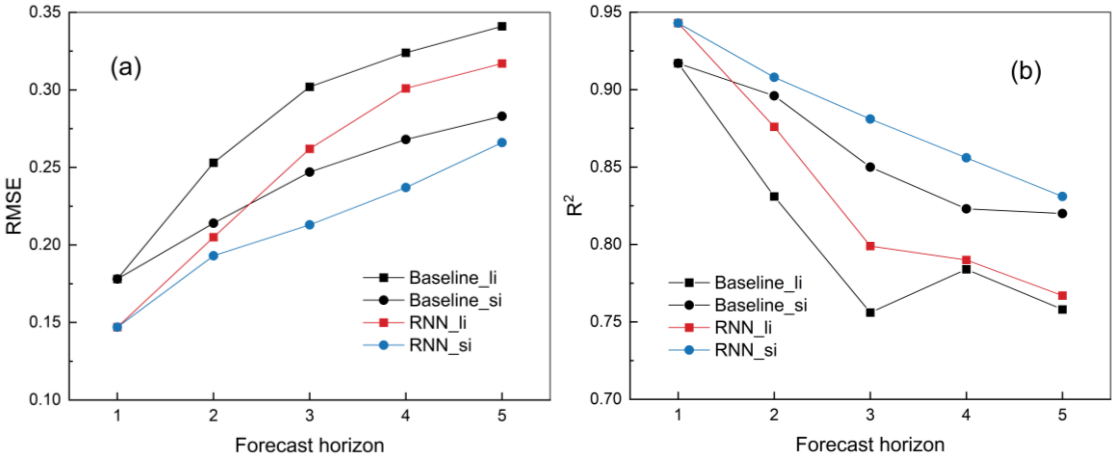


Figure 6.9 Effects of forecast horizon in the *N*-step forecast of (a) RMSE and (b) R^2

6.3.4 Baseline model

The accuracy in predicting the future penetration rate depends on the time or distance to the future. Such a result makes sense, as we would have expected that the rock quality and geological conditions further ahead of the TBM cutterhead would have less influence on its current operational data.

Machine learning models generally display undesirable time lags between measured and predicted results (Shan et al., 2022). The possible reason is that the time lag is a characteristic of time series forecasting of penetration rates, which is motivated by a random walk theory. Random walk assumes that the next value, for example, of the stock market or price, has a fifty-fifty per cent chance to grow or drop. The difference between successive steps statistically shows a normal distribution.

Accordingly, the increment in TBM penetration rate over time is plotted in histograms and probability density estimations in Figure 6.10. The distribution of the penetration rate increment is in a quasi-normal distribution and is almost symmetrical about the x-axis of zero, which means a fifty-fifty per cent chance that the next value is lower or higher than the current value. The increment of zero (no change) has the highest likelihood, learned in the training process, so the most likely prediction of the next value is like a copy of the last value. Therefore, time lags are observed between measured and predicted results.

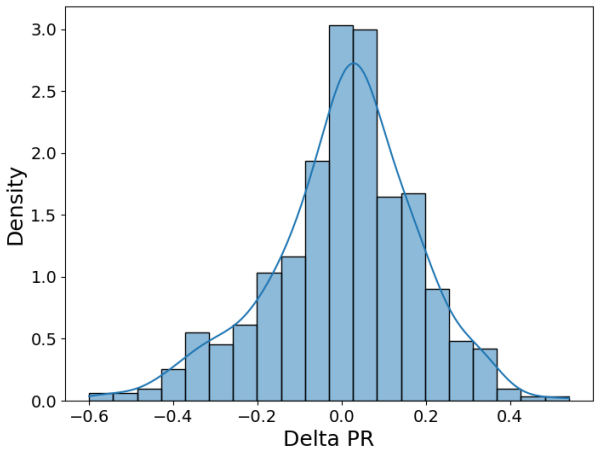


Figure 6.10 Histograms and probability density estimations of the penetration rate increment

If a copy of the last input is set as a baseline, the coefficient of determination in baselines (R^2_{bs}) is shown in the last column of Table 6.1. In the one-step forecast, the baseline result is surprisingly good at 0.917, and the gap between Model 6.1 and its baseline is only 0.026. In the N^{th} step forecast, R^2_{bs} decreases from 0.917 for the first step forecast to 0.552 for the fifth step forecast in Figure 6.7(b). The gaps between RNN models and baselines

become larger with an increasing forecast horizon. Regarding the *N*-step forecast in Figure 6.9(b), R^2_{bs} shows a similar trend, decreasing from 0.917to 0.758 for long inputs and from 0.917 to 0.820 for short inputs. It is noted that the baseline results for short inputs are again better than those for long inputs.

Because the penetration rate increment distribution is likely to be a quasi-normal distribution, predicting the penetration rate in the future can be a quasi-random walk. The problem of *N* time lags in the N^{th} step forecast is very challenging. Furthermore, deep learning algorithms, RNN and LSTM, help to learn the potential relationship between time series and to build better models than their baselines.

6.4 Multivariate model

Using historical data, multivariate models aim to predict the next penetration rate (1.5 m ahead). The dataset from Changsha metro trains time series forecasting models, and their model performance is evaluated by RMSE and R^2 on the dataset from Zhengzhou metro, as summarized in Table 6.2. The hyperparameters for the near-optimal models are presented in Table 6.3. In the table, $\{PR_t\}$ refers to time series data up to the recent penetration rate, and PR_{t+1} represents the penetration rate in the next step. Hidden size 1 refers to the hidden size derived from time series PR, while hidden size 2 refers to the hidden size derived from the last step of geological and operational parameters in the advanced RNN model.

Table 6.2 Model performance in multivariate models

Model	Output	Input	Method	RMSE	R^2
6.23	PR_{t+1}	$\{PR_t\}$	RNN	0.4034	0.6741
6.24	PR_{t+1}	$\{PR_t, GEO_t\}$	RNN	0.4046	0.6718
6.25	PR_{t+1}	$\{PR_t, OP_t\}$	RNN	0.4117	0.6593
6.26	PR_{t+1}	$\{PR_t, GEO_t, OP_t\}$	RNN	0.4053	0.6638
6.27	PR_{t+1}	$\{PR_t\}, GEO_t, OP_t$	Advanced RNN	0.3952	0.6913

Table 6.3 Hyperparameters in multivariate models

Model	Batch size	Time step	Hidden size 1	Hidden size 2	Number of layers	Learning rate
6.23	32	3	20	-	1	0.005
6.24	64	3	50	-	1	0.01
6.25	64	5	100	-	1	0.01
6.26	32	5	100	-	1	0.01
6.27	32	3	100	50	1	0.01

6.4.1 Historical data-based model

The effects of data smoothing will be discussed in Section 6.6, and EMA processes multivariate models in Section 6.4, which is different from univariate models using SMA in Section 6.3. As a result, Model 6.23, a univariate RNN model, is used for comparison.

Figure 6.11 shows the penetration rate from Changsha, where a full blue line represents measured values, and a dotted orange line represents predicted values. The univariate RNN model perfectly fits the next-step penetration rate in the training data. The trained

model is evaluated using unseen data from Zhengzhou. The test results closely follow the trend of measured data with RMSE of 0.4034. The value of R^2 is equal to 0.6741, indicating a correlation between measured and predicted values.

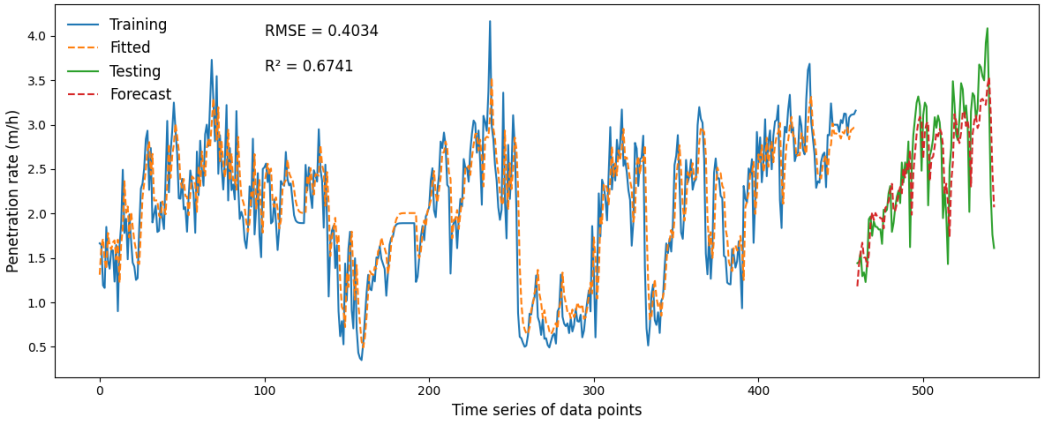


Figure 6.11 Measured and predicted results using historical penetration rate

Apart from the univariate model, multivariate models take into account the impact of other geological and operational parameters on TBM performance. Model 6.24 incorporates geological parameters (GEO) in Figure 6.12, while Model 6.25 incorporates operational parameters (OP) in Figure 6.13, in which good agreements between the measured and predicted values are observed. When geological parameters are added to Model 6.24, the results are with RMSE of 0.4046 and R^2 of 0.6718. In contrast, Model 6.25 with operational parameters has a lower performance with RMSE of 0.4147 and R^2 of 0.6593.

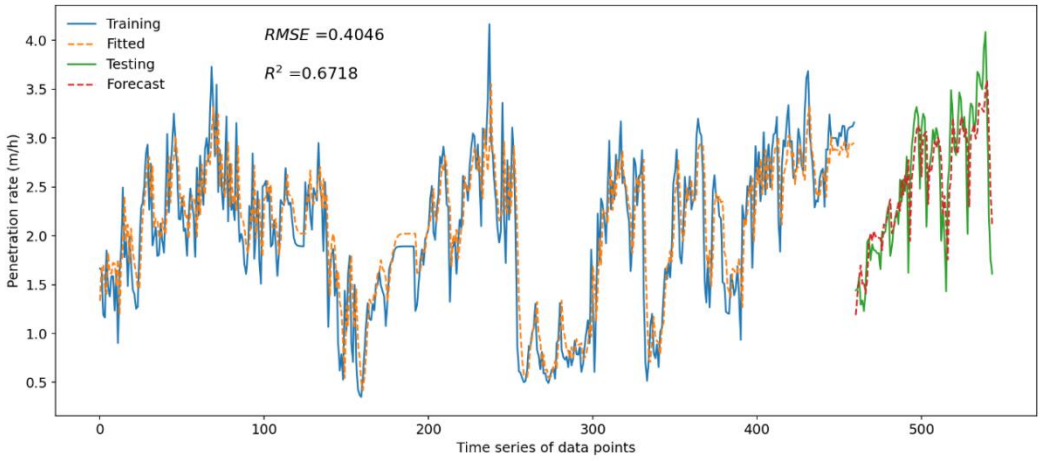


Figure 6.12 Measured and predicted results using historical penetration rate and geological data

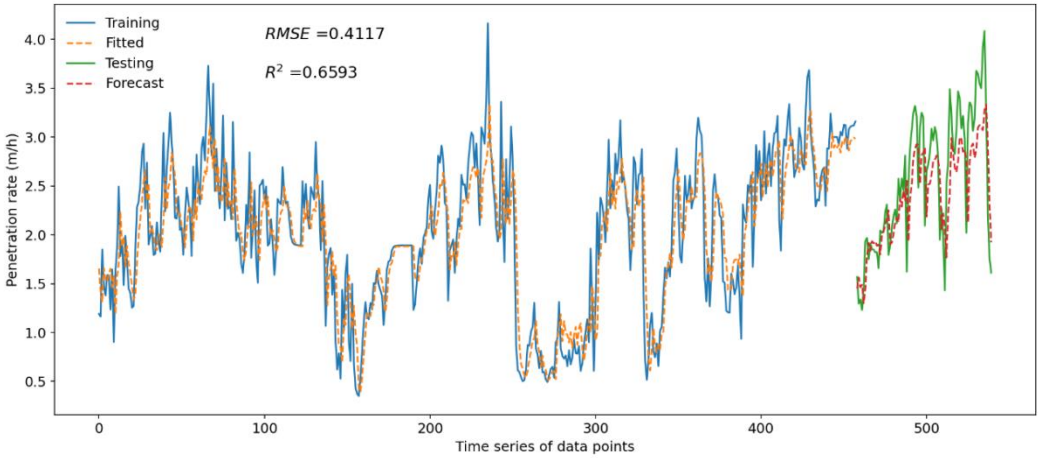


Figure 6.13 Measured and predicted results using historical penetration rate and operational data

Model 6.26 further incorporates both GEO and OP with the best combination of hyperparameters: a window size of 3, a hidden size of 20, a learning rate of 0.005, a number of layers of 1, and a batch size of 32. In Figure 6.14, training and validation losses decrease dramatically at the beginning and then slightly decrease after the epoch of 10. The embedded Figure 6.14 provides a detailed view of the changes in the loss function

between epochs 4 and 14. The early stopping saves a converged model at the epoch of 62 to prevent overfitting, which yields the minimum validation loss. During the last 20 epochs (62–82), the validation loss no longer decreases, showing that the saved model is subjected to validation data without overfitting.

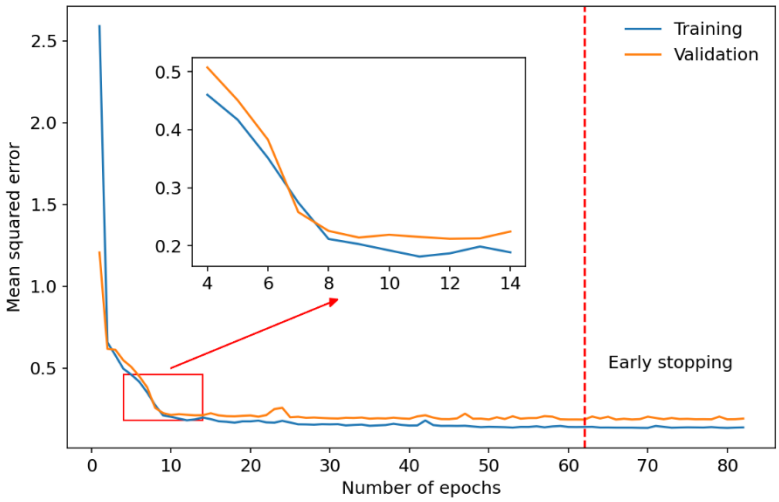


Figure 6.14 Loss function against epochs in training and validation

Figure 6.15 shows the measured penetration rate and the predicted values, resulting in RMSE of 0.4053 and R^2 of 0.6638. However, the multivariate models do not perform much better than the univariate model (Model 6.23). It is counter-intuitive that incorporating other parameters does not improve the accuracy of time series forecasting, probably due to the additional parameters increasing model complexity with helpful and irrelevant information. This irrelevant information negatively affects feature extraction and degrades model performance.

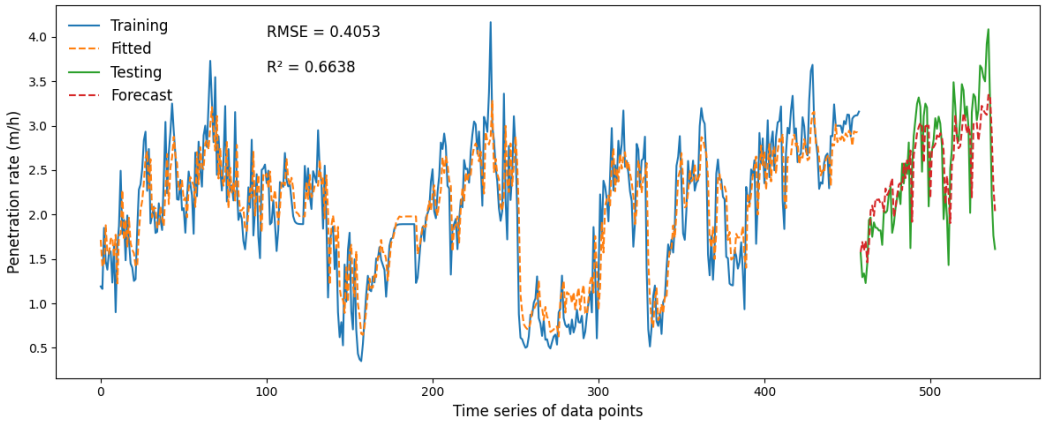


Figure 6.15 Measured and predicted results using historical penetration rate and geological and operational data

6.4.2 Last-step covariate model

The advanced RNN reconfigures the inputs into two components: a time series of PR and the last-step GEO and OP. The time series of penetration rate $\mathrm{PR}_{t-2}, \mathrm{PR}_{t-1}, \mathrm{PR}_t$ are fed into RNN one by one, producing h_t in the recent hidden layer. At the same time, the last-step geological and operational parameters $\mathrm{GEO}_t, \mathrm{OP}_t$ are fully connected to extract hidden features $\bar{h}_t$. The output PR_{t+1} is then fully connected with h_t and $\bar{h}_t$ in a linear transformation.

Figure 6.16 illustrates measured and predicted values by the advanced RNN. Among these models, Model 6.27 has the lowest RMSE at 0.3952 and the highest R^2 at 0.6913. Therefore, the advanced RNN model successfully improves the model performance because the last-step geological and operational parameters are closer to the future penetration rate.

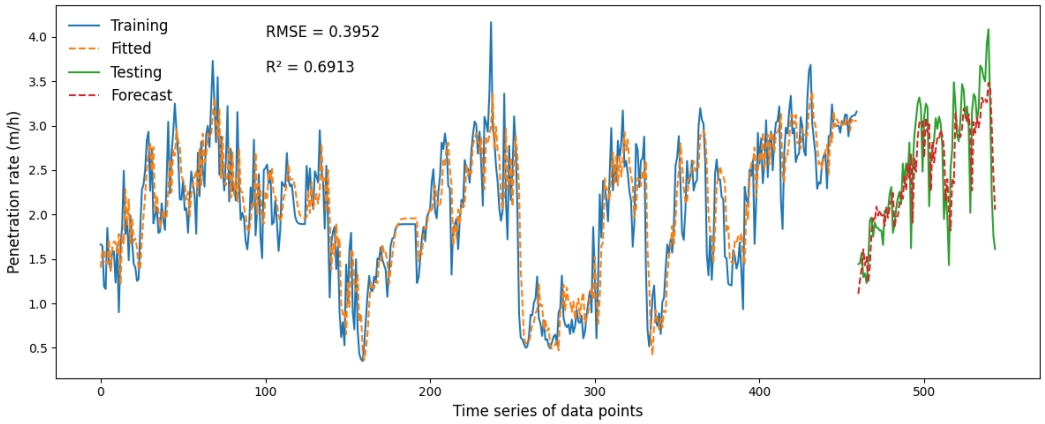


Figure 6.16 Measured and predicted results using historical penetration rate and the last-step geological and operational data

6.5 Sensitivity analysis

Good generalisation is observed as the predicted values closely follow the trends of the measured values in training and testing. Machine learning models are highly non-linear, so the relationship between the input and output is usually poorly understood. The Sobol method is widely used to study the impacts of independent inputs on the output.

For example, in Model 6.27, the input includes the last-three-step penetration rate and last-step geological and operational parameters. There are four steps in the method: (1) defining the range of parameters from 0 to 1 because all inputs are min-max normalised; (2) generating 1024 samples from the pseudo-random Sobol sequence to create 24,576 (=2×(11+1)×1024) parameter sets in total; (3) running the parameter sets through Model 6.27 to calculate the output; (4) sending the output back to calculate Sobol indices. Figure

6.17 displays the first-order Sobol index in Model 6.27. The time series of penetration rate is the most important parameter, with $S_{\{\mathrm{PR}_t\}}$ =0.8681, and the last-step penetration rate is the most important time, with S_{PR_t} =0.7191. The range of the training data is usually expected to cover the test data, but the geological conditions between Changsha and Zhengzhou are quite different. According to the results of S_{GEO_t} =0.0429 and S_{OP_t} =0.0861, it is found that other geological and operational inputs in Model 6.27 have little effect on predicting penetration rate in the future.

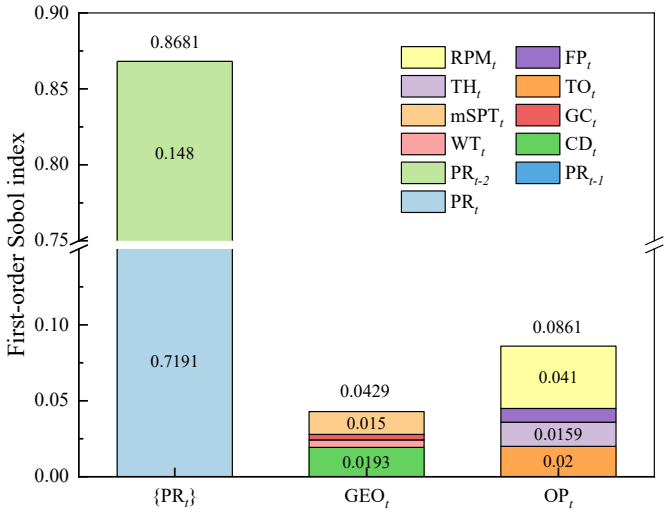


Figure 6.17 First-order Sobol index for one-step forecasts in Model 6.27

6.6 Effects of data smoothing

Noise in a time series can obscure its intrinsic characteristics. After data smoothing by EMA, the probability density estimations of residuals are in normal distributions with a mean of 0, as shown in Figure 6.2(c). However, the standard deviation of residuals

increases as the smoothing factor decreases. Although a time series of noise is in a normal distribution, the normally distributed residuals are not necessarily noise. No clear distinction between noise and clean data exists, as both are vibrations in the sequence. A reasonable assumption is that normally distributed residuals represent different levels of noise, and the smoothed data remains clean to varying extents.

Regarding the effects of data smoothing, additional models are conducted in time series forecasting. The original data ($\alpha = 1$) are processed by EMA with various smoothing factors ranging from 0.67 to 0.2. In Figure 6.18, the correlation of determination R^2 between the original and smoothed data gradually decreases from 0.9456 to 0.8807, 0.7858, and 0.6718 with decreasing smoothing factor. Notably, the smoothing factor of 0.67 has the least smoothing effect, while a value of 0.2 results in the most significant smoothing effect.

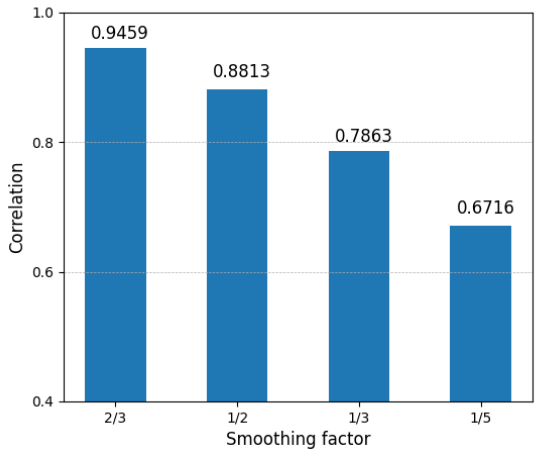


Figure 6.18 Correlation between original and smoothed data of varying smoothing factors

We should exclude the effects of other parameters and focus solely on changes in the time series when predicting the next penetration rate. Figure 6.19 shows the results of evaluation metrics against varying smoothing factors, where the algorithm architecture and learning process are the same as in Model 6.23. As the smoothing factor decreases, RMSE decreases from 0.6164 to 0.4034, 0.3092, 0.2026, and 0.1239. R^2 is 0.4061 in the original data and becomes 0.9610 when α = 0.2 in the smoothed data.

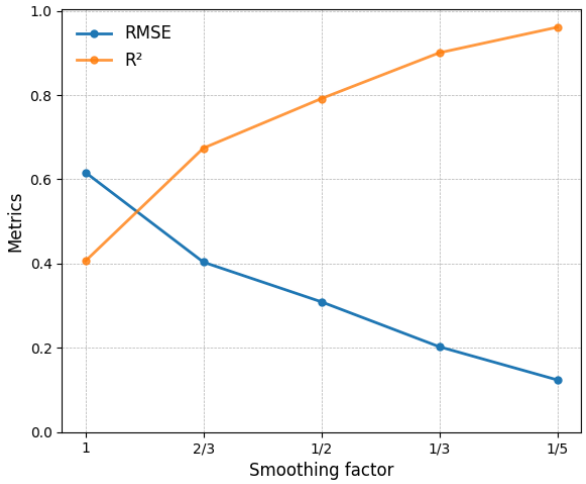


Figure 6.19 Effects of data smoothing on one-step forecast evaluated by RMSE and R^2

On the one hand, the next-step penetration rate is easier to predict when the time series is smoother. For example, the lowest smoothing factor (α =0.2) removes the largest extent of noise by EMA, resulting in the best model performance with RMSE of 0.1239 and R^2 of 0.9610. On the other hand, over-simplified data can also eliminate underlying characteristics, raising a question about whether smoothed data can represent actual data characteristics.

It is a dilemma that the original data can contain too much noise and be challenging in predicting the next-step penetration rate, while over-simplified data can lose important characteristics in the time series. In order to balance these factors, we conservatively choose a smoothing factor of 0.67 in the tests mentioned above. This allowed us to remove some of the noise while preserving the important features of the data as much as possible.

6.7 Summary

This Chapter aims to build deep learning models to predict multi-step penetration rates (0–7.5 m) using univariate models and to predict one-step penetration rates (0–1.5 m) using multivariate models. Forecasting TBM performance, if successful, would play a significant role in project time management and cost control during tunnel construction. The trained models are generalised to different geological conditions using training data from Changsha and test data from Zhengzhou.

With the given training and evaluation data, one-step forecasts of TBM performance are acceptable with high confidence and predicted results closely follow the measured data, even at sharp peaks and valleys. The one-step forecast can transcend the very different geological conditions associated with the training and evaluation data, indicating good robustness of the deep learning methods.

In the N^{th} step forecast, the accuracy decreases with an increasing forecast horizon from

the TBM cutterhead. It is found that N time lags generally exist between measured and predicted results, corresponding to the N^{th} step forecast. These time lags are present in all methods studied, including the recursive RNN. Alternatively, the N-step forecast explores the average penetration rate in the next N steps using different length inputs. The reliability of the N-step forecast also depends on the forecast horizon. Amongst the methods, the N-step forecast with short inputs yields the greatest results, and its performance over other alternatives gets more pronounced as the number N increases.

We set a baseline that copies the last input as the prediction, motivated by random walk theory. The baseline results are almost acceptable and a little worse than the predicted results of RNN models, especially in the on-step forecast. This is because the distribution of the penetration rate increment is a quasi-normal distribution. At the same time, the baseline properly explains the N time lags in the N^{th} step forecast.

However, multivariate RNN models, incorporating the time series of geological and operational parameters, perform slightly worse than the univariate RNN model. We develop an advanced RNN that splits the inputs into a time series of penetration rate and last-step other parameters, which successfully improves model accuracy.

According to the sensitivity analysis, time series penetration rate is the most important parameter, while other parameters, including geological conditions, have little impact on the time series forecast. Geological conditions can become relevant if they do not vary

much from training to test data.

Data smoothing can effectively remove noise in the time series. It is found that smoothed data are easier to predict than original data. However, over-simplified data can lose real characteristics in the time series. It is a balance between noise reduction and actual data preservation in time series forecasting.

Chapter 7. Cutterhead Torque and Thrust Force Forecasting

7.1 Introduction

The Yinsong Water Diversion Project was excavated using an open-type TBM to transfer water, with a total distance of 17.50 km. The tunnel lithology consists mainly of limestone, diorite, tuff, and granite, and rock mass classification is categorised into five classes, namely hydropower classification (HC).

This Chapter utilises deep learning techniques to predict cutterhead torque (TO) and thrust force (TH) simultaneously, as summarized in Table 7.1. The dataset, derived from the Yinsong water diversion project, is preprocessed into low-frequency data in Section 4.3.2, where each sample represents the average of a stable boring cycle. Models 7.1 and 7.2 are regression inputting geological parameters of HC and FZ, and setting values of PR_set and RPM_set, respectively. Models 7.3–7.5 are one-step forecasts comparing RNN, LSTM, and GRU algorithms based on historical operational parameters of PR, RPM, TH, and TO. Model 7.6 encompasses operational parameters and setting values to predict the next step TO and TH in the future.

Table 7.1 Model performance for forecasting cutterhead torque and thrust force

Model	Output	Input	Method	TO		TH	
				MAPE	R^2	MAPE	R^2
7.1	TO, TH	HC, FZ	ANN	0.401	0.337	0.258	0.290

7.2	TO, TH	PR_set, RPM_set	ANN	0.335	0.654	0.193	0.702
7.3	TO_{t+1}, TH_{t+1}	$\{TO_t, TH_t, PR_t, RPM_t\}$	RNN	0.168	0.773	0.093	0.820
7.4	TO_{t+1}, TH_{t+1}	$\{TO_t, TH_t, PR_t, RPM_t\}$	LSTM	0.170	0.770	0.091	0.820
7.5	TO_{t+1}, TH_{t+1}	$\{TO_t, TH_t, PR_t, RPM_t\}$	GRU	0.167	0.773	0.093	0.821
7.6	TO_{t+1}, TH_{t+1}	$\{TO_t, TH_t, PR_t, RPM_t\}$, $PR\_set_{t+1}, RPM\_set_{t+1}$	Advanced RNN	0.150	0.808	0.088	0.852

7.2 Modelling process

In addition to the sequential data, some parameters are not time-dependent but have an impact on future values. The inputs can be categorised into sequential and non-sequential data. While RNN, LSTM and GRU are adept at handling sequential data owing to the loop architecture, they cannot process non-sequential data. Therefore, we propose an advanced recurrent neural network (advanced RNN in Section 3.2.8) combined with sequential and non-sequential inputs. The sequential inputs are historical operational parameters, while the non-sequential inputs are setting values during the TBM operation.

To address the issue of varying units and magnitudes among parameters, the min-max normalisation is applied, scaling low-frequency data between 0 and 1. This normalisation stabilises gradient descent and facilitates faster convergence of the model. It is worth noting that the outputs are scaled back to the original scale after building models.

In the training process, machine learning models, especially advanced RNN, are applied to build a model in 60% of training data, with the Adam optimiser and loss function of

the mean squared error (MSE). The training process involves 50 epochs to iteratively update weights and biases. A dropout rate is implemented as a regularisation technique to prevent overfitting, randomly setting a fraction of input units to zero during training (Srivastava et al., 2014).

A separate validation dataset of 20% is used to tune hyperparameters not trained in the training process. These hyperparameters include time step [1, 2, 3, 5, 7, 10], number of layers [1, 2, 3], learning rate [0.001, 0.01], hidden size [8, 16, 32, 64, 128], dropout rate [0, 0.2, 0.5], and batch size [32, 64, 128]. A series of 60 random trials are conducted by the random search method, which is computationally efficient in finding the best combination of hyperparameters.

As a result, a near-optimal model is built and evaluated in 20% of test data, assessing the model performance and generalisation ability. The near-optimal model is evaluated by MAPE and R^2.

7.3 Regression model

7.3.1 Geological parameter-based model

During the training process, operational parameters of PR, RPM, TO, and TH are inaccessible. The model based on geological parameters is impractical for real-world application, even having high accuracy. Since geological conditions are measured via site

investigation and available before tunnelling, we apply a model based on geological parameters before the start of the project, expressed as Eq. 7.1.

$$\mathrm{TO}, \mathrm{TH} = f(\mathrm{HC}, \mathrm{FZ}) \qquad 7.1$$

where geological parameters of HC and FZ are the input vector, and TO and TH are the output vector. The weight matrix and bias are the arguments to be trained by the machine learning algorithm $f(x)$, which employs ANN here.

Figure 7.1 illustrates the testing performance of the near-optimal model, with MAPE of 0.401 and R^2 of 0.337 for torque and MAPE of 0.258 and R^2 of 0.290 for thrust. Given that geological conditions are sparsely sampled with localised information, the predicted results appear stepwise. Therefore, these geological conditions fail to capture the temporal changes in TBM performance.

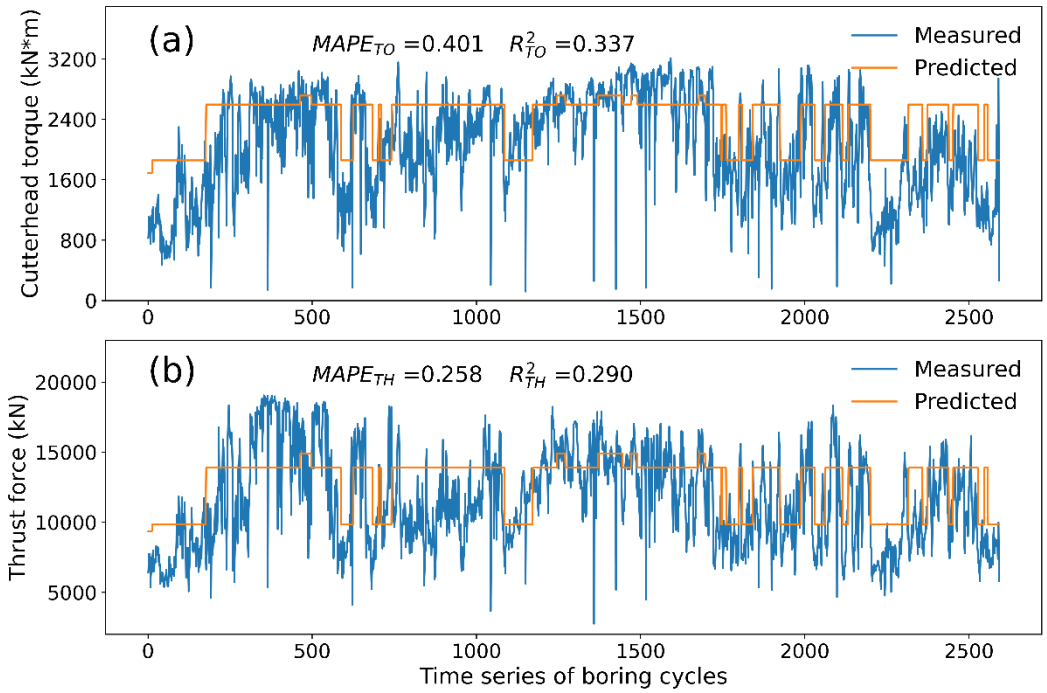


Figure 7.1 Measured and predicted results for predicting (a) cutterhead torque and (b) thrust force in the geological parameter-based model

7.3.2 Setting value-based model

Before proceeding to the next boring cycle, PR_set and RPM_set are predetermined, taking into account operators' experience and limited geological conditions. When predicting the next-step torque and thrust, the setting values for the next step are already known in advance. When it comes to multi-step forecasts, PR_set and RPM_set are not accessible because they are far away from the cutterhead. The setting values, while effective for one-step forecasts, are not as well-suited for multi-step forecasts. The model based on setting values is represented by Eq. 7.2

$$\mathrm{TO}_{t+1}, \mathrm{TH}_{t+1} = f(\mathrm{PR\_set}_{t+1}, \mathrm{RPM\_set}_{t+1}) \qquad 7.2$$

where PR_set_{t+1} and RPM_set_{t+1} are the setting values for the upcoming cycle and form the input vector. The cutterhead torque and thrust force in the next cycles, TO_{t+1} and TH_{t+1}, are the output vector. The machine learning algorithm used in this case is the ANN model.

Figures 7.1(a) and 7.2(b) illustrate that the predicted results generally align with the trend of measured results during testing, characterised by MAPE of 0.335 and R^2 of 0.654, and MAPE of 0.193 and R^2 of 0.702, respectively. This degree of alignment suggests that PR_set and RPM_set indeed have an influence on the performance of both torque and thrust.

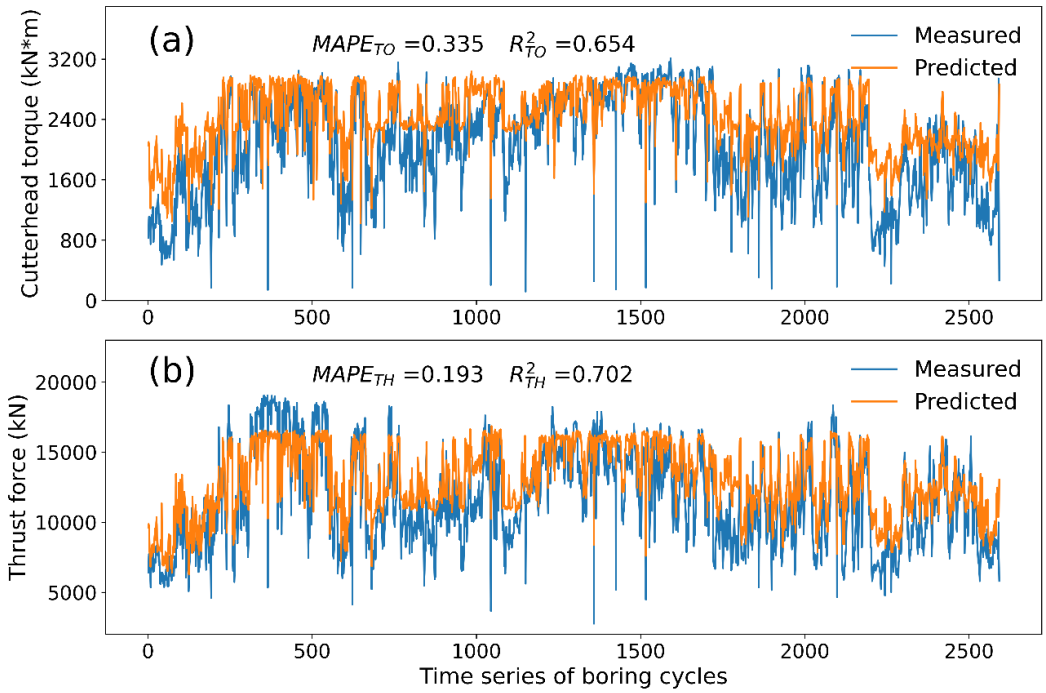


Figure 7.2 Measured and predicted results for predicting (a) cutterhead torque and (b) thrust force in the setting value-based model

7.4 Time series forecasting

7.4.1 Operational parameter-based model

One-step forecasts in low frequency are to simultaneously predict the cutterhead torque and thrust force in the next boring cycle, around 1.14 metres ahead of the cutterhead. Time series forecasting uses historical operational parameters as feature vectors to predict the next-step TBM performance in Eq. 7.3. Through a random search process, the combination of hyperparameters that yields a near-optimal RNN model is determined. These hyperparameters include a time step of 7, a number of layers of 3, a learning rate of 0.001, a hidden size of 32, a dropout rate of 0, and a batch size of 32.

$$\mathrm{TO}_{t+1}, \mathrm{TH}_{t+1} = f(\{\mathrm{PR}_t, \mathrm{RPM}_t, \mathrm{TO}_t, \mathrm{TH}_t\}) \quad 7.3$$

The trained RNN model is then evaluated using unseen data, constituting 20% of the dataset, as depicted in Figure 7.3. It is observed that the predicted results in orange demonstrate a close alignment with the measured results in blue, with MAPE of 0.168 and R^2 of 0.773 for torque and MAPE of 0.093 and R^2 of 0.820 for thrust. It is imortant to acknowledge that historical operational parameters are highly related to the future TBM performance, especially in the one-step forecasts.

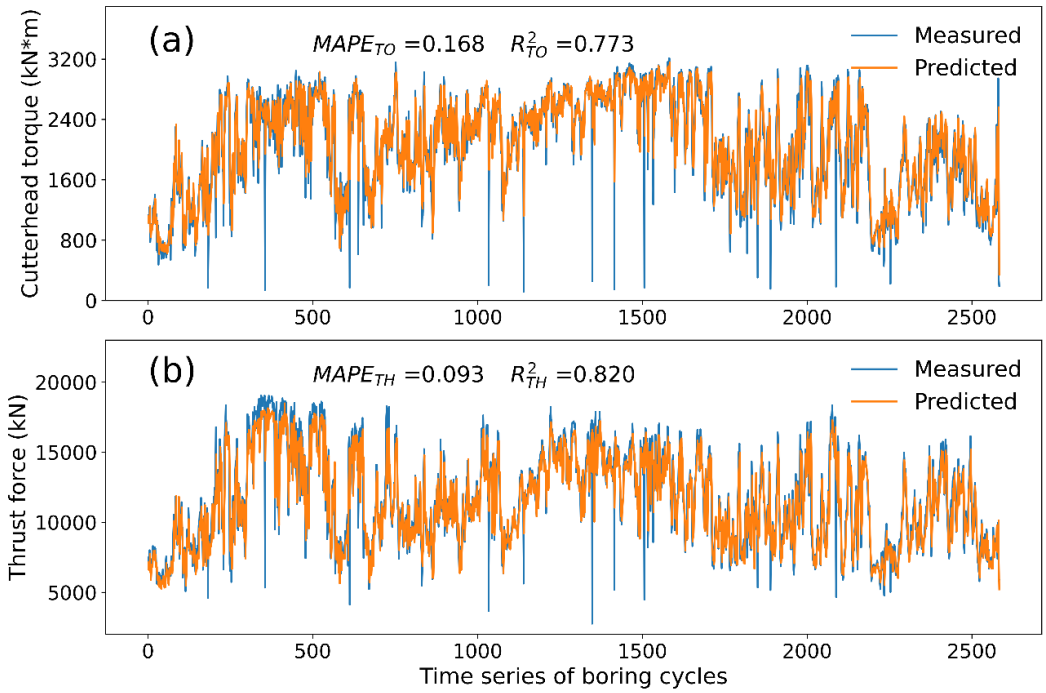


Figure 7.3 Measured and predicted results for time series forecasting of (a) cutterhead torque and (b) thrust force in the RNN model

7.4.2 Comparison between deep learning algorithms

LSTM and GRU are two popular variants of RNNs designed to address the vanishing gradient problem, allowing them to capture long-term dependencies in time series more effectively. To compare these with the RNN model, LSTM and GRU models are built for one-step forecasts. In Table 7.2, MAPEs for torque are 0.168, 0.170, and 0.167, and corresponding R^2 of 0.773, 0.770, and 0.773 in the RNN, LSTM, and GRU models, respectively. MAPEs for thrust are 0.093, 0.091, and 0.093, and corresponding R^2 of 0.820, 0.820, and 0.821 in the RNN, LSTM, and GRU models, respectively. These results reveal negligible differences among these algorithms, suggesting that long-term

information has limited influence on future TBM performance, supported by sensitivity analysis in Section 7.5. It implies that the time series of TBM performance behaves like a random walk (Shan et al., 2022), lacking long-term trends and periodicity.

Table 7.2 Comparison between RNN, LSTM, and GRU for cutterhead torque and thrust force forecasts

Algorithms	**TO**		**TH**	
	MAPE	**R^2**	**MAPE**	**R^2**
RNN	0.168	0.773	0.093	0.820
LSTM	0.170	0.770	0.091	0.820
GRU	0.167	0.773	0.093	0.821

7.4.3 Setting value and operational parameter-based model

PR_set and RPM_set are manually set on the control panel, directly influencing the actual penetration rate and revolutions in TBM tunnelling. When predicting the next-step TBM performance, the setting values for the next step are already known beforehand. As setting values are aware context, the proposed model is referred to as the advanced RNN model in Eq. 7.4.

$$\mathrm{TO}_{t+1}, \mathrm{TH}_{t+1} = f(\{\mathrm{PR}_t, \mathrm{TBM}_t, \mathrm{TO}_t, \mathrm{TH}_t\}, \mathrm{PR\_set}_{t+1}, \mathrm{RPM\_set}_{t+1}) \qquad 7.3$$

In the training process, the advanced RNN model reconfigures the inputs into two components: sequential inputs of historical data and non-sequential inputs of setting values. The historical data (PR, RPM, TO, and TH) are fed into the RNN cell, producing hidden features. At the same time, the setting values (PR_set and RPM_set) are fully

connected to extract hidden features. These hidden features are concatenated in a fully connected layer, ultimately producing the next-step torque and thrust.

The optimal combination of hyperparameters for the advanced RNN model includes a time step of 7, a number of layers of 2, a learning rate of 0.001, a hidden size of 32 for historical operational parameters, a hidden size of 32 for setting values, a dropout rate of 0, and a batch size of 32. In Figure 7.4, measured and predicted results are illustrated, having improved MAPE of 0.150 and R^2 of 0.808 for torque and improved MAPE of 0.088 and R^2 of 0.852 for thrust.

Figure 7.4 shows testing performance for one-step forecasts employing the setting value-based model, operational parameter-based model, and setting value and operational parameter-based model. In comparisons, the setting value and operational parameter-based model, or advanced RNN model, demonstrates superior performance, with a reduction in MAPE from 0.335 and 0.168 to 0.150 for torque and a reduction in MAPE from 0.193 and 0.093 to 0.088 for thrust in Figure 7.5(a). Similarly, for R^2, the advanced RNN model significantly improves in both cutterhead torque and thrust force forecasts, as depicted in Figure 7.5(b).

In general, the accuracy of forecasting thrust force is higher than that of forecasting cutterhead torque in the models. This discrepancy can be attributed to the fact that the cutterhead torque time series exhibits more rapid changes than the thrust force. These

rapidly changing points, or minimal features, can often be overlooked by machine learning models, leading to less accurate predictions.

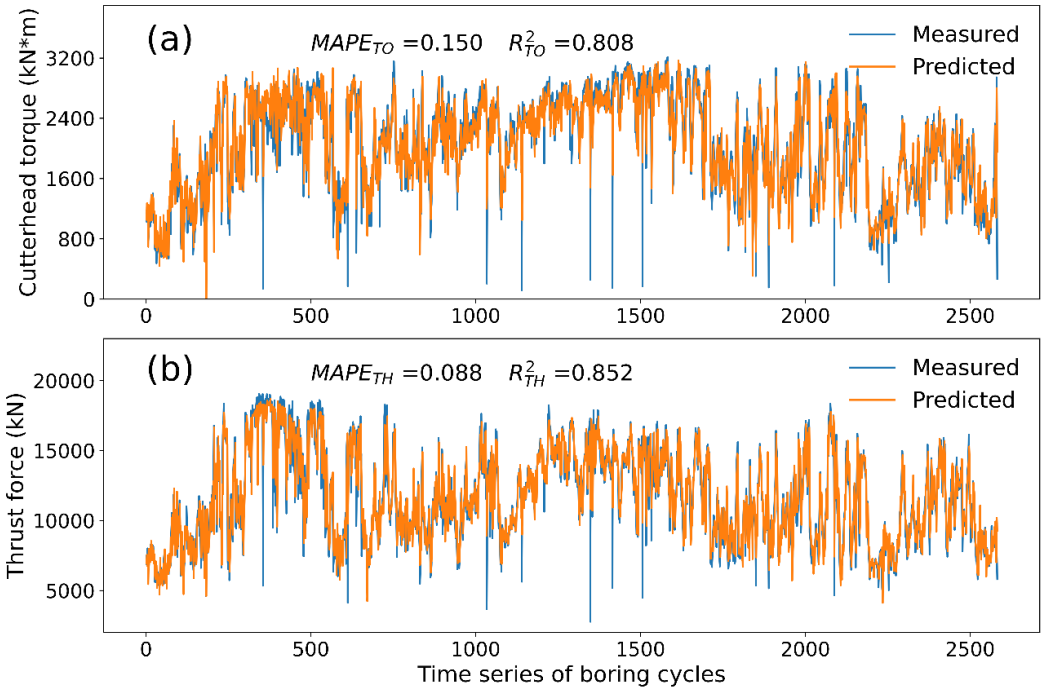


Figure 7.4 Measured and predicted results for time series forecasting of (a) cutterhead torque and (b) thrust force in the advanced RNN model

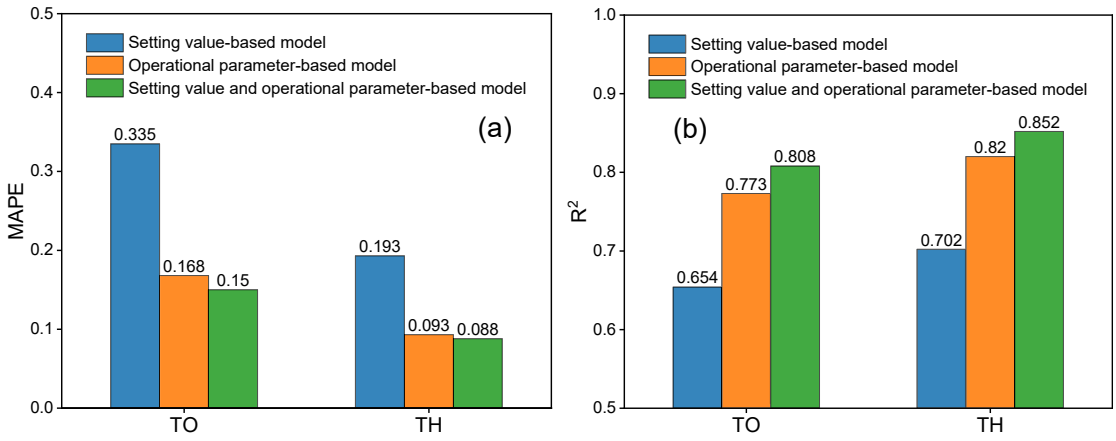


Figure 7.5 Comparing models on input parameters in terms of (a) MAPE and (b) R^2

7.5 Sensitivity analysis

From a quantitative perspective, the Sobol method is employed to analyse the sensitivity of the input parameters and observe their variations in model output (Sobol, 1990). The near-optimal advanced RNN model accounts for four operational parameters of the last seven steps and the two setting values for the next step. We investigate the importance of the input parameters, where the total Sobol index equals the sum of the Sobol indexes across the last seven steps. When forecasting cutterhead torque on the left side of Figure 7.6, RPM_set is the most important parameter with a Sobol index of 0.4684, while the historical RPM contributes merely 0.0565 to the index. TO is the second-order sensitive parameter with a Sobol index of 0.2420. For forecasting thrust force on the right side of Figure 7.6, TH itself exerts a significant influence on the next-step thrust with a Sobol index of 0.4675, followed by RPM_set with an index of 0.3540. It is found that TBM performance in the next step is greatly related to the setting values, especially RPM_set.

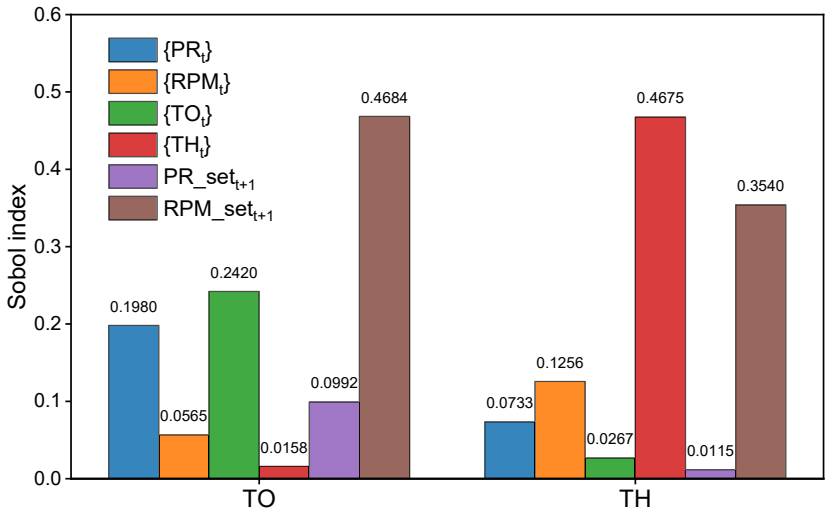


Figure 7.6 Sobol index of input parameters in advanced RNN model

To analyse the effect of time steps, heat maps of the Sobol index for torque and thrust are presented in Figures 7.7(a) and 7.7(b), respectively. The Sobol indexes from the last seven-step to three-step are below 0.01, suggesting their minimal influence on future TBM performance. This indicates that operational parameters further away from the current cutterhead have negligible effects, thus explaining why LSTM and GRU, which are capable of long-term dependencies, perform equivalently to RNN. The Sobol index of the last-step operational parameters outperforms that of other time steps. In detail, for forecasting cutterhead torque, the last-step thrust has the highest Sobol index of 0.23, while for forecasting thrust force, the last-step thrust force scores the highest Sobol index of 0.44.

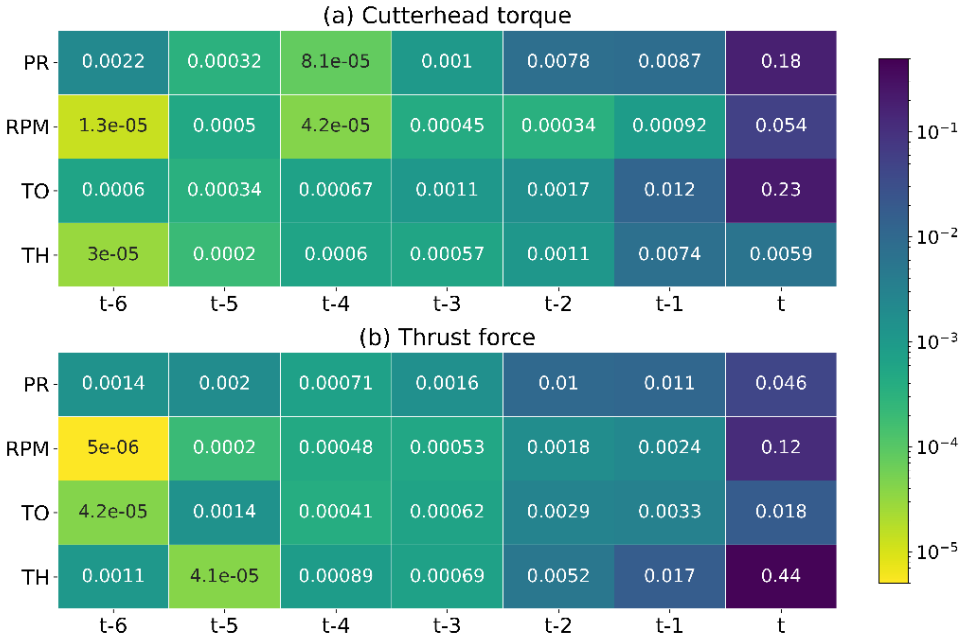


Figure 7.7 Heatmaps of Sobol index for forecasting (a) cutterhead torque and (b) thrust force in the advanced model

7.6 Summary

Chapter 7 utilises deep learning algorithms, especially the RNN model, to simultaneously predict cutterhead torque and thrust force. The dataset, derived from the Yinsong Water Diversion Project, is preprocessed into low-frequency data where each sample represents the average of a stable boring cycle. The input parameters in the modelling process encompass operational parameters of PR, RPM, TH, and TO, geological parameters of HC and FZ, and setting values of PR_set and RPM_set.

Geological parameters, HC and FZ, fail to fit TBM performance due to their sparse sampling. Considering PR_set and RPM_set are predetermined, the setting values-based model is able to fit TO and TH that align with the general trends. It shows that setting values have an influence on the performance of torque and thrust.

In terms of time series forecasting, RNN, LSTM, and GRU models that use historical operational parameters as input vectors are compared. LSTM and GRU perform equivalently to RNN because operational parameters distant from the current cutterhead have negligible effects on future TBM performance.

We propose an innovative advanced RNN model, which integrates the historical

operational parameters and current setting values. This model exhibits significant improvement in both cutterhead torque and thrust force forecasts. In sensitivity analysis, RPM_set plays an essential role in both cutterhead torque and thrust force forecasts. Additionally, the last-step thrust and torque themselves exert a significant influence on their respective future values.

Chapter 8 Rock Mass Classification

8.1 Introduction

The Yinsong Water Diversion Project was excavated using an open-type TBM to transfer water, with a total distance of 17.50 km. The tunnel lithology consists mainly of limestone, diorite, tuff, and granite, and rock mass classification is categorised into five classes, namely hydropower classification (HC), as detailed in Table 4.6. Unknown rock mass classification may lead to inappropriate operation and even low safety and efficiency of excavation. As a result, an accurate and reliable prediction method of rock mass classification is required in TBM tunnelling.

In this Chapter, we apply machine learning algorithms to classify rock masses of HC using operational data such as PR, RPM, TO, and TH in Eq. 8.1. The dataset, derived from the Yinsong water diversion project, is preprocessed into low-frequency data in Section 4.3.2, where each sample represents the average of a stable boring cycle. In Table 8.1, Model 8.1 is an RF classifier trained on labelled data, Model 8.2 incorporates the oversampling technique, SMOTE, to ensure balanced classes during training, and Model 8.3 leverages both labelled and unlabelled data by an RF-based self-training classifier.

$$\mathrm{HC} = f(\mathrm{PR}, \mathrm{RPM}, \mathrm{TO}, \mathrm{TH}) \qquad 8.1$$

Table 8.1 Model performance for rock mass classification

Model	Method	HC	PRC	REC	F1
8.1	RF	II	0.25	0.2	0.222
		III	0.585	0.6	0.593
		IV	0.615	0.6	0.608
		V	0.667	0.8	0.723
		Accuracy			0.589
8.2	SMOTE-RF	II	0.286	0.4	0.333
		III	0.625	0.5	0.556
		IV	0.605	0.65	0.627
		V	0.5	0.8	0.615
		Accuracy			0.578
8.3	RF-based self-training	II	1.0	0.2	0.333
		III	0.703	0.65	0.675
		IV	0.6	0.75	0.667
		V	1	0.4	0.571
		Accuracy			0.656

8.2 Data processing

During TBM excavation, the rock mass classification serves as a guide for operators to adjust operational parameters. Based on the dataset from the Songhua water diversion project, several studies have endeavoured to predict hydropower classification (HC) using advanced classifiers (Liu et al., 2020b; Bo et al., 2022; Hou et al., 2022). However, a prominent limitation of rock mass classification lies in the approximate fit of the HC into other segments. As depicted in Figure 8.1, rock mass classification is assumed to be consistent between two boreholes, even when they are spaced over 200m apart. Such

fitted labels are coarse classification as inputs in Chapter 7 but might not accurately represent the actual geological conditions as outputs, leading to degraded model performance and difficulty in evaluation.

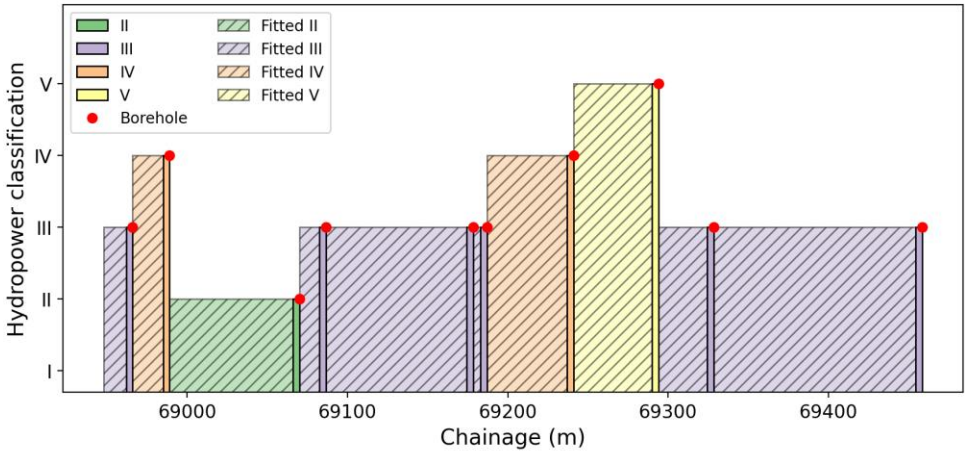


Figure 8.1 Fitted and actual rock mass classification in the Yinsong water diversion tunnel

It is assumed that the rock mass classification in the vicinity of boreholes (up to 2m) is relatively accurate, while the classification beyond this range remains ambiguous. Following this premise, 447 samples have been accurately labelled around 275 boreholes, leaving the remaining 12515 samples unlabelled. Figure 8.2 shows a pie chart detailing the distribution of four classes: 6.5% for class II, 39.4% for class III, 47.2% for class IV, and 6.9% for class V. Notably, the majority of the rock mass, 86.6%, falls between classes III and IV. Since the fitted label is likely to indicate a wrong rock mass classification, the actual labels can often be inferred from operational parameters.

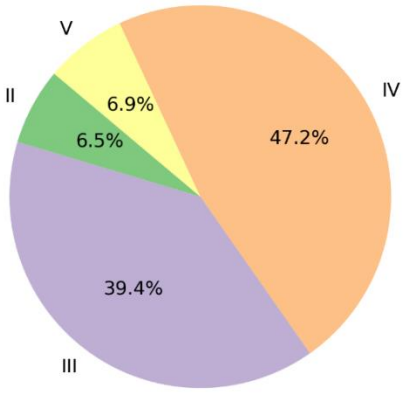


Figure 8.2 Pie chart on rock mass classification proportion

Likewise, input features in classification involve operational parameters such as PR, RPM, TO, and TH. To account for discrepancies in units and magnitudes among these parameters, min-max normalisation is applied in Eq. 5.2, scaling the low-frequency data to fall between 0 and 1.

Concurrently, the outputs, rock mass classification, are numerically encoded. Of the 447 labelled data, a random 80% is allocated for training, while the remaining 20% is reserved for testing. Figures 8.3(a) and 8.3(b) display counts of observations for each class in the training and test sets, respectively. Notably, the distributions are consistent between the training and test data, with classes III and IV being the most prevalent.

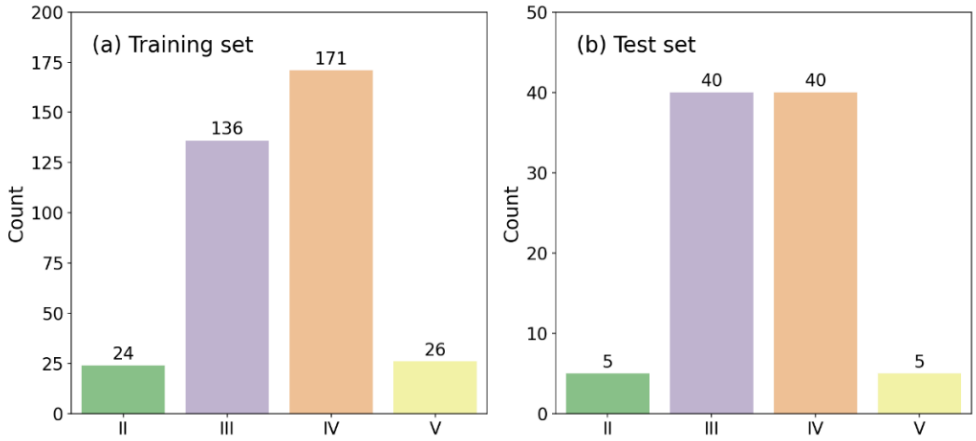


Figure 8.3 Counts of rock mass classification in (a) training and (b) test

8.3 Rock mass classification in supervised learning

8.3.1 Random forest classifier

Supervised learning requires labelled data to train machine learning models, while unsupervised learning works with the inherent input features without labelled output. In this Section, we employ 357 labelled data in supervised learning and build a near-optimal RF classifier, adding 5-fold cross-validation to avoid overfitting. A series of 30 random trials are conducted to find the optimal hyperparameters: 50 trees, a maximum tree depth of 10, a bifurcation criterion of at least two samples for internal nodes, and a single sample minimum at leaf nodes.

Figure 8.4 presents a confusion matrix between predicted and actual labels in the test, yielding an average accuracy of 0.589. In detail, PRC, REC, and F1 metrics are 0.25, 0.2, 0.222 for class II; PRC, REC, and F1 metrics are 0.585, 0.6, 0.593 for class III; PRC,

REC, and F1 metrics are 0.615, 0.6, 0.608 for class IV; PRC, REC, and F1 metrics are 0.667, 0.8, 0.723 for class V. Model performance is poor in minority classes, especially in class II, where only one out of five instances is correctly predicted (REC = 0.2).

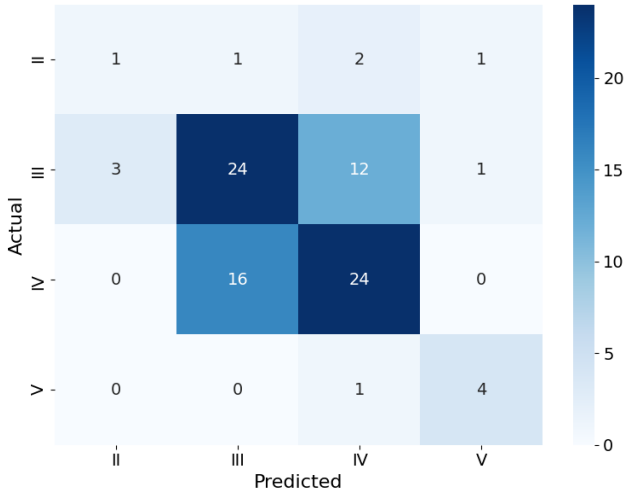


Figure 8.4 Confusion matrix of RF classifier in test

8.3.2 SMOTE-random forest classifier

The Synthetic Minority Over-sampling Technique (SMOTE) is a technique used to oversample the minority class in classification problems with imbalanced datasets. When using SMOTE, apply it to the training data and leave the test data untouched to get an unbiased estimate of the model performance on unseen data. The minority of classes II and V are generated by randomly selecting 5 neighbours, reaching 100 samples after oversampling in Figure 8.5.

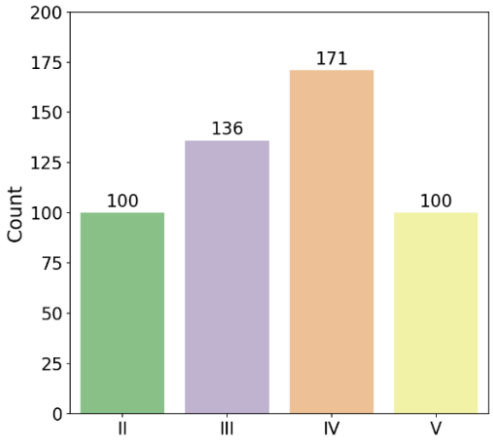


Figure 8.5 Counts of rock mass classification after oversampling in training

Oversampled train data, 507 samples in total, are used to build a SMOTE-RF classifier. In evaluation, the SMOTE-RF classifier is assessed by the same labelled test data, with a confusion matrix in Figure 8.6. Detailed PRE, REC, and F1 metrics are provided in Table 8.1, and the average accuracy is 0.578. Interestingly, the accuracy of the SMOTE-RF classifier performs worse compared with the RF classifier (0.578 <0.589). Oversampling can lead to improved accuracy for the minority class at the potential cost of reduced accuracy for other classes. The effect of SMOTE on model accuracy is problem-dependent and should be thoughtful about the problem at hand.

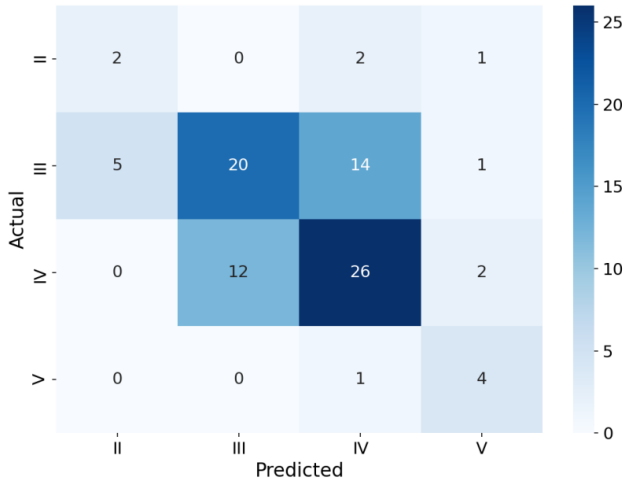


Figure 8.6 Confusion matrix of SMOTE-RF classifier in test

8.4 Rock mass classification in semi-supervised learning

Semi-supervised learning, a hybrid machine learning algorithm, utilises both labelled and unlabelled data for training, leveraging the unlabelled data to achieve better model performance. From the dataset, 357 data are labelled and allocated for initial training, 12,515 data remain unlabelled for generating pseudo-labels, and 90 labelled data are served for testing.

In self-training, an initial model is trained on the labelled data using a random forest classifier as the base estimator. Subsequent to this initial phase, the classifier predicts labels for the unlabelled data. The samples from the unlabelled dataset, exhibiting prediction confidence (as measured by probability) exceeding a threshold of 0.75, are added into the training set with their pseudo-labels. This augmented training set, comprising original labelled data and high-confident pseudo-labelled data, forms the

basis for subsequent model training. This iterative process continues for a total of five cycles, culminating in an enhanced RF-based self-training classifier. It is noteworthy to mention that not every unlabelled sample might be used for self-training, especially if the model is not confident in its predictions for some of the samples. Table 8.2 shows the number of label and pseudo-labels and their accuracies throughout each iteration, where iteration 0 indicates the base estimator. By the fifth iteration, a significant proportion of the unlabelled data, 11,530 samples, have been assimilated as confident pseudo-labels. The average accuracy of the classifiers witnesses an increment from 0.589 to 0.656.

Table 8.2 Number of trained data in each iteration

Iteration	**Number of trained samples**	**Accuracy**
0	357 labels	0.589
1	357 labels +4150 pseudo-labels	0.589
2	357 labels +8422 pseudo-labels	0.622
3	357 labels +10308 pseudo-labels	0.6
4	357 labels +11223 pseudo-labels	0.611
5	357 labels +11530 pseudo-labels	0.656

Based on the confusion matrix in Figure 8.7, classification metrics are calculated: PRC, REC, and F1 metrics are 1.0, 0.2, 0.333 for class II; PRC, REC, and F1 metrics are 0.703, 0.65, 0.675 for class III; PRC, REC, and F1 metrics are 0.6, 0.75, 0.667 for class IV; PRC, REC, and F1 metrics are 1.0, 0.4, 0.571 for class V.

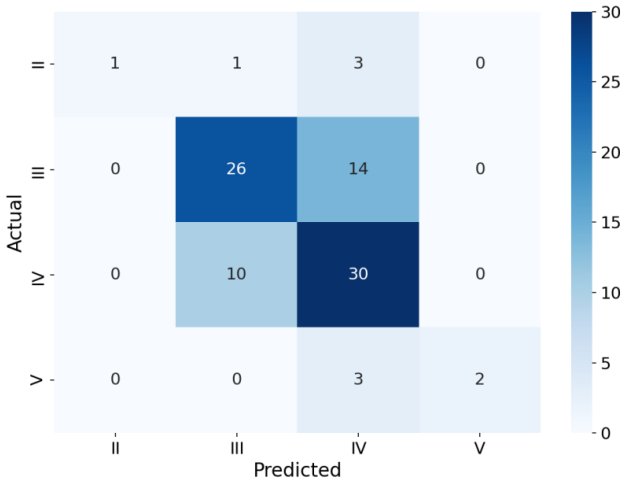


Figure 8.7 Confusion matrix of RF-based self-training classifier in test

8.5 Summary

In this Chapter, both supervised and semi-supervised learning techniques are employed to classify the rock mass classification of HC. The dataset, derived from the Yinsong Water Diversion Project, undergoes preprocessing to produce low-frequency data, with each sample representing the average value of a stable boring cycle. Operational parameters, specifically PR, RPM, TH, and TO, are included as input features for the modelling process.

An RF classifier, optimised by random search, is built using limited labelled data with acceptable accuracy. An alternative method employs the SMOTE-RF classifier, which aims to balance the dataset by oversampling the minority classes, though this yields no substantial performance improvement.

The self-training method offers a strategy to harness unlabelled data by generating confident pseudo-labels, facilitating enhanced training processes. Remarkably, by the fifth iteration of the RF-based self-training classifier, accuracy is significantly improved, rising from 0.589 to 0.656.

Chapter 9 Conclusion and Recommendations

9.1 Conclusion

Performance forecast and regression of tunnel boring machines remain a cornerstone of modern tunnelling, and this thesis offers an in-depth exploration of the topic. Through a detailed literature review, we navigated the multidimensional landscape of TBM performance regression, time series forecasting, and rock mass classification. In the thesis, employed methodologies for data processing, machine learning, evaluation metrics and sensitivity analysis are elaborated. Four TBM datasets—the Pahang-Selangor raw water tunnel, Changsha metro line, Zhengzhou metro line, and Yinsong water diversion tunnel—are summarized with data size and geological and operational parameters. Using machine learning algorithms, we build a causal relationship in penetration rate regression in Chapter 5. In Chapter 6, the near future (0–7.5 m) penetration rate is forecasted in different geological conditions. In Chapter 7, we forecast cutterhead torque and thrust force in the next boring cycle, approximately 1.14 metres ahead of the cutterhead. We conduct rock mass classification by semi-supervised learning in Chapter 8.

9.1.1 Conclusion on operational efficiency

In the realm of TBM performance regression, the RF model is the most suitable for predicting penetration rate with accuracy compared with the SVM and ANN models.

Regarding input parameters, the model with all five geological parameters is the best. In addition, the challenge is not merely about prediction but about achieving practical and actionable insights. While models encompassing both operational and geological parameters may score high on accuracy, their applicability in real-world scenarios is questioned due to the inaccessibility of operational parameters during the training process.

The employment of deep learning models for predicting penetration rates in the future is valuable. In univariate models, one-step forecasts have exhibited commendable robustness and confidence. In contrast, as the forecast horizon extends in multi-step forecasts, accuracy diminishes, indicating inherent challenges in long-term forecasts. Utilising a baseline, where the last input is copied as the prediction, offers moderately reliable results. This is particularly evident in the one-step forecast and highlights the relevance of the random walk theory. While multivariate RNN models incorporate both geological and operational parameters, their performance is not as optimal as univariate RNN models. Improvements, especially in the inputs into a time series of penetration rate and last-step other parameters, are achieved in the advanced RNN model. In sensitivity analysis, time series penetration rate emerges as the most crucial determinant, while geological conditions exert minimal influence on time series forecasting. Data smoothing can improve accuracy by negating noise. However, the balance is vital, as excessive simplification can lead to the loss of actual data characteristics.

9.1.2 Conclusion on safety

We focus on using deep learning algorithms to predict the cutterhead torque and thrust force simultaneously in the low-frequency data. Geological parameters HC and FZ, due to their infrequent sampling, do not successfully correlate with the cutterhead torque and thrust force. When comparing RNN, LSTM, and GRU models for time series forecasting, no significant advantage is found in using the more complex LSTM or GRU models over RNN. The advanced RNN model, which corporates historical operational parameters and current setting values (specifically PR_set and RPM_set), showcases a marked improvement in forecasting the cutterhead torque and thrust force. In sensitivity analysis, RPM_set is a vital parameter influencing forecasts for both cutterhead torque and thrust force. Furthermore, the last-step thrust and torque also significantly determine their upcoming values.

Both supervised and semi-supervised learning techniques are used to classify the rock mass classification. An RF classifier, fine-tuned using random search, delivers acceptable accuracy even with limited labelled data. While the SMOTE-RF classifier is designed to balance the dataset by oversampling minority classes, it does not offer substantial improvements in classification performance. The self-training method allows for leveraging unlabelled data. By generating confident pseudo-labels, the RF-based self-training classifier enhances the training process and reaches a significant jump in

accuracy by the fifth iteration.

9.2 Recommendation

Though powerful, the lack of interpretability in machine learning, known as the black box problem, poses a significant limitation. Machine learning excels at capturing complex patterns and relationships in data, but understanding the rationale behind their decisions can be challenging. For example, sensitivity analysis can provide insights into the importance of input features.

Another crucial challenge lies in the limited applicability of machine learning models developed and validated on different datasets. To gain confidence in the industry, it is necessary to validate and generalise these models across different datasets. The essence of this challenge is the lack of publicly available TBM operational data in countries like Australia: very few companies are willing to share and publish these data due to commercial interest. As tunnelling data becomes more accessible, using larger data for training can further improve the reliability and robustness of machine learning models in future projects, providing valuable feedback to the industry.

The thesis underscores the synergy between data processing, machine learning, and domain knowledge. Various factors, from geological parameters to machine settings, influence TBM performance. However, the key lies in skillfully applying advanced

models in tandem with domain expertise. The field is primed for further study, hinting at the development of more advanced models that can smoothly handle the intricacies of TBM operations. Future research could pivot to hybrid models, enhance data processing methods, or even explore real-time adaptive systems, ensuring that TBM performance regression remains an ever-evolving and exciting frontier.

References

Armaghani, D. J., Koopialipoor, M., Marto, A. & Yagiz, S. (2019) Application of several optimization techniques for estimating TBM advance rate in granitic rocks. *Journal of Rock Mechanics and Geotechnical Engineering* **11(4)**:779-789.

Armaghani, D. J., Mohamad, E. T., Narayanasamy, M. S., Narita, N. & Yagiz, S. (2017) Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. *Tunnelling and Underground Space Technology* **63**:29-43.

Armaghani, D. J., Yagiz, S., Mohamad, E. T. & Zhou, J. (2021) Prediction of TBM performance in fresh through weathered granite using empirical and statistical approaches. *Tunnelling and Underground Space Technology* **118**:104183.

Bai, X.-D., Cheng, W.-C. & Li, G. (2021) A comparative study of different machine learning algorithms in predicting EPB shield behaviour: a case study at the Xi'an metro, China. *Acta Geotechnica*:1-20.

Barton, N. (2000) *TBM tunnelling in jointed and faulted rock.* Crc Press.

Barton, N., Lien, R. & Lunde, J. (1974) Engineering classification of rock masses for the design of tunnel support. *Rock mechanics* **6**:189-236.

Benardos, A. & Kaliampakos, D. (2004) Modelling TBM performance with artificial neural networks. *Tunnelling and Underground Space Technology* **19(6)**:597-605.

Bengio, Y., Simard, P. & Frasconi, P. (1994) Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks* **5(2)**:157-166.

Bergstra, J. & Bengio, Y. (2012) Random search for hyper-parameter optimization. *Journal of machine learning research* **13(2)**.

Bieniawski, Z. T. (1989) *Engineering rock mass classifications: a complete manual for engineers and geologists in mining, civil, and petroleum engineering.* John Wiley & Sons.

Bo, Y., Liu, Q., Huang, X. & Pan, Y. (2022) Real-time hard-rock tunnel prediction model for rock mass classification using CatBoost integrated with Sequential Model-Based Optimization. *Tunnelling and Underground Space Technology* **124**:104448.

Box, G. E. & Pierce, D. A. (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association* **65(332)**:1509-1526.

Breiman, L. (2001) Random forests. *Machine learning* **45**:5-32.

Bruland, A. (1998) *Hard rock tunnel boring.* Norwegian University of Science and Technology, Trondheim, Norway.

Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. (2002) SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research* **16**:321-357.

Chen, R.-P., Zhang, P., Kang, X., Zhong, Z.-Q., Liu, Y. & Wu, H.-N. (2019) Prediction of

maximum surface settlement caused by earth pressure balance (EPB) shield tunneling with ANN methods. *Soils and Foundations* **59(2)**:284-295.

Chen, Z., Zhang, Y., Li, J., Li, X. & Jing, L. (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.

Erharter, G. H. & Marcher, T. (2021) On the pointlessness of machine learning based time delayed prediction of TBM operational data. *Automation in Construction* **121**:103443.

Feng, S., Chen, Z., Luo, H., Wang, S., Zhao, Y., Liu, L., Ling, D. & Jing, L. (2021) Tunnel boring machines (TBM) performance prediction: A case study using big data and deep learning. *Tunnelling and Underground Space Technology* **110**:103636.

Fu, X., Feng, L. & Zhang, L. (2022) Data-driven estimation of TBM performance in soft soils using density-based spatial clustering and random forest. *Applied Soft Computing* **120**:108686.

Fu, X., Wu, M., Ponnarasu, S. & Zhang, L. (2023) A hybrid deep learning approach for dynamic attitude and position prediction in tunnel construction considering spatio-temporal patterns. *Expert Systems with Applications* **212**:118721.

Fu, X. & Zhang, L. (2021) Spatio-temporal feature fusion for real-time prediction of TBM operating parameters: A deep learning approach. *Automation in Construction* **132**:103937.

Gao, B., Wang, R., Lin, C., Guo, X., Liu, B. & Zhang, W. (2021) TBM penetration rate prediction based on the long short-term memory neural network. *Underground space* **6(6)**:718-731.

Gao, M.-Y., Zhang, N., Shen, S.-L. & Zhou, A. (2020) Real-time dynamic earth-pressure regulation model for shield tunneling by integrating GRU deep learning method with GA optimization. *IEEE Access* **8**:64310-64323.

Gao, X., Shi, M., Song, X., Zhang, C. & Zhang, H. (2019) Recurrent neural networks for real-time prediction of TBM operating parameters. *Automation in Construction* **98**:225-235.

Gomes, A. R., Chapman, B., Chapman, N. & Cortes, F. (2021) Development of the geotechnical baseline report for the snowy 2.0 pumped storage project. In *17th Australasian Tunnelling Conference-2020+ 1: 2020 Vision: Innovating the next 50 years (ATS 2020) Innovating the next 50 years (ATS 2020)*.) Australasian Institute of Mining and Metallurgy Melbourne, pp. 983-1000.

Gong, Q. & Zhao, J. (2009) Development of a rock mass characteristics model for TBM penetration rate prediction. *International Journal of Rock mechanics and mining Sciences* **46(1)**:8-18.

Grima, M. A., Bruines, P. & Verhoef, P. (2000) Modeling tunnel boring machine performance by neuro-fuzzy methods. *Tunnelling and Underground Space Technology* **15(3)**:259-269.

Guo, D., Li, J., Jiang, S.-H., Li, X. & Chen, Z. (2022) Intelligent assistant driving method for tunnel boring machine based on big data. *Acta Geotechnica* **17(4)**:1019-1030.

Hamidi, J. K., Shahriar, K., Rezai, B. & Rostami, J. (2010) Performance prediction of hard rock TBM using Rock Mass Rating (RMR) system. *Tunnelling and Underground Space Technology* **25(4)**:333-345.

Hassanpour, J., Rostami, J. & Zhao, J. (2011) A new hard rock TBM performance prediction model for project planning. *Tunnelling and Underground Space Technology* **26(5)**:595-603.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J. & Scholkopf, B. (1998) Support vector machines. *IEEE Intelligent Systems and their applications* **13(4)**:18-28.

Hochreiter, S. & Schmidhuber, J. (1997) Long short-term memory. *Neural computation* **9(8)**:1735-1780.

Hoeffding, W. (1992) A class of statistics with asymptotically normal distribution. In *Breakthroughs in statistics*.) Springer, pp. 308-334.

Hou, S., Liu, Y. & Yang, Q. (2022) Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning. *Journal of Rock Mechanics and Geotechnical Engineering* **14(1)**:123-143.

Huang, X., Zhang, Q., Liu, Q., Liu, X., Liu, B., Wang, J. & Yin, X. (2022) A real-time prediction method for tunnel boring machine cutter-head torque using bidirectional long short-term memory networks optimized by multi-algorithm. *Journal of Rock Mechanics and Geotechnical Engineering* **14(3)**:798-812.

Innaurato, N., Mancini, A., Rondena, E. & Zaninetti, A. (1991) Forecasting and effective TBM performances in a rapid excavation of a tunnel in Italy. In *ISRM Congress*.) ISRM, pp. ISRM-7CONGRESS-1991-200.

Jain, A. K., Mao, J. & Mohiuddin, K. M. (1996) Artificial neural networks: A tutorial. *Computer* **29(3)**:31-44.

Javad, G. & Narges, T. (2010) Application of artificial neural networks to the prediction of tunnel boring machine penetration rate. *Mining Science and Technology (China)* **20(5)**:727-733.

Jung, J.-H., Chung, H., Kwon, Y.-S. & Lee, I.-M. (2019) An ANN to predict ground condition ahead of tunnel face using TBM operational data. *KSCE Journal of Civil Engineering* **23**:3200-3206.

Kannangara, K. P. M., Zhou, W., Ding, Z. & Hong, Z. (2022) Investigation of feature contribution to shield tunneling-induced settlement using Shapley additive explanations method. *Journal of Rock Mechanics and Geotechnical Engineering* **14(4)**:1052-1063.

Koopialipoor, M., Fahimifar, A., Ghaleini, E. N., Momenzadeh, M. & Armaghani, D. J. (2020) Development of a new hybrid ANN for solving a geotechnical problem related to tunnel boring machine performance. *Engineering with Computers* **36**:345-357.

Lau, S.-C., Lu, M. & Ariaratnam, S. T. (2010) Applying radial basis function neural networks to estimate next-cycle production rates in tunnelling construction. *Tunnelling and Underground Space Technology* **25(4)**:357-365.

Leu, S.-S. & Adi, T. J. W. (2011) Probabilistic prediction of tunnel geology using a Hybrid Neural-HMM. *Engineering Applications of Artificial Intelligence* **24(4)**:658-665.

Li, J., Li, P., Guo, D., Li, X. & Chen, Z. (2021) Advanced prediction of tunnel boring machine performance based on big data. *Geoscience Frontiers* **12(1)**:331-338.

Li, S., Liu, B., Xu, X., Nie, L., Liu, Z., Song, J., Sun, H., Chen, L. & Fan, K. (2017) An overview of ahead geological prospecting in tunneling. *Tunnelling and Underground Space Technology* **63**:69-94.

Lin, S.-S., Shen, S.-L. & Zhou, A. (2022a) Real-time analysis and prediction of shield cutterhead torque using optimized gated recurrent unit neural network. *Journal of Rock Mechanics and Geotechnical Engineering* **14(4)**:1232-1240.

Lin, S., Zhang, N., Zhou, A. & Shen, S. (2022b) Time-series prediction of shield movement performance during tunneling based on hybrid model. *Tunnelling and Underground Space Technology* **119**:104245.

Lin, Y. (1999) Study of BQ formula in national standard of quantitative classification for basic quality of rock mass. *Chinese Journal of Geotechnical Engineering* **21(4)**:481-485.

Liu, B., Wang, R., Guan, Z., Li, J., Xu, Z., Guo, X. & Wang, Y. (2019) Improved support vector regression models for predicting rock mass parameters using tunnel boring machine driving data. *Tunnelling and Underground Space Technology* **91**:102958.

Liu, B., Wang, R., Zhao, G., Guo, X., Wang, Y., Li, J. & Wang, S. (2020a) Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm. *Tunnelling and Underground Space Technology* **95**:103103.

Liu, B., Wang, Y., Zhao, G., Yang, B., Wang, R., Huang, D. & Xiang, B. (2021a) Intelligent decision method for main control parameters of tunnel boring machine based on multi-objective optimization of excavation efficiency and cost. *Tunnelling and Underground Space Technology* **116**:104054.

Liu, F. T., Ting, K. M. & Zhou, Z.-H. (2008) Isolation forest. In *2008 eighth ieee international conference on data mining*.) IEEE, pp. 413-422.

Liu, Q., Wang, X., Huang, X. & Yin, X. (2020b) Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data. *Tunnelling and Underground Space Technology* **106**:103595.

Liu, Z., Li, L., Fang, X., Qi, W., Shen, J., Zhou, H. & Zhang, Y. (2021b) Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network. *Automation in Construction* **125**:103647.

Mahdevari, S., Shahriar, K., Yagiz, S. & Shirazi, M. A. (2014) A support vector regression model for predicting tunnel boring machine penetration rates. *International*

Journal of Rock mechanics and mining Sciences **72**:214-229.

Mokhtari, S. & Mooney, M. A. (2020) Predicting EPBM advance rate performance using support vector regression modeling. *Tunnelling and Underground Space Technology* **104**:103520.

Ozdemir, L. (1977) Development of theoretical equations for predicting tunnel boreability.) Colorado School of Mines.

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

Pourtaghi, A. & Lotfollahi-Yaghin, M. (2012) Wavenet ability assessment in comparison to ANN for predicting the maximum surface settlement caused by tunneling. *Tunnelling and Underground Space Technology* **28**:257-271.

Qin, C., Shi, G., Tao, J., Yu, H., Jin, Y., Lei, J. & Liu, C. (2021) Precise cutterhead torque prediction for shield tunneling machines using a novel hybrid deep neural network. *Mechanical Systems and Signal Processing* **151**:107386.

Qin, C., Shi, G., Tao, J., Yu, H., Jin, Y., Xiao, D., Zhang, Z. & Liu, C. (2022) An adaptive hierarchical decomposition-based method for multi-step cutterhead torque forecast of shield machine. *Mechanical Systems and Signal Processing* **175**:109148.

Qu, T., Di, S., Feng, Y., Wang, M. & Zhao, T. (2021) Towards data-driven constitutive modelling for granular materials via micromechanics-informed deep learning. *International Journal of Plasticity* **144**:103046.

Roberts, S. (2000) Control chart tests based on geometric moving averages. *Technometrics* **42(1)**:97-101.

Rostami, J. (1997) *Development of a force estimation model for rock fragmentation with disc cutters through theoretical modeling and physical measurement of crushed zone pressure.* Colorado School of Mines Golden.

Rostami, J. & Ozdemir, L. (1993) A new model for performance prediction of hard rock TBMs. In *Proceedings of 1993 rapid excavation and tunneling conference*.), pp. 793-809.

Roxborough, F. F. & Phillips, H. R. (1975) Rock excavation by disc cutter. In *International journal of rock mechanics and mining sciences & geomechanics abstracts*.) Elsevier, vol. 12, pp. 361-366.

Salimi, A., Rostami, J. & Moormann, C. (2019) Application of rock mass classification systems for performance estimation of rock TBMs using regression tree and artificial intelligence algorithms. *Tunnelling and Underground Space Technology* **92**:103046.

Salimi, A., Rostami, J., Moormann, C. & Delisio, A. (2016) Application of non-linear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs. *Tunnelling and Underground Space Technology*

58:236-246.

Salimi, A., Rostami, J., Moormann, C. & Hassanpour, J. (2022) Introducing Tree-Based-Regression Models for Prediction of Hard Rock TBM Performance with Consideration of Rock Type. *Rock Mechanics and Rock Engineering*:1-23.

Sapigni, M., Berti, M., Bethaz, E., Busillo, A. & Cardone, G. (2002) TBM performance estimation using rock mass classifications. *International Journal of Rock mechanics and mining Sciences* **39(6)**:771-788.

Schuster, M. & Paliwal, K. K. (1997) Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing* **45(11)**:2673-2681.

Shan, F., He, X., Armaghani, D. J. & Sheng, D. (2023a) Effects of data smoothing and recurrent neural network (RNN) algorithms for real-time forecasting of tunnel boring machine (TBM) performance. *Journal of Rock Mechanics and Geotechnical Engineering*.

Shan, F., He, X., Armaghani, D. J., Zhang, P. & Sheng, D. (2022) Success and challenges in predicting TBM penetration rate using recurrent neural networks. *Tunnelling and Underground Space Technology* **130**:104728.

Shan, F., He, X., Armaghani, D. J., Zhang, P. & Sheng, D. (2023b) Response to Discussion on "Success and challenges in predicting TBM penetration rate using recurrent neural networks" by Georg H. Erharter, Thomas Marcher. *Tunnelling and Underground Space Technology*:105064.

Shan, F., He, X., Xu, H., Armaghani, D. J. & Sheng, D. (2023c) Applications of Machine Learning in Mechanised Tunnel Construction: A Systematic Review. *Eng* **4(2)**:1516-1535.

Shen, S.-L., Elbaz, K., Shaban, W. M. & Zhou, A. (2022) Real-time prediction of shield moving trajectory during tunnelling. *Acta Geotechnica* **17(4)**:1533-1549.

Shi, G., Qin, C., Tao, J. & Liu, C. (2021) A VMD-EWT-LSTM-based multi-step prediction approach for shield tunneling machine cutterhead torque. *Knowledge-Based Systems* **228**:107213.

Snowdon, R., Ryley, M. & Temporal, J. (1982) A study of disc cutting in selected British rocks. In *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*.) Elsevier, vol. 19, pp. 107-121.

Sobol, I. M. (1990) On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie* **2(1)**:112-118.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014) Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* **15(1)**:1929-1958.

Sun, W., Shi, M., Zhang, C., Zhao, J. & Song, X. (2018) Dynamic load prediction of tunnel boring machine (TBM) based on heterogeneous in-situ data. *Automation in Construction* **92**:23-34.

Suwansawat, S. & Einstein, H. H. (2006) Artificial neural networks for predicting the

maximum surface settlement caused by EPB shield tunneling. *Tunnelling and Underground Space Technology* **21(2)**:133-150.

Wang, L., Wang, L., Zhang, W., Meng, X., Liu, S. & Zhu, C. (2024) Time series prediction of reservoir bank landslide failure probability considering the spatial variability of soil properties. *Journal of Rock Mechanics and Geotechnical Engineering*.

Wang, R., Li, D., Chen, E. J. & Liu, Y. (2021a) Dynamic prediction of mechanized shield tunneling performance. *Automation in Construction* **132**:103958.

Wang, X., Zhu, H., Zhu, M., Zhang, L. & Ju, J. W. (2021b) An integrated parameter prediction framework for intelligent TBM excavation in hard rock. *Tunnelling and Underground Space Technology* **118**:104196.

Wu, Z., Wei, R., Chu, Z. & Liu, Q. (2021) Real-time rock mass condition prediction with TBM tunneling big data using a novel rock–machine mutual feedback perception method. *Journal of Rock Mechanics and Geotechnical Engineering* **13(6)**:1311-1325.

Xiong, W., Wang, J. & Wu, M. (2023) Data-driven constitutive modelling of granular soils considering multiscale particle morphology. *Computers and Geotechnics* **162**:105699.

Xu, C., Liu, X., Wang, E. & Wang, S. (2021) Prediction of tunnel boring machine operating parameters using various machine learning algorithms. *Tunnelling and Underground Space Technology* **109**:103699.

Yagiz, S. (2002) Development of rock fracture and brittleness indices to quantify the effects of rock mass features and toughness in the CSM model basic penetration for hard rock tunneling machines.

Yagiz, S. (2008) Utilizing rock mass properties for predicting TBM performance in hard rock condition. *Tunnelling and Underground Space Technology* **23(3)**:326-339.

Yagiz, S. (2009) Assessment of brittleness using rock strength and density with punch penetration test. *Tunnelling and Underground Space Technology* **24(1)**:66-74.

Yagiz, S. & Karahan, H. (2015) Application of various optimization techniques and comparison of their performances for predicting TBM penetration rate in rock mass. *International Journal of Rock mechanics and mining Sciences* **80**:308-315.

Yang, B., Yin, K., Lacasse, S. & Liu, Z. (2019) Time series analysis and long short-term memory neural network to predict landslide displacement. *Landslides* **16**:677-694.

Yang, H., Song, K. & Zhou, J. (2022a) Automated recognition model of geomechanical information based on operational data of tunneling boring machines. *Rock Mechanics and Rock Engineering* **55(3)**:1499-1516.

Yang, H., Wang, Z. & Song, K. (2022b) A new hybrid grey wolf optimizer-feature weighted-multiple kernel-support vector regression technique to predict TBM performance. *Engineering with Computers*:1-17.

Yang, J., Yagiz, S., Liu, Y.-J. & Laouafa, F. (2022c) Comprehensive evaluation of machine learning algorithms applied to TBM performance prediction.

Underground space **7(1)**:37-49.

Yu, H., Tao, J., Qin, C., Liu, M., Xiao, D., Sun, H. & Liu, C. (2022) A novel constrained dense convolutional autoencoder and DNN-based semi-supervised method for shield machine tunnel geological formation recognition. *Mechanical Systems and Signal Processing* **165**:108353.

Yu, H., Tao, J., Qin, C., Xiao, D., Sun, H. & Liu, C. (2021) Rock mass type prediction for tunnel boring machine using a novel semi-supervised method. *Measurement* **179**:109545.

Zeng, J., Roy, B., Kumar, D., Mohammed, A. S., Armaghani, D. J., Zhou, J. & Mohamad, E. T. (2021) Proposing several hybrid PSO-extreme learning machine techniques to predict TBM performance. *Engineering with Computers*:1-17.

Zhang, N., Zhang, N., Zheng, Q. & Xu, Y.-S. (2022) Real-time prediction of shield moving trajectory during tunnelling using GRU deep neural network. *Acta Geotechnica* **17(4)**:1167-1182.

Zhang, P., Chen, R., Dai, T., Wang, Z. & Wu, K. (2021a) An AIoT-based system for real-time monitoring of tunnel construction. *Tunnelling and Underground Space Technology* **109**:103766.

Zhang, P., Chen, R. & Wu, H. (2019a) Real-time analysis and regulation of EPB shield steering using Random Forest. *Automation in Construction* **106**:102860.

Zhang, P., Wu, H.-N., Chen, R.-P. & Chan, T. H. (2020a) Hybrid meta-heuristic and machine learning algorithms for tunneling-induced settlement prediction: A comparative study. *Tunnelling and Underground Space Technology* **99**:103383.

Zhang, P., Wu, H., Chen, R., Dai, T., Meng, F. & Wang, H. (2020b) A critical evaluation of machine learning and deep learning in shield-ground interaction prediction. *Tunnelling and Underground Space Technology* **106**:103593.

Zhang, P., Yin, Z.-Y. & Jin, Y.-F. (2021b) State-of-the-art review of machine learning applications in constitutive modeling of soils. *Archives of Computational Methods in Engineering* **28(5)**:3661-3686.

Zhang, Q., Hu, W., Liu, Z. & Tan, J. (2020c) TBM performance prediction with Bayesian optimization and automated machine learning. *Tunnelling and Underground Space Technology* **103**:103493.

Zhang, Q., Liu, Z. & Tan, J. (2019b) Prediction of geological conditions for a tunnel boring machine using big operational data. *Automation in Construction* **100**:73-83.

Zhang, X., Zhu, C., He, M., Dong, M., Zhang, G. & Zhang, F. (2021c) Failure mechanism and long short-term memory neural network model for landslide risk prediction. *Remote Sensing* **14(1)**:166.

Zhou, C., Xu, H., Ding, L., Wei, L. & Zhou, Y. (2019) Dynamic prediction for attitude and position in shield tunneling: A deep learning method. *Automation in Construction* **105**:102840.

Zhou, J., Qiu, Y., Armaghani, D. J., Zhang, W., Li, C., Zhu, S. & Tarinejad, R. (2021a) Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic techniques. *Geoscience Frontiers* **12(3)**:101091.

Zhou, J., Qiu, Y., Zhu, S., Armaghani, D. J., Khandelwal, M. & Mohamad, E. T. (2021b) Estimation of the TBM advance rate under hard rock conditions using XGBoost and Bayesian optimization. *Underground space* **6(5)**:506-515.