Title
Developing AI Systems for EPB TBM Utilizing Sensing Data and Machine Learning

Author
Apoji, Dayu

Publication Date
2023

Peer reviewed|Thesis/dissertation

Developing AI Systems for EPB TBM Utilizing Sensing Data and Machine Learning

by

Dayu Apoji

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Civil and Environmental Engineering

and the Designated Emphasis

in

Computational and Data Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kenichi Soga, Chair
Professor Dimitrios Zekkos
Professor Giles Hooker

Spring 2023

Developing AI Systems for EPB TBM Utilizing Sensing Data and Machine Learning

Abstract

Developing AI Systems for EPB TBM Utilizing Sensing Data and Machine Learning

by

Dayu Apoji

Doctor of Philosophy in Civil and Environmental Engineering

and the Designated Emphasis in

Computational and Data Science and Engineering

University of California, Berkeley

Professor Kenichi Soga, Chair

This dissertation presents a development of an integrated framework of artificial intelligence (AI) systems for earth pressure balance (EPB) tunnel boring machine (TBM) tunneling. The framework is constructed based on the feedback loop control system. The AI systems are developed using machine learning algorithms and structured to follow the human cognitive model, i.e., sensing, perceiving, and decision-making. The development of the systems is conducted in three parts.

The first part discusses the characteristics of TBM data and the effects of data preparation on data-driven models. This is achieved by (i) proposing a knowledge-based EPB TBM feature taxonomy and (ii) investigating the effects of data aggregation and feature selection on prediction models. The investigation shows that models developed using different data aggregation levels produce comparable prediction trends and similar feature importance rank in the conditions of sufficient observations and predictors. However, models with a coarser aggregation level may appreciate higher prediction performance due to the lower variance. The investigation also shows that the knowledge-guided TBM feature selection offers benefits over embedded machine learning-based feature selections. The developed model can produce relatively consistent feature importance in different tunneling cases, indicating better generalizability of the model.

The second part proposes AI systems that perceive tunneling environments in real-time based on the streams of sensor data during tunneling operation. This is achieved by developing (i) a supervised AI system to interpret the encountered geologic conditions, (ii) an unsupervised AI system to detect the encountered geologic anomalies, and (iii) a supervised AI system that connects TBM data to the ground monitoring data and estimates tunneling-induced

ground movements. The proposed geologic interpretation system uses either Random Forests (RF) classification or regression algorithms to infer the geologic transitions along the tunnel alignment. The proposed geologic anomaly detection system combines Principal Component Analysis (PCA) to project the data into a lower dimension space and Local Outlier Factor (LOF) to measure the degree of the anomaly of the projected data points. The proposed tunneling-induced ground movement estimation system uses RF regression to approximate any shape of ground movements solely based on TBM operation data and tunnel spatial geometries without prior assumptions on the ground movement shape, geologic material parameters, and the expected ground loss.

The third part proposes AI systems that model tunneling and its decision-making processes. This is achieved by introducing (i) a combination of probabilistic graph modeling and structure learning algorithms as a tool to systematically explore the causal effect interactions contained in TBM operation data and (ii) a multi-output supervised AI system to determine multiple steering control parameters simultaneously and steer the TBM along the tunnel alignment. The study shows that Bayesian Network Structure Learning (BNSL) can potentially be used to model the interactions of human operator decisions, TBM behaviors, and ground conditions in an integrated representation. The study also shows that the multivariate Random Forest (MRF) algorithm can concurrently make multiple decisions on TBM steering control parameters during driving.

This dissertation has demonstrated that TBM data contain valuable information that can be extracted to benefit tunneling operations. Due to the complex relationships within the data, the nonlinear and nonparametric machine learning models offer advantages over the conventional linear and parametric models. This development is envisioned to be a building block that advances the development of autonomous TBM technology.

To my family

Contents

Contents **ii**

List of Figures **viii**

List of Tables **xiii**

1 Introduction **1**
1.1 Background 1
1.2 Problem Statements 6
1.3 Objectives 8
1.4 Dissertation Outline 9

2 EPB TBM Tunneling **11**
2.1 Introduction 11
2.2 Historical Development 11
2.3 TBM Overview 13
2.4 EPB TBM Components 15
2.4.1 Cutter Head 15
2.4.2 Thrust 17
2.4.3 Articulation 18
2.4.4 Excavation Chamber 18
2.4.5 Ground Conditioning 18
2.4.6 Muck Discharge 19
2.4.7 Backfill Grouting 20
2.5 EPB TBM Tunneling Process 21
2.5.1 Performance Indicators 21
2.5.2 Tunneling Process Control 21
2.5.3 Data Management 26

3 Control System for EPB TBM Tunneling **28**
3.1 Introduction 28
3.2 Dynamical System 28

3.3 Feedback Control System . . . 29
3.4 Controller . . . 30
3.4.1 Linear Control . . . 31
3.4.2 Model-Based Control . . . 31
3.5 Intelligent Control System for TBM Tunneling . . . 33
3.5.1 Direction Control . . . 34
3.5.2 Earth Pressure Balance Control . . . 35
3.5.3 Integrated Control with Minimal Human Intervention . . . 36
3.5.4 Integrated Control with No Human Intervention . . . 36
3.6 Proposed Framework for EPB TBM . . . 37
3.6.1 Cognitive Model . . . 37
3.6.2 AI-Based System . . . 38

4 Machine Learning for EPB TBM Tunneling 40
4.1 Introduction . . . 40
4.2 Machine Learning . . . 40
4.2.1 Overview . . . 40
4.2.2 Algorithms and Selection . . . 42
4.2.3 Mathematical Basis . . . 44
4.2.4 Modeling Consideration . . . 48
4.3 Machine Learning for EPB TBM Tunneling Applications . . . 50
4.3.1 Geologic Interpretation . . . 50
4.3.2 Ground Movement Estimation . . . 51
4.3.3 Tunneling Process Modeling . . . 52

5 Effects of Data Aggregation on Prediction Models 53
5.1 Introduction . . . 53
5.1.1 Background . . . 53
5.1.2 Related Works . . . 55
5.1.3 Objectives . . . 56
5.2 Data . . . 56
5.2.1 Tunneling Cases . . . 56
5.2.2 Data Description and Preparation . . . 57
5.3 Methods . . . 58
5.3.1 Statistical Parameters . . . 63
5.3.2 Aggregation Ratio . . . 63
5.3.3 Model Setup . . . 64
5.3.4 Prediction Methods . . . 65
5.3.5 Model Hyperparameters . . . 66
5.3.6 Model Evaluation and Feature Importance . . . 66
5.4 Results and Discussion . . . 67
5.4.1 Statistical Parameter Changes Over Dynamic Sequence . . . 67

5.4.2 Effects on Statistical Parameters . . . 69
5.4.3 Effects on Model Predictability . . . 69
5.4.4 Data Reduction and Information Loss . . . 72
5.4.5 Effects on Model Computability . . . 73
5.4.6 Effects on Model Stability . . . 75
5.5 Conclusions . . . 80

6 Effects of Feature Selection on Prediction Models 82
6.1 Introduction . . . 82
6.1.1 Background . . . 82
6.1.2 Related Works . . . 83
6.1.3 Objectives . . . 84
6.2 Data . . . 84
6.2.1 Tunneling Cases . . . 84
6.2.2 Data Description and Preparation . . . 85
6.3 Methods . . . 85
6.3.1 Hierarchical Feature Selection . . . 87
6.3.2 Multicollinearity Analysis . . . 88
6.3.3 Model Setup . . . 89
6.3.4 Prediction Methods . . . 90
6.3.5 Model Hyperparameters . . . 91
6.3.6 Model Evaluation and Feature Importance . . . 91
6.4 Results and Discussion . . . 93
6.4.1 Generalized Feature Taxonomy . . . 93
6.4.2 Multicollinearity Problems . . . 95
6.4.3 Effects on Model Predictability . . . 99
6.4.4 Effects on Model Computability . . . 102
6.4.5 Effects on Model Stability . . . 102
6.5 Conclusions . . . 107

7 Interpreting Geologic Conditions during Tunneling 109
7.1 Introduction . . . 109
7.1.1 Background . . . 109
7.1.2 Related Works . . . 110
7.1.3 Objectives . . . 112
7.2 Data . . . 113
7.2.1 Tunneling Case . . . 113
7.2.2 Geologic Conditions . . . 114
7.2.3 Data Description and Preparation . . . 115
7.3 Methods . . . 119
7.3.1 Model Setup . . . 120
7.3.2 Prediction Methods . . . 122

7.3.3 Model Hyperparameter . . . 123
7.3.4 Feature Importance . . . 124
7.4 Results and Discussion . . . 124
7.4.1 Effects of Hyperparameters . . . 124
7.4.2 Classification-Based Interpretation . . . 125
7.4.3 Regression-Based Interpretation . . . 129
7.4.4 Feature Importance . . . 130
7.4.5 The Importance of Feature Interactions . . . 133
7.5 Conclusions . . . 133

8 Detecting Geologic Anomalies during Tunneling 136
8.1 Introduction . . . 136
8.1.1 Background . . . 136
8.1.2 Related Works . . . 137
8.1.3 Objectives . . . 137
8.2 Data . . . 138
8.2.1 Tunneling Case and Geologic Conditions . . . 138
8.2.2 Data Description and Preparation . . . 138
8.3 Methods . . . 139
8.3.1 Model Setup . . . 139
8.3.2 Principal Component Analysis . . . 140
8.3.3 Local Outlier Factor . . . 141
8.4 Results and Discussion . . . 141
8.4.1 PCA Projection . . . 141
8.4.2 Local Outlier Factors . . . 142
8.4.3 Anomaly Detection . . . 146
8.5 Conclusions . . . 147

9 Connecting TBM to the Ground Responses 149
9.1 Introduction . . . 149
9.1.1 Background . . . 149
9.1.2 Related Works . . . 149
9.1.3 Objectives . . . 152
9.2 Data . . . 153
9.2.1 Tunneling Case and Geologic Conditions . . . 153
9.2.2 Ground Monitoring Data . . . 153
9.2.3 TBM Data . . . 155
9.2.4 Data Integration . . . 155
9.3 Methods . . . 155
9.3.1 Model Setup . . . 157
9.3.2 Prediction Methods . . . 158
9.3.3 Model Hyperparameters . . . 158

9.3.4 Feature Importance . . . 159
9.3.5 Segmentation Analysis . . . 159
9.4 Results and Discussion . . . 161
9.4.1 Effects of Hyperparameters . . . 161
9.4.2 Static Random Model Predictions . . . 162
9.4.3 Dynamic Sequential Model Predictions . . . 164
9.4.4 Three-Dimensional Expansion . . . 166
9.4.5 Ground Response Segmentation . . . 166
9.4.6 Feature Importance for Overall Response . . . 168
9.4.7 Feature Importance for Segmental Response . . . 169
9.5 Conclusions . . . 170

10 Modeling TBM Excavation Process 173
10.1 Introduction . . . 173
10.1.1 Background and Related Works . . . 173
10.1.2 Objectives . . . 174
10.2 Data . . . 174
10.2.1 Tunneling Case and Geologic Conditions . . . 174
10.2.2 Data Description and Preparation . . . 175
10.3 Methods . . . 176
10.3.1 Bayesian Networks . . . 176
10.3.2 Structure Learning . . . 176
10.3.3 Model Setup . . . 176
10.4 Results and Discussion . . . 177
10.4.1 Ground Conditioning System . . . 177
10.4.2 Interactions of Excavation Features . . . 178
10.4.3 Effects of Soil Labels . . . 181
10.4.4 Feature Interactions along Tunnel Alignment . . . 183
10.5 Conclusions . . . 184

11 Modeling TBM Steering Control Decisions 186
11.1 Introduction . . . 186
11.1.1 Background and Related Works . . . 186
11.1.2 Objectives . . . 187
11.2 Data . . . 188
11.2.1 Tunneling Case and Geologic Conditions . . . 188
11.2.2 Data Description and Preparation . . . 188
11.3 Methods . . . 188
11.3.1 Steering Control Parameters . . . 188
11.3.2 Model for Steering Control Decisions . . . 191
11.3.3 Model Setup . . . 192
11.3.4 Prediction Method . . . 195

11.3.5 Model Evaluation 196
11.4 Results and Discussion 196
11.4.1 Steering Control Decisions: Human Operator versus Machine Learning 196
11.4.2 Incremental versus Non-Incremental Learning 198
11.5 Conclusions 200

12 Conclusions **203**
12.1 Summary and Conclusions 203
12.2 Future Works 206

Bibliography **209**

List of Figures

1.1 Increasing world urban populations, 1950-2050 (United Nations, 2019). 2
1.2 Cost components of (a) road tunnels and (b) railway tunnels (N. Efron & Read, 2012). 3
1.3 Factors affecting tunneling cost: (a) final cost contributors and (b) cost reduction opportunities (N. Efron & Read, 2012). 4
1.4 Control panel in TBM navigation cabin, from Dornfeld (2014). 5
1.5 Outline of this dissertation 10

2.1 TBM classification based on the ground supports. 13
2.2 Selection of TBM shield face support method in different soil particle distribution (British Tunneling Society, 2005). 14
2.3 Soil conditioning of EPB TBM tunneling in different ground types (Boundaries are only indicative) (EFNARC, 2005) 15
2.4 Overview of EPB TBM components, modified from WSDOT (2013) 16
2.5 Schematic diagram of foam production system and the typical composition, adopted from Maidl et al. (2013) 20
2.6 Excavation process control (or management) of cutting face conditions in EPB TBM tunneling (JSCE, 2016) 23
2.7 Example of EPB regulation flow diagram, adopted from Maidl et al. (2013). . . 25

3.1 Fundamental scheme of a dynamical system. 29
3.2 Tunneling as a dynamical system. 29
3.3 Control systems: (a) open loop, and (b) closed (feedback) loop. 30
3.4 EPB TBM tunneling in a feedback control system framework. 31
3.5 Model-based predictive control. 32
3.6 Human cognitive model. 38
3.7 Implementation of the cognitive model for intelligent EPB TBM systems 38
3.8 Intelligent control system for EPB TBM 39

4.1 General taxonomy of machine learning. 41
4.2 Model flexibility vs. interpretability, modified from James et al. (2013). 43
4.3 Concept of model underfitting and overfitting. 49
4.4 Example of robust machine learning workflow. 49

4.5 Data splitting with cross-validation. 50

5.1 Example of a feature of TBM operation data plotted in different data aggregation levels, following different data preparation methods from several previous studies. 54
5.2 Longitudinal plot of the SR99 data sets in both fine and coarse data aggregation levels. 59
5.3 Longitudinal plot of the NSLP1-SB data sets in both fine and coarse data aggregation levels. 60
5.4 Longitudinal plot of the NSLP1-NB data sets in both fine and coarse data aggregation levels. 61
5.5 Schematic diagram of the prediction model. 64
5.6 Schematic diagram of dynamic sequential data splitting for model training and testing. 65
5.7 Statistical parameters of the considered EPB TBM features from all the tunneling cases. 68
5.8 Aggregation ratios of the considered EPB TBM features from all the tunneling cases. 70
5.9 Comparison of prediction performance between prediction models trained using fine and coarse data aggregation levels. 71
5.10 Effects of data reduction on the prediction performance of the models. 73
5.11 Effects of data size on the computation time (left panel) and speed-up (right panel) of the RF models. 74
5.12 Visualization of feature importance comparison plot and the indication to stable and unstable ranks. 75
5.13 Examples of prediction models with relatively stable feature importance ranks. Diagonal patterns of the data points can be observed. 76
5.14 Examples of prediction models with unstable feature importance ranks. The data points were scattered without any identifiable patterns 78
5.15 Coefficient of determination (R^2) produced from feature importance comparison plots of all prediction models. 79

6.1 Correlation networks of EPB TBM data: (a) SR99, (b) NSLP1-SB, (c) NSLP1-NB 86
6.2 Schematic diagram of the prediction model. 89
6.3 Schematic diagram of static random data splitting for model training and testing. 90
6.4 Effects of ntrees and mtry hyperparameters on models with different feature sets and tunnelling cases. 92
6.5 Effects of mtry and minimum node size hyperparameters on models with different feature sets and tunnelling cases. 93
6.6 Generalized taxonomy of EPB TBM features. The generalized features are highlighted in yellow. 94
6.7 Correlation networks of EPB TBM data with the hierarchical feature selection: (a) SR99, (b) NSLP1-SB, (c) NSLP1-NB. 96

6.8 Variance inflation factors (VIF) of the EPB TBM data with and without the hierarchical feature selection. 97
6.9 Extreme errors of OLS predictions due to severe multicollinearity in the data. . 98
6.10 Comparison of prediction value distributions produced by models with different prediction methods, feature sets, and tunneling cases. 100
6.11 Comparison of prediction values produced by models with different prediction methods, feature sets, and tunneling cases. 101
6.12 Computation performance of an RF model with different numbers of predictor features 103
6.13 Comparison of top predictors identified in cutter torque models with different prediction methods, feature sets, and tunneling cases. 104
6.14 Comparison of top predictors identified in thrust force models with different prediction methods, feature sets, and tunneling cases. 106

7.1 Geologic map of the SR99 tunnel (WSDOT, 2010b, 2010c). 115
7.2 Examples of EPB TBM operation data along the tunnel alignment: features related to (a) forces, (b) speeds, (c) volumes. Background colors represent the soil (classification) labels. 119
7.3 Schematic diagram of the classification-based prediction approach. 120
7.4 Schematic diagram of the regression-based prediction approach. 121
7.5 Schematic diagram of the static random model. 122
7.6 Schematic diagram of the dynamic sequential model. 123
7.7 Effects of ntrees and mtry hyperparameter (top panels) and effects of mtry and minimum node size hyperparameters (bottom panels) on models constructed with different split rules, i.e., the Gini (left panels) and Extratrees (right panels). . . 126
7.8 Geologic interpretation along the tunnel chainages produced by the static random classification model. Black bars denote classification errors. 127
7.9 Geologic interpretation along the tunnel chainages produced by the dynamic sequential classification model. Black bars denote classification errors. 128
7.10 Geologic interpretation along the tunnel chainages produced by the dynamic sequential regression model. 129
7.11 Comparison of geologic conditions interpreted by human (the geologic map) and by the AI system. 131
7.12 Feature importance ranks produced by (a) impurity, (b) permutation, and (c) conditional permutation methods. 132
7.13 Soil classification accuracy when RF models were trained using a single predictor feature only. 134

8.1 Examples of 2D PCA projections of the EPB TBM data at different chainages along the tunnel alignment. 143
8.2 Example of 3D PCA projection of the EPB TBM data (at Chainage 28730). . . 144

8.3 Interpreting PCA results: (a) Scree plot to visualize the information in each PC score, (b) biplot to visualize the contribution of each variable feature to PC1 and PC2. 145
8.4 Examples of degree of anomaly of projected data points: (a) current EPB TBM location inside clusters, (b) current EPB TBM location outside clusters (i.e., potentially at new ground conditions). 146
8.5 Longitudinal plot of the geologic anomaly measurement: LOF of every new EPB TBM data point projection along the tunnel alignment. 147

9.1 Locations of MPBX instrumentation along the tunnel alignment 154
9.2 Tunneling-induced ground movement data relative to the TBM head position. . 157
9.3 Schematic illustration of tunnel spatial geometries. 158
9.4 Schematic diagram of the (a) static random and (b) dynamic sequential models. 159
9.5 Effects of RF hyperparameters on the OOB samples of the training data set. . . 161
9.6 Tunneling-induced ground movement predictions with different types of responses at selected MPBX testing data set: settlement (panels in the left column), heaving (the middle column), and relatively stable responses (the right column). 163
9.7 Absolute error of every monitoring point relative to the TBM head distance from all predictions in the testing data set. 164
9.8 Simulation of real-time tunneling-induced ground movement predictions using the dynamic sequential model. 165
9.9 Mean absolute error of predictions along the tunnel alignment. 166
9.10 Three-dimensional expansion of the tunnelling-induced ground movement predictions. 167
9.11 Ground response segments based on the segmentation analysis. Each segment produces MAE < 0.2 mm and is shown as a distance between the two red dashed boundary lines. 168
9.12 Feature importance rank of the overall ground response model 169
9.13 Feature importance rank of each ground response segment. 171

10.1 Interactions of features related to ground conditioning volumes constructed by the structure learning algorithm. Red links denote true relationship. 178
10.2 Ground conditioning volumes along the tunnel alignment: (a) normalized data, (b) probabilistic strength of the links, and (c) probabilistic direction of the links. 179
10.3 Examples of iteration during structure learning to find a BN graph configuration with the highest BIC score. 180
10.4 Interactions of features related excavation constructed by the structure learning algorithm: (a) chainage 23000 ft, (b) chainage 28000 ft. Red links denote true relationship as calculated; blue links denote sensible interactions based on available literatures; black links denote possible interactions. 181

10.5 Interactions of features related excavation with soil labels constructed by the structure learning algorithm: (a) chainage 23000 ft, (b) chainage 28000 ft. (See Figure 10.4 caption for links color legend.) 182
10.6 Features related to penetration rate along the tunnel alignment: (a) normalized data, (b) probabilistic strength of the links, and (c) probabilistic direction of the links. 183
10.7 Features related to chamber pressure along the tunnel alignment: (a) normalized data, (b) probabilistic strength of the links, and (c) probabilistic direction of the links. 185

11.1 Schematic diagram of TBM steering systems. 190
11.2 Configuration of the thrust and articulation jack groups. 191
11.3 Schematic diagram of the feedback system for TBM steering control. 192
11.4 TBM movements as a dynamical system. 193
11.5 Inverse of TBM movement process that outputs the steering control parameters. 193
11.6 Schematic diagram of the model input-output mapping (training phase). 194
11.7 Schematic diagram of the prediction model input and output (testing phase). . . 195
11.8 Simulation of steering control decisions at five selected ring segments: human operator vs. machine learning. 197
11.9 Simulation of steering control decisions at six consecutive ring segments without incremental learning. 199
11.10 Simulation of steering control decisions at six consecutive ring segments with incremental learning. 201

List of Tables

2.1 Typical KPIs in a tunneling operation instruction sheet. 22

4.1 Summary of machine learning tasks and algorithms in this dissertation 44

5.1 List of EPB TBM features selected for the data aggregation analysis. 62
5.2 Summary of the observation data points. 62

6.1 Dimension of the data sets before and after data preparation. 87
6.2 Overall prediction performance of different models and feature sets. 99

7.1 Borehole labels for the classification model. 116
7.2 Soil labels for the regression model. 117
7.3 List of EPB TBM features used in the supervised geologic interpretation model. 118

8.1 List of EPB TBM features used in the unsupervised geologic anomaly detection method. 138
8.2 Algorithm of the unsupervised geologic anomaly detection method. 140

9.1 Comparison of machine learning algorithms for estimating tunneling-induced ground movements from previous studies. 151
9.2 List of EPB TBM features used in the tunnelling-induced ground movement estimation model. 156
9.3 Algorithm of the segmentation analysis. 160

10.1 List of EPB TBM features used in the tunneling excavation process modeling . 175

11.1 List of EPB TBM features and positions used in the steering control decision model. 189

Acknowledgments

I would like to express my gratitude to all those who have contributed to the completion of my doctoral study and dissertation. First, I would like to express my sincere gratitude to Professor Kenichi Soga for his mentorship during this period. This mentorship has shaped my research philosophy and inspired me to have a vision far ahead of time and to be not afraid to explore novel scientific areas, break the boundary of knowledge, and create innovations. He directed me to this research topic and patiently provided guidance, advice, and resources during the research work.

Also, I would like to thank Professor Dimitrios Zekkos and Professor Giles Hooker, who served on my doctoral dissertation committee, for reviewing this manuscript and providing valuable discussions during the preparation. Their views are essential due to the interdisciplinary domain of this research problem. In addition, I would like to extend my thank to Professor Jonathan Bray, Professor Robert Kayen, and Professor Bin Yu, who served as the chair and members of my doctoral qualifications committee and provided valuable insights during the early stages of my research proposal.

This research was formed by discussions with many colleagues from the industry who work in various positions (tunnel experts, engineers, and operators) and in various institutions (tunnel owners, contractors, consultants, and specialists). The shared knowledge was instrumental in this research and greatly appreciated. It would not be possible to list all of them here. However, special thanks are given to collaborators who were willing to spend much of their time guiding me during some periods of the work: Yuji Fujita, Satoru Yamauchi, and their colleagues at Enzan Koubou Co. Ltd., Dr. Zhangwei Ning and his colleagues at Sixense Inc., and Hiroshi Kogi and his colleagues at Shimizu Corp. and Shimizu Institute. Their involvement was critical to bridging this research to be relevant to industrial perspectives and applications.

This research was conducted using data obtained from the Washington State Department of Transportation (WSDOT) and PT. Mass Rapid Transit Jakarta (MRTJ). The availability of the data was critical in this research and greatly appreciated.

This doctoral study was made possible by a scholarship from the Indonesia Endowment Fund for Education, as well as awards and appointments from the Department of Civil and Environmental Engineering, University of California Berkeley. Financial resources from these institutions are acknowledged.

Graduate student life is not only about academic and research activities, but also about building relationships and memories with fellow scholars. During this period, I was fortunate to meet many inspiring people and make friends with them. It would not be possible to list all of them here. However, special thanks are given to fellow graduate students who served on the Geo-Bear Leadership with me during 2019-2020: Ezra Setiasabda Tjung, Grigorios Lavrentiadis, and Andrew Yeskoo; and during 2020-2021: Zorana Mijic, Xinyi Qian, Camilo Pinilla-Ramos, and Bodhinanda Chandra. I thank Peter Hubbard and Franklin Olaya, who started the program at the same time and shared efforts in completing the first-year courses; and Renjie Wu, who started the AIML sub-group in the Soga Research Group with me and

provided early comments on my research during the group meetings. I would also thank a longtime friend Daniel Hutabarat, who persuaded me to come to Berkeley. I believe this was a pivotal decision in my life.

Finally, a doctoral study is a long journey. The journey can be intellectually stimulating and fulfilling, with many fascinating discoveries. But it can also be challenging and exhausting, with many uncertainties. Therefore, I would like to thank and dedicate this endeavor to my parents and all of my family for their unwavering support throughout this journey. Especially to Tita and Kal, whose presence at Berkeley has given me moments of happiness that I will forever cherish.

Chapter 1

Introduction

1.1 Background

The global population is projected to increase and will be about 10 billion in 2050 (United Nations, 2022). Due to urbanization, most of the increase will take place in urban areas. As shown in Figure 1.1, more than 50% people have been living in urban areas since 2008. This number is projected to reach 70% by 2050 (United Nations, 2019). The increase in the global population and the trend of urbanization can bring challenges to society, such as scarcity of space, infrastructure inadequacy, transportation congestion, and other environmental issues. Underground space, including tunnels, has been seen as one of the key solutions to tackle these challenges (Broere, 2016; Kaliampakos, 2016). It is envisioned that more tunnels are required in urban areas to meet the increasing demands of urban space, alleviate transportation congestion, and save the on-ground space for green environments.

Tunneling technology has been advancing in the past several decades to meet the demand. The development of face pressurized tunnel boring machines (TBM), such as earth pressure balance (EPB) and slurry pressure balance (SPB) TBM, has spurred the growth of urban tunneling, which is more challenging since it is commonly done in soft ground conditions under densely populated regions. Nevertheless, this growth is often impeded by tunneling costs. The cost of underground structures has always been significantly higher than on-ground and elevated structures (Kaliampakos et al., 2016). According to Godard and Hugonnard (1989), underground structures could cost two to three times higher than on-ground structures. In unfavorable sites, this cost could soar to six to ten times. Similarly, Fox (2000) reported that underground urban rapid transit systems could cost four to six times higher than on-ground systems.

Not only that, estimating the cost of tunnels is challenging. N. Efron and Read (2012) conducted a tunnel database analysis and a survey on the cost of tunneling to various stakeholders (i.e., clients, consultants, contractors, and cost estimators). They found that the actual tunneling cost is almost always inflated from the estimated cost. This happened globally in most of the continents. The results also show that the cost of excavation is the most

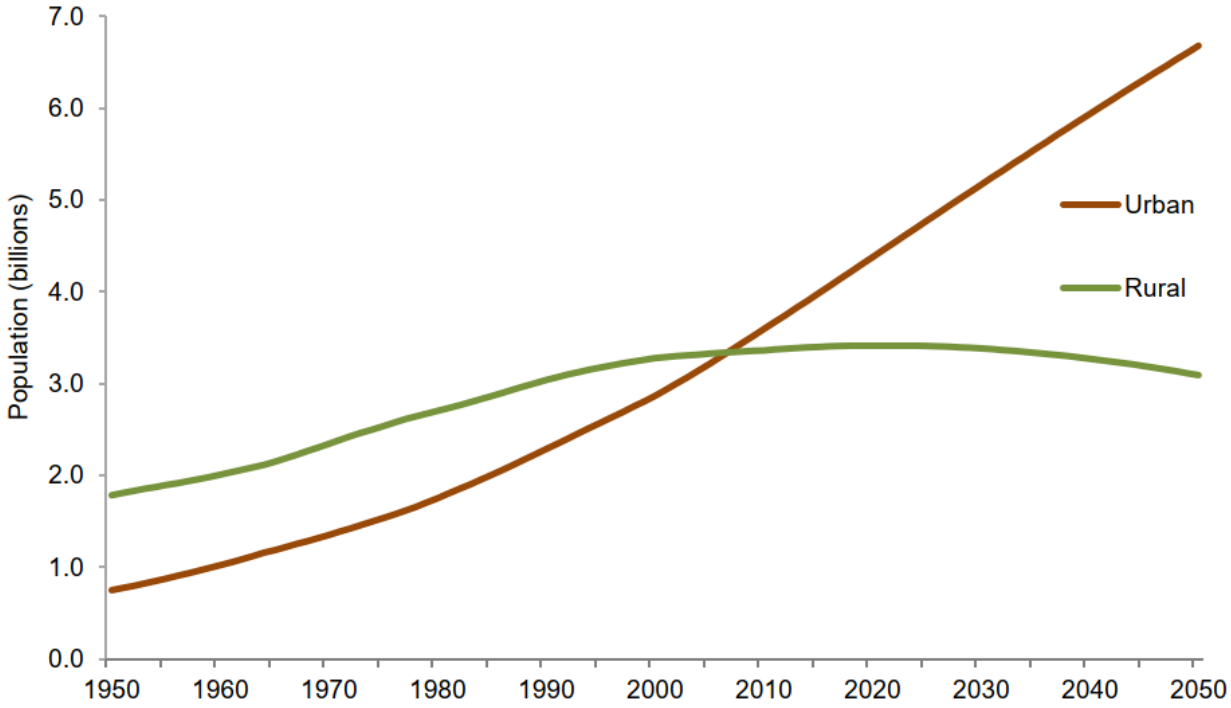

Figure 1.1: Increasing world urban populations, 1950-2050 (United Nations, 2019).

significant component in tunneling, which can be more than 50% of the total cost (Figure 1.2). Most stakeholders identified that the geologic condition is the first major cost contributor. Most stakeholders also believe that reducing the geologic uncertainty through site investigation can be the most potential cost reduction opportunity (Figure 1.3). Confirming the survey, Paraskevopoulou and Boutsis (2020) provided some cases where inadequate geologic information could overrun tunneling costs by up to 77%. Membah and Asa (2015) used statistical methods to identify 40 tunneling cost factors based on carefully selected 39 articles published from 1988 to 2013. They concluded that the complexity of tunneling processes and the uncertainty of geologic conditions are the top contributing factors that inflate tunnel construction costs. These studies have corroborated the main problem of tunneling, i.e., unpredictable tunneling operations due to variability and uncertainty of the geologic conditions.

Tunneling operations are always complicated and uncertain. During tunneling, TBM operators must be aware of changing ground conditions. At the same time, they have to

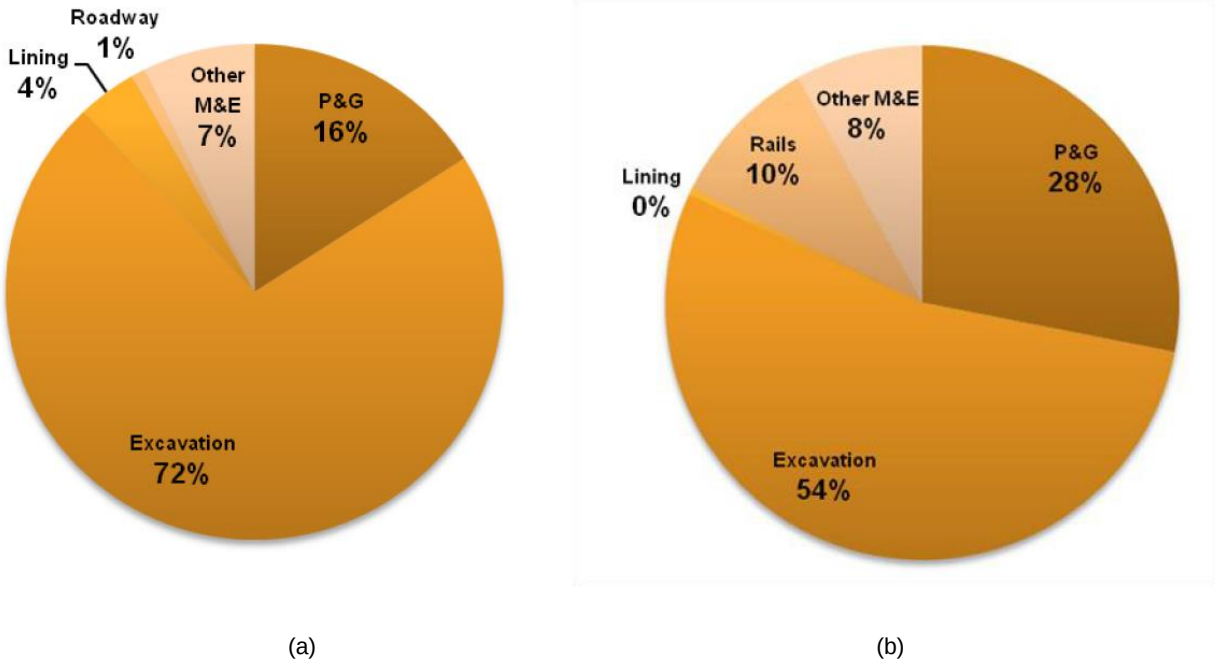

Figure 1.2: Cost components of (a) road tunnels and (b) railway tunnels (N. Efron & Read, 2012).

make real-time decisions on TBM control parameters. Furthermore, unlike driving a car, there is no window in the cabin (Figure 1.4). All the decisions can only be made based on interpreting the tunneling operation data. This data is generated every few seconds by numerous sensors installed on the TBM. For humans, continuously making real-time interpretations, judgments, and decisions in various control tasks based on numerous streaming data are laborious and may lead to inconsistency, bias, and errors.

Furthermore, the advancement of sensor and data acquisition technologies has boosted the amount of tunneling operation data. This increases the amount of information acquired during tunneling and thus may reduce tunneling uncertainty. However, due to the massive volume and data acquisition velocity, the data cannot be handled by the operators and engineers intuitively in real-time during tunneling. Moreover, the data is difficult to be interpreted due to its inherent complexity. They are products of causal effect interactions of the ground conditions, TBM behaviors, and human control decisions. Therefore, at the

Figure 1.3: Factors affecting tunneling cost: (a) final cost contributors and (b) cost reduction opportunities (N. Efron & Read, 2012).

moment, the data cannot be fully utilized. A more systematic approach is required to process the data.

The emergence of artificial intelligence (AI) techniques and the increase in computing performance in the past few years have opened extensive opportunities for managing and interpreting "big data" (Jordan & Mitchell, 2015). AI systems can be seen as computer-based systems that exhibit intelligent behavior, such as perceiving and inferring information from data. Machine learning is a subset of AI that defines the ability of an algorithmic system to "learn" tasks without being explicitly programmed. Instead of following a set of fixed rules, machine learning algorithms can independently detect hidden structures in the data, map the patterns, and identify complex systems (Brunton et al., 2016). This advancement also promotes the expansion of data-driven modeling and the development of autonomous systems. This technology can potentially optimize the utilization of massive tunneling data.

Some studies have implemented machine learning algorithms to transform TBM operation

Figure 1.4: Control panel in TBM navigation cabin, from Dornfeld (2014).

data into tunneling information, such as geologic conditions. Q. Zhang et al. (2019) used three machine learning algorithms, i.e., Support Vector Machine (SVM), Random Forests (RF), and k-Nearest Neighbors (kNN), to rock mass classification systems. For the same purpose, Zhao et al. (2019) evaluated the performance of eight machine learning algorithms, i.e., kNN, Bayesian Linear Regression (BLR), SVM, Decision Tree (DT), RF, CatBoost, Extreme Gradient Boosting (XGBoost), and Artificial Neural Network (ANN). Recently, Erharter et al. (2020) continued the development by proposing a real-time rock classification system during tunneling using ensemble Long Short-Term Memory (LSTM).

Machine learning has also been used to estimate tunneling-induced ground responses. The early development was dominated by the implementation of ANN to estimate the maximum ground settlement, such as by Shi et al. (1998), Suwansawat and Einstein (2006), and Boubou et al. (2010). Later, more variations of machine learning algorithms were used for this task, such as the Adaptive Neuro-Fuzzy-based Inference System algorithm (ANFIS) (Bouayad & Emeriault, 2017) and Multivariate Adaptive Regression Splines (MARS) (Goh et al., 2018).

Recently, P. Zhang, Wu, Chen, Dai, et al. (2020) used RF and LSTM to investigate the interactions between a shield TBM and the ground in both cross and longitudinal sections.

Besides modeling the tunneling environments (ground conditions and responses), machine learning has also been used to develop intelligent TBM control systems. Most studies proposed predictive models to estimate chamber pressures for the EPB regulation system, such as using the Least Squares-SVM (LS-SVM) (X. Liu et al., 2011; X. Liu & Zhang, 2019), a neural network-based algorithm with a Gated Recurrent Unit (GRU) (Gao et al., 2020), and Backpropagation Neural Networks (BPPN) in Dual Heuristic Programming (DHP) (X. Liu et al., 2020). Other studies proposed predictive models to estimate the optimal point of the resultant force produced by thrust jacks for the directional control guidance system, such as using the gradient boosting machine (Wada et al., 2021).

These studies show that the implementation of AI and machine learning is becoming more prevalent in tunneling. The use of machine learning allows flexible data-driven modeling without any predefined assumption. This can be beneficial for tunneling since the process is governed by complex causal effect interactions of ground conditions, machine behaviors, and human decisions. Machine learning can also exploit latent information in massive data, enabling better utilization of tunneling operation data.

1.2 Problem Statements

Despite the emergence of AI and machine learning implementation in tunneling and its encouraging results, some potential issues have been identified.

Limited studies discuss the characteristics of TBM data and evaluate the effects of data preparation methods on data-driven models.

- **Understanding the effects of data aggregation**. Many studies have been done to develop prediction models based on TBM operation data. However, less effort has been given to understanding the characteristics of TBM data and the effects of data preparation on the models. There is no agreement on how the observation data points should be prepared (or aggregated). The previous prediction models were developed using different data aggregation levels. The effects of TBM data aggregation on the prediction models have not been identified.

- **Understanding the effects of feature selection**. Similarly, there is no standard approach to selecting TBM features as inputs in a data-driven model. The previous models were developed using different inputs selected by feature selection algorithms or pre-selected by the model developers. The effects of TBM feature selections on prediction models have not been identified.

Developing data-driven models for inferring the environments during tunneling remains an open question.

- **Perceiving ground conditions**. Some studies have employed machine learning to develop geologic classification models based on TBM operation data. However, most studies focused on rock mass classification in discrete categories. This may not be realistic since the actual ground materials, especially soils, do not have a clear boundary between the categories. Furthermore, the discrete classification system cannot identify the gradual transitions of geologic deposits. In addition, most previous studies used only a limited number of TBM features as predictors. This might lead to underfitted models, where the predictors are not adequate to capture the complexity of the tunneling process.
- **Perceiving ground responses**. Some studies have used machine learning to connect TBM operations to tunneling-induced ground movements. However, most models were merely developed to estimate the maximum ground movements. Less attention has been given to investigating the effects of TBM operations on longitudinal ground movements. The progressions of longitudinal ground movements should result from the interaction between different TBM control parameters and the ground at different TBM passing phases (i.e., before, during, and after TBM passing).

The tunneling sequence must be considered when developing a deployable data-driven model.

- **Modeling considerations**. Some studies did not provide an exhaustive discussion about their modeling and deployment considerations. For example, a static model with a randomized data splitting scheme will not be deployable due to the sequential process in tunneling. The dynamic sequential process in tunneling should be an important consideration in developing a data-driven model for tunneling.
- **Model computability**. The dynamic sequential process means the machine learning model needs to be continuously retrained at every TBM advance. A computation bottleneck may arise if the time required for training is longer than the TBM advancing time. However, less attention has been given to this computability issue. Many previous studies favored employing more computationally expensive machine learning algorithms.

The goal of developing data-driven decision-making systems in tunneling should not be centered merely on prediction accuracy.

- **Model predictability**. Most studies focused on increasing the model's prediction performance by selecting and arranging machine learning algorithms. All the studies argued that their proposed model and the selected machine learning algorithms produced excellent prediction performance. The comparative studies also resulted in different conclusions. This indicates that, with adequate preparation and training, any machine learning algorithms may produce reasonable estimations. This also implies that the selection of machine learning algorithms may not be the critical issue in developing a data-driven model in tunneling.

- **Model interpretability**. Machine learning models have been notoriously seen as "black box" models. However, more attention has been given to developing less interpretable models, such as using neural network-based algorithms. Less effort has been devoted to opening the "black box". Furthermore, less effort has also been given to leveraging machine learning to elucidate complex interactions in tunneling processes and generate new scientific discoveries.

- **Model integration**. Most studies focused on modeling a part of the tunneling process, e.g., steering, advancing, and earth pressure balancing processes. Less attention has been given to building an integrated model for tunneling processes.

1.3 Objectives

This study aims to develop an integrated framework of AI systems for EPB TBM tunneling. The framework is constructed based on a feedback loop control system, which is the foundation of autonomous systems. The AI systems are developed using data-driven models and structured to follow the human cognitive model, i.e., sensing, perceiving, and decision-making. Machine learning algorithms are used to build the data-driven models and to leverage information from massive sensing data produced during tunneling operations.

The specific objectives of this study are listed below.

1. To better understand the characteristic of tunneling operation data and the effects of data preparation (i.e., data aggregation and feature selections) on data-driven models. This is expected to be a foundation to facilitate a unified method of TBM data preparation.

2. To develop AI systems that can perceive tunneling environments in real-time based on the streaming sensing data during tunneling. This includes systems to (i) infer the encountered geologic conditions and (ii) estimate tunneling-induced ground responses. This is expected to provide systematic and objective tools to perceive tunneling environments based on data.

3. To develop AI systems that can model the complexity and decision-making in the tunneling process based on data in integrated and interpretable representation. This representation is essential to exploit causal effect interactions of the ground conditions, machine behaviors, and human decisions during tunneling. This is expected to expand the current tunneling knowledge and can be extended to develop autonomous decision-making systems for TBM control parameters.

The result of this study is expected to answer two main challenges in tunneling, i.e., (i) reducing the uncertainty by exploiting latent information in massive tunneling data that are not fully utilized in current practices and (ii) facilitating more quantitative and systematic TBM control systems to minimize laborious, subjective, and biased decision-making.

This development is envisioned to be a building block that improves tunneling performance, reduces risk and cost, and advances the fundamental development of autonomous TBM technology.

1.4 Dissertation Outline

This dissertation consists of 12 chapters (including this introduction) and can be divided into four main parts, i.e., fundamentals and reviews, sensing data, perceiving systems, and decision-making systems. Figure 1.5 maps the position of each part and chapter in the tunneling AI framework. More descriptions of each part and chapter are discussed below.

- **Part I – Fundamentals and Reviews**. This part provides the fundamentals and reviews of different domain knowledge required for this study. Chapter 2 describes tunneling and EPB TBM system components. Chapter 3 describes dynamical systems and control and discusses the proposed AI framework. Chapter 4 describes machine learning and discusses the current implementation of machine learning in TBM tunneling operations.

- **Part II – Sensing Data**. This part investigates the characteristics of tunneling operation data and the effects of different data preparation methods on data-driven models. Chapter 5 investigates the effects of data aggregation on its statistical characteristics and the behaviors of the prediction models. Chapter 6 presents a proposed taxonomy of EPB TBM features, demonstrates the multicollinearity among the features, and investigates the effects of different feature selections on the behaviors of the prediction models.

- **Part III – Perceiving Systems**. This part proposes AI systems to perceive tunneling environments in real-time based on the streaming sensing data during tunneling. Chapter 7 and Chapter 8 present the proposed supervised geologic interpretation and unsupervised geologic anomaly detection systems for EPB TBM tunneling, respectively. Chapter 9 presents the proposed data-driven model for estimating tunneling-induced ground movements in real-time during tunneling.

- **Part IV – Decision-Making Systems**. This part proposes AI systems to model the complexity and decision-making in the tunneling process. Chapter 10 presents the implementation of probabilistic graph modeling and structure learning to explore interactions of EPB TBM features during tunneling and model the causal effect decision-making process in integrated and interpretable representation. Chapter 11 presents the proposed data-driven model for autonomous TBM steering control parameters.

- **Part V – Conclusion**. Chapter 12 provides the summary and conclusion of this study. This chapter also provides some discussions on future research directions.

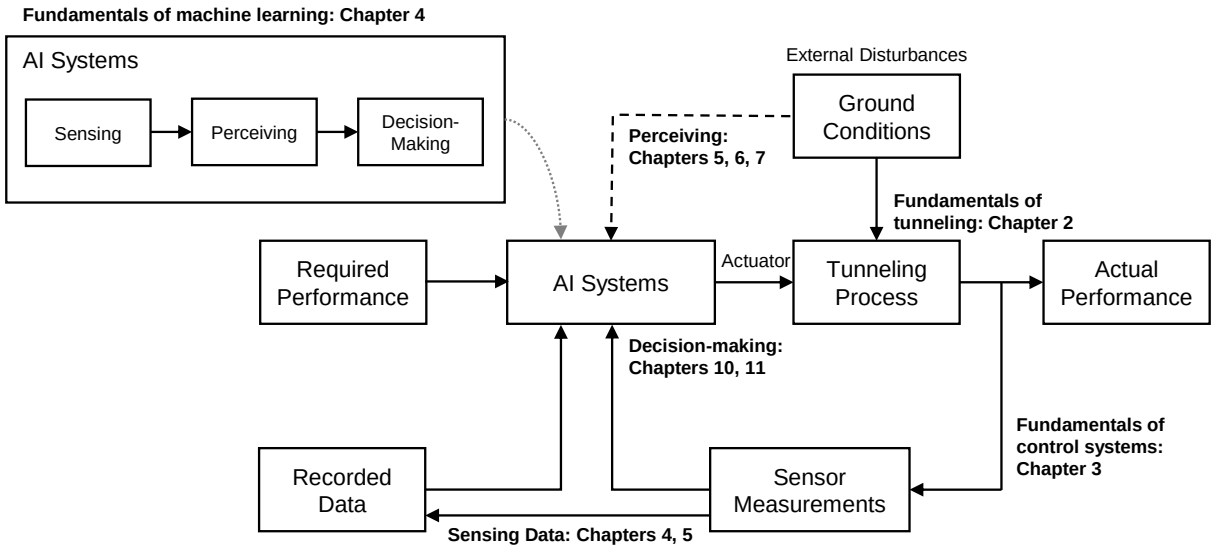

Figure 1.5: Outline of this dissertation

.

Chapter 2

EPB TBM Tunneling

2.1 Introduction

This chapter provides the fundamentals of mechanized tunneling methods and operation, particularly the earth pressure balance (EPB) tunnel boring machine (TBM) tunneling. This chapter briefly discusses the historical development of mechanized tunneling and the overview of TBM. Then, this chapter provides more detailed descriptions of EPB TBM, i.e., the components and operation process. This understanding is essential to build intelligent tunneling systems.

2.2 Historical Development

Developments of EPB TBMs can be traced back to about a century ago. The process started with the invention of shield tunneling, which then progressed to mechanized excavation and pressurized face tunneling. More discussions on the development of TBM tunneling can be found in Guglielmetti (2008), Koyama (2003), Kurosawa (1981), Maidl et al. (2013), and Stack (1982). These developments are briefly summarized below.

- **Shield tunneling**. The shield tunneling method was invented in the early 19th century by M. I. Brunel. The shield was rectangular and was divided into compartments, with a worker working independently in each of the compartments. After an excavation section was completed, the shield was driven forward using hydraulic cylinders. This concept was developed into a screw shield excavation process with immediate lining and was implemented in the Thames tunnel project in London (1825 – 1843). Despite the completion, this construction project suffered several water inflows. The circular shield and the use of circular cast iron segments as the tunneling lining were introduced in 1869 by J. H. Greathead. This method was successfully implemented without substantial problems with the ground and groundwater conditions.

- **Mechanized excavation**. The shield method enabled the integration of mechanized excavation in tunneling. The first patent of a shield machine was granted in 1876 to J. D. Brunton and G. Brunton. The shield had a hemispherical rotating cutting head built up of single plates. A better design was proposed and patented in 1896 by J. Price. This was the first that use a rotating cutting head inside a shield.

- **Compressed air pressure balance**. The method of compressed air was proposed in 1830 by T. Cochrane. This method excavated the tunnel in a working space under increased pressure. This method could solve the problem of controlling water on soft ground in underwater tunnel construction and was implemented in tunnel construction in Antwerp (1879) and New York (1880). The compressed air method was combined with the shield method to construct the London Underground (1886). However, the compressed air method was associated with several problems, e.g., the health of the workers (as they had to work back and forth under different pressure conditions), the non-uniform earth pressure in large-diameter tunnel construction, and the issue in soils with higher permeability.

- **Fluid pressure balance**. The fluid-supported shield method was proposed by J. H. Greathead to deal with the compressed air problem. The method used fluid to excavate and transport the soils hydraulically as slurry. Following this development, the design of a slurry-supported face was proposed by E. Gardner and successfully implemented in a sewage tunnel in 1959. The slurry shield machine with a cutting wheel and hydraulic muck (remolded excavated soils) removal was first used in Japan in 1967. The use of bentonite suspension for active face support was introduced in 1960 by Schneidereit. The first shield machine with a bentonite-supported face was developed and used in Germany in 1974.

- **Earth pressure balance**. The development of earth pressure balance shield machines started after the slurry shield machine. The development was motivated by stringent environmental regulations in many large cities in Japan, such as groundwater and air pollution, and to avoid unfavorable working environments under compressed air conditions. The first design was developed in 1963 by a Japanese company and intended to excavate soft and flowing soil beneath the groundwater table. The development of earth pressure balance shields started in the early 1970s in Japan, and the first tunneling was carried out in 1974 in Tokyo.

This brief discussion shows continuously changing technological advancement in tunneling practices. This development does not stop at this point, as it will continue following the changing needs and challenges faced by society.

2.3 TBM Overview

The challenges in tunneling and the development of tunneling technologies have triggered various tunneling methods and types of machines. There are several different ways to categorize tunneling machines, such as by the International Tunneling Association (ITA, 2000), the German Association for Underground Construction (TVM) (Maidl et al., 2013), and the Japanese Society of Civil Engineers (JSCE, 2016). These categorization systems depend on factors such as tunneling methods, ground supports, and components of the machines. However, many of the tunneling machines can essentially be categorized as TBMs. In principle, a TBM can be defined as a machine that excavates the ground using a combination of cutter rotation and thrust advance. Depending on the geologic conditions of the tunnel alignment, the excavation process may or may not require ground support. Figure 2.1 presents general categories of TBMs based on the provided ground supports.

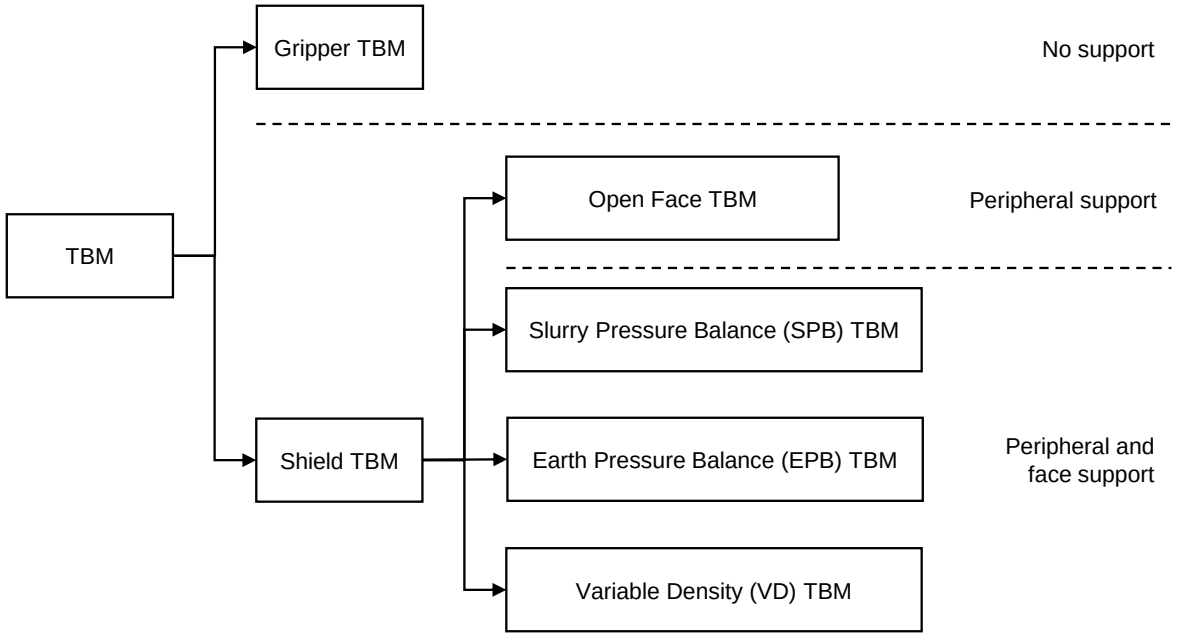

Figure 2.1: TBM classification based on the ground supports.

In hard rock tunneling, the excavation can be carried out without face and peripheral supports. In this case, the TBM thrust force is applied by radial grippers acting on the sides of the tunnel. Peripheral ground supports (i.e., TBM shields) can be required to minimize risk in unpredictable geologic conditions such as weathered rocks and soils. Open-face shield TBM can be used in hard cohesive soils, where the face can stand up without face support.

In soft ground tunneling, shield TBM with face support is required. The support can be provided mainly by slurry or earth pressure balancing (Figure 2.2).

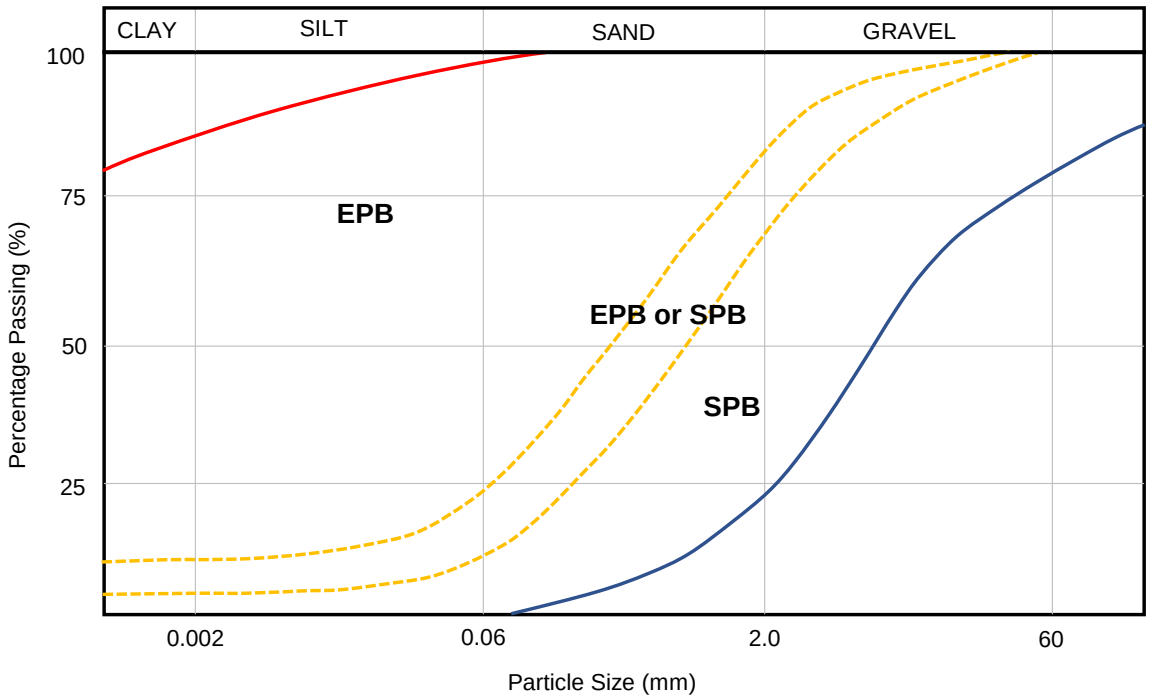

Figure 2.2: Selection of TBM shield face support method in different soil particle distribution (British Tunneling Society, 2005).

The slurry pressure balance (SPB) method provides support on the excavation face using a pressurized fluid, e.g., bentonite suspensions. The density and viscosity of the fluid should be able to be varied according to the ground conditions. The earth pressure balance (EPB) method provides support on the excavation face using remolded excavated soils (muck). Additives can be applied on the excavation face and the chamber to condition the muck into favorable properties. The selection of the face pressure support is governed by soil types along the tunneling alignment. SPB is more appropriate for high-permeability soils such as sand and gravel. EPB is optimal for cohesive soils that are dominated by clay and silts. Recently, variable density (VD) TBM has been developed to deal with a broader range of soil types.

2.4 EPB TBM Components

In principle, EPB TBMs support the excavated face by pressuring the slurry formed by the remolded excavated material in the cutter chamber. Typically, this system is suitable for tunneling in soils with fine contents higher than 30% and optimal for soft clay-to-silt and silt-to-sand soils (Maidl et al., 2013). Ground conditioning may be required to improve the flowing properties of the excavated soil (Figure 2.3). The main components of an EPB TBM consist of (i) cutter head, (ii) thrust, (iii) articulation, (iv) excavation chamber, (v) ground conditioning, (vi) muck discharge, (vii) backfill grouting, and (viii) segment erection systems (Figure 2.4). These components, except the segment erection system, which is out of the scope of this study, are briefly discussed in the following.

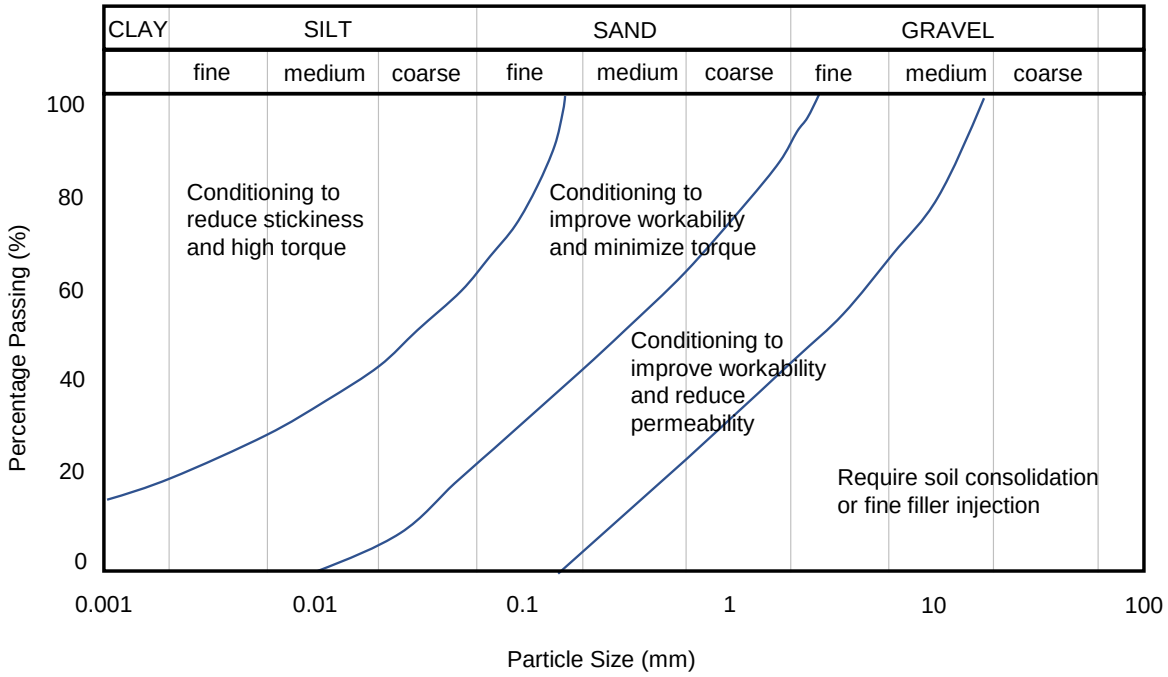

Figure 2.3: Soil conditioning of EPB TBM tunneling in different ground types (Boundaries are only indicative) (EFNARC, 2005)

2.4.1 Cutter Head

The cutter head system is an essential excavation component in any TBMs. It can be designed in various types (e.g., rotary, non-circular), shapes (e.g., flat, semi-dome, and dome),

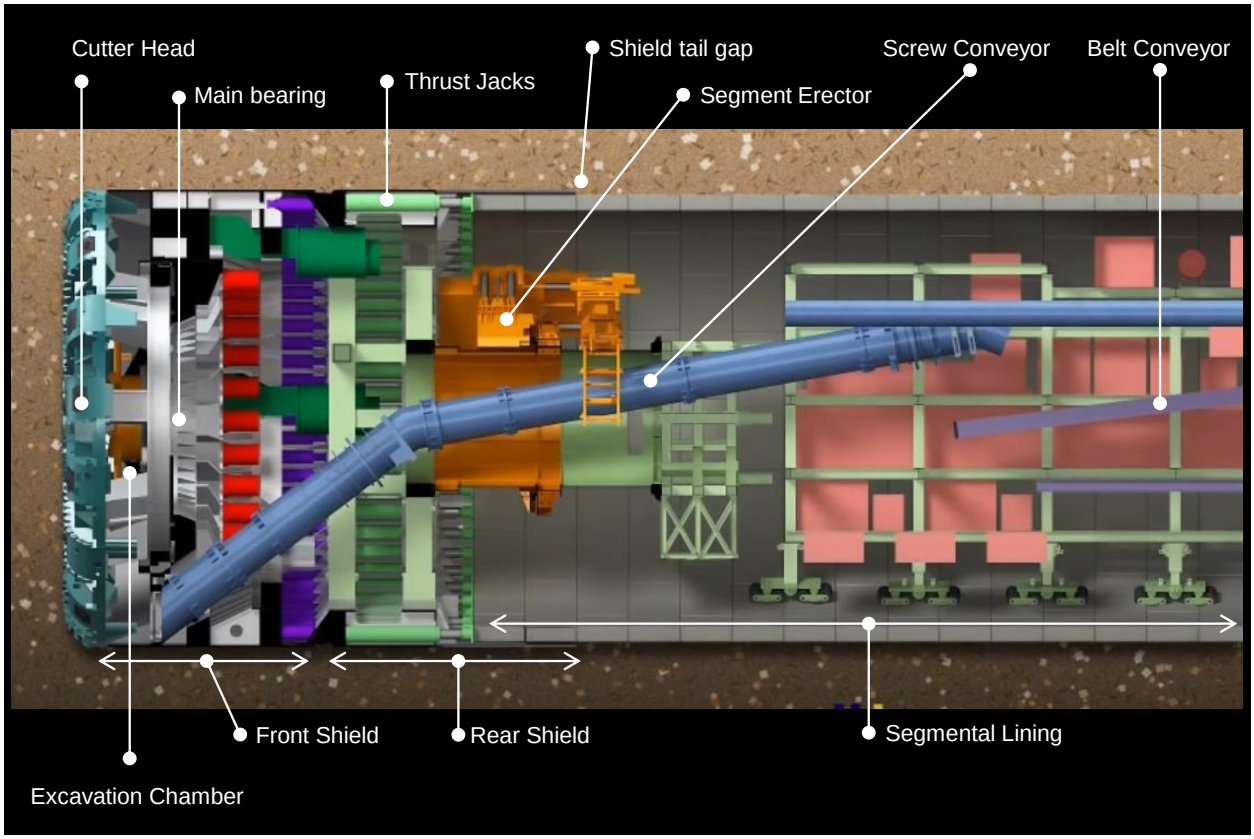

Figure 2.4: Overview of EPB TBM components, modified from WSDOT (2013)

materials, and bits layouts. The design should consider the ground conditions and the expected stability at the cutting face. It should adequately deliver two main cutter head capabilities, i.e., the cutter torque and rotation speed.

- **Cutter torque**. The required torque is mainly determined as a function of the TBM outer diameter and the ground conditions. The torque should be adequate to overcome all resistance forces during the rotation, such as the soil resistance, main bearing, cutter drive unit seal, and other mechanical losses. An appropriate allowance should also be provided due to the inherent uncertainty in tunneling processes. Typically, the allowance is estimated empirically based on the TBM type, shield diameter, and soil conditions (JSCE, 2016). Also, as a general rule, the maximum cutter rotation speed should be achieved by only delivering less than half of the maximum torque capacity (Maidl et al., 2013).

- **Cutter rotation speed**. The required cutter rotation speed (in rpm) is determined

from the cutter circumferential speed (in m/min) and the diameter of the cutter head. Typical cutter circumferential speed can be 15 to 25 mm/min (JSCE, 2016). However, it depends on the target advance rate and the ground conditions. Note that increasing the cutter rotation speed may expedite the cutter tools wearing. Therefore, the operational cutter rotation speed for a target advance rate is commonly designed to be as slow as possible (Maidl et al., 2013).

- **Overcutting**. Overcutting devices include the over-cutter and the copy-cutter tools. The over-cutter increases the excavation diameter outward from the cutter head. This is useful for reducing the thrust forces. The copy-cutter increases a selected side of the excavation diameter outward, which allows easier shield maneuvering. The over-cutter blades can be activated by hydraulic jacks. The degree, range, and stroke length are adjusted by the operator.

2.4.2 Thrust

The thrust system is an essential component to advance and maneuver the TBM. The thrust system consists of thrust force, thrust jack arrangement, thrust stroke length, and thrust jack speed.

- **Thrust force**. The thrust force capacity is mainly determined as the sum of all the resisting forces. The resistance forces against driving a shield TBM consist of friction between the shield and the ground, the pressure acting on the cutting face, direction changes such as when curving, friction in the tail, and trailing gears (JSCE, 2016). Similar to the cutter torque capacity, an appropriate allowance should also be provided due to the inherent uncertainty in the ground conditions and tunneling processes, which is typically an empirical factor (Maidl et al., 2013). The design of thrust force per jack and the required number of thrust jacks depends on various considerations, but mainly the outer diameter of the TBM, the required total thrust force, the segment structure, and the tunnel alignment. Typical thrust force per jack can be 500 to 1500 kN for small and medium-diameter shields and 2000 to 5000 kN for large-diameter shields (JSCE, 2016). Constant control of the thrust force is vital to avoid overloading.

- **Thrust jack arrangement and speed**. The thrust jacks are positioned parallel to the longitudinal axis of the shield and typically arranged in equal spaces to ensure uniform force over the circumference of the segment ring. The stroke length should fit the segment width with an additional allowance for margin. The operating speed should be based on the target advance rate. The typical operating speed in a straight trajectory (all jacks are in operation) is about 40 to 60 mm/min. However, the actual advance rate can be 20 to 45 mm/min at a straight trajectory, 15 to 35 mm/min at a curving trajectory, and 60 to 100 mm/min for high-speed tunneling (JSCE, 2016).

2.4.3 Articulation

The articulation system is required to control the TBM position and maneuver in curved tunnel alignment. In general, there are two types of articulation systems, as described in the following.

(i) Passive articulation system. In a passive articulation system, the shield is mainly steered by regulating the thrust jack to achieve a required jack stroke difference (JSD). The thrust jack can be regulated by (i) selecting the patterns of active thrust jacks (i.e., jack activation-controlled TBMs) or (ii) adjusting different pressures on the active thrust jacks (i.e., pressure-controlled TBMs).

(ii) Active articulation system. In an active articulation system, besides the JSD, the maneuvers can also be controlled by specifying the angle of the articulation unit. This unit is essentially a number of articulation jacks that connect the front and rear TBM shields and allows the TBM to maneuver in a smaller curve radius.

2.4.4 Excavation Chamber

The excavation chamber is a critical component in EPB TBM. The chamber pressure should be maintained to balance the earth pressure to ensure excavation stability and prevent excessive ground movements. Too low pressures may cause excessive ground settlements, water inflow, and even instabilities at the excavation face (i.e., failures). In contrast, too high pressures may increase the cutter torque and thrust force, reduce the advance rate, and result in ground heaving. The chamber pressure can be regulated by adjusting the thrust jack speed (that controls the contraction and expansion of the chamber) and screw conveyor rotation speed (that controls the muck volumes discharged from the chamber). Some essential elements of the excavation chamber are (i) the bulkhead, which should have adequate strength and water tightness, (ii) the mixing blades, which are installed behind the cutter head to mix the excavated soils and the additives into the muck with favorable properties, and (iii) pressure sensors, which are critical as measurement tools in balancing the target earth pressures.

2.4.5 Ground Conditioning

The ground conditioning system is a critical component in EPB TBM tunneling to (i) reduce the permeability and shear strength of the excavated soil, (ii) enhance the plastic flow condition of muck in the excavation chamber, and (iii) reduce adhesion of the excavated soil to the shield (JSCE, 2016). Favorable muck properties may reduce cutter bits abrasion, cutter head torque, and screw conveyor torque. The key components of a ground conditioning system consist of additive, foam production, and injection systems.

- **Additives**. The common types of additives to be used as ground conditioners are:

  - Clay and bentonite minerals, to reduce the permeability and increase the fluidity of the excavated soils.
  - Foam, to reduce the adhesion and increase the fluidity and water tightness of the excavated soils.
  - Polymer, to absorb water or to increase the viscosity of the excavated soils.

- **Foam production**. Foam is the primary ground conditioner in EPB TBM tunneling. Foam solutions typically consist of a combination of water with surfactant (foaming agent) and stabilizer (polymer). The mixture is specified based on several ground conditioning parameters, such as
  - the concentration of foam agent in water (Cf),
  - the concentration of polymer agent in water (Cp),
  - the foam expansion ratio (FER), and
  - the foam injection ratio (FIR).

  Foam is produced in a foam generator by mixing the foam solution with compressed air. Figure 2.5 shows the principle of foam production in EPB TBMs. Foam bubbles fill the soil particles, reducing the soil's bulk density and shear strength. This condition produces favorable muck properties, i.e., lower stiffness and better fluidity. This enables better face pressure control and muck flow during discharging.

- **Injection system**. The ground conditioning injection system consists of pumps and nozzles. The nozzles are typically installed on both the cutter head (to condition the excavated soils) and the excavation chamber (to condition the muck). This system should allow control over ranges of injection pressures and rates as different ground conditions may require different additive recipes and ground conditioning treatments. The operators can specify the ground conditioner injection volumes by adjusting the injection pump, while the injection rate is typically adjusted according to the advance rate automatically.

2.4.6 Muck Discharge

The muck discharge system transports the excavated materials from the excavation chamber to the muck dumping system. It consists of two main components, i.e., the screw and the belt conveyors.

- **Screw conveyor**. The screw conveyor is the primary discharge system in an EPB TBM. The screw not only transfers the muck from the excavation chamber to the belt conveyor but also functions to regulate chamber pressures. This can be done by adjusting the rotation speed to increase or decrease the amount of discharged muck

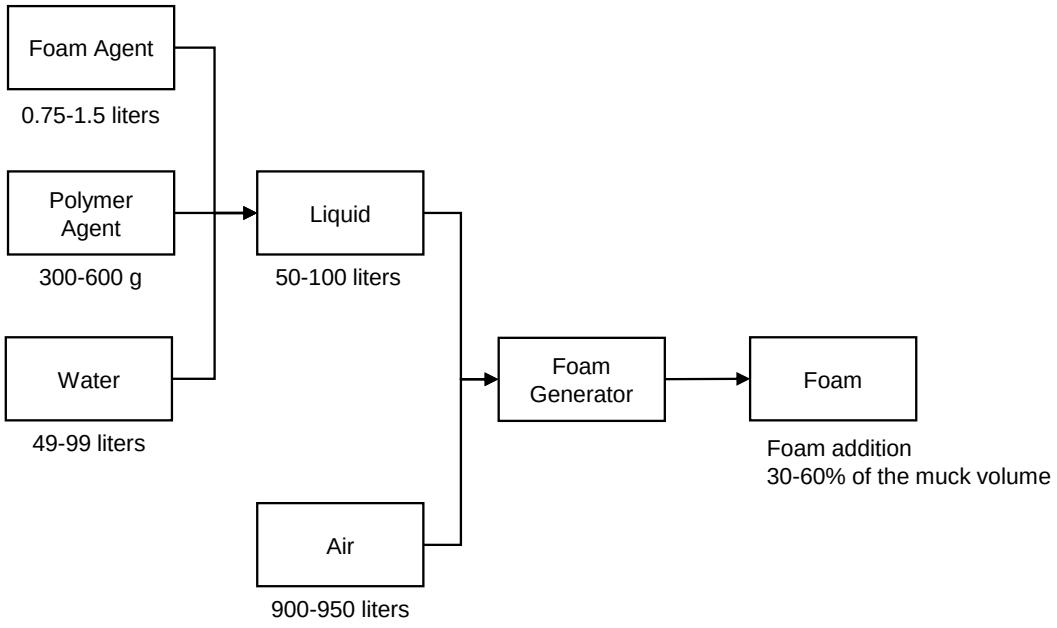

Figure 2.5: Schematic diagram of foam production system and the typical composition, adopted from Maidl et al. (2013)

from the chamber. Additional screw conveyors (i.e., secondary screw conveyors) can be installed if required, e.g., in tunneling under high water pressure or excavating soils with excessive gravel content.

- **Belt conveyor**. The belt conveyor transports discharged muck from the screw conveyor gate to the muck dumping system. This is where the operators and engineers can observe and evaluate the muck conditions. However, since the excavated soils spend some time in the excavation chamber, it should be noted that muck conditions on the belt conveyor cannot fully represent the encountered ground conditions at the excavation face.

2.4.7 Backfill Grouting

The backfill grouting system is required to fill the shield tail void. The grout material should have adequate fluidity to fill the gaps and promptly achieve the intended strength with only a little volume loss. The grout should be injected into the tail void and uniformly applied on the segment ring. The grouting should be carried out simultaneously with the shield advance to ensure the stability of the segment ring and to limit ground movements. The

injection can be controlled by pressure or volume. Typically, the grout injection pressure is specified as the current face pressure with an additional 200 kN/m2 (JSCE, 2016). Due to grout penetration, overcut, and other factors, the injected grout volume can be 130 to 170% of the theoretical tail void volume (JSCE, 2016).

2.5 EPB TBM Tunneling Process

TBM tunneling generally involves two main phases, i.e., the excavation (or advancing) phase, and the segment lining erection phase. The advancing phase is a continuous examination process that requires (i) specifying various target performances, (ii) adjusting control parameters, (iii) monitoring tunneling conditions, and (iv) readjusting control parameters according to the monitored tunneling conditions. This process is briefly discussed below.

2.5.1 Performance Indicators

The target performance of a tunneling process is specified as tunneling key performance indicators (KPIs). In EPB TBM tunneling, the KPIs may include several targets, such as chamber pressure, advance rate, and ground conditioning injection. The KPIs are commonly prepared by the tunnel engineers and handed to the operators via tunneling instruction sheets. Typical KPIs in the instruction sheet are listed in Table 2.1. The KPIs have to be determined for every tunneling cycle based on various bases, such as the tunnel design (e.g., target tunnel alignment), theoretical analysis (e.g., target earth pressures), or empirical performance from the previous cycles (e.g., ranges of ground conditioning injections). Determining the KPIs also involves judgments from the engineers since it is difficult to capture all complexities of ground-machine interactions. Therefore, it can be subjective and biased (Hammerer, 2015).

Some studies have been done to explore more systematic approaches to model the tunneling process and determine the performance indicators. Numerical modeling tools such as the Finite Element Method (FEM) enable computational simulation to investigate the tunneling process at some level (Kasper & Meschke, 2004, 2006a, 2006b; Komiya et al., 1999; Meschke, 2018). Nevertheless, building FEM models continuously during tunneling may not be practical due to the laborious modeling process and expensive computation cost. Recent advancements utilize data-driven approaches to model the tunneling process. This approach becomes more popular considering the availability of massive data produced during the tunneling (Meschke et al., 2019).

2.5.2 Tunneling Process Control

Determining and adjusting TBM control parameters to achieve the specified KPIs involve substantial judgments from the operators since there is no exact procedure to drive a TBM. Furthermore, the selected control parameters may or may not achieve the specified target

Table 2.1: Typical KPIs in a tunneling operation instruction sheet.

Process	KPI
Excavation	Maximum cutter torque
	Excavation volume
Advancing	Maximum thrust force
	Target advance rate
Steering	Target shield pitching
	Target jack stroke difference
Ground conditioning	Foam mix ratio & injection volume
	Polymer mix ratio & injection volume
Earth pressure balancing	Operating pressure range
Backfill grouting	Tail grout volume and pressure

performance due to unknown ground conditions and complex ground-machine interactions. Therefore, the operators have to monitor the TBM behaviors based on tunneling operation data and readjust the control parameters whenever necessary to meet the target performance. The readjustments must be performed continuously during tunneling. Figure 2.6 presents an example of excavation process control (management) in EPB TBM tunneling (JSCE, 2016). The figure shows that the excavation process consists of several connected partial process control. The key process control may include tasks related to (i) advancing and steering, (ii) excavation, which includes ground conditioning, (iii) earth pressure balancing, and (iv) ground movement controls. These control tasks are briefly discussed in the following.

2.5.2.1 Advancing and Steering Controls

The navigation system of a TBM includes information about the TBM spatial position and attitudes (i.e., pitch and yaw). Both are defined in both global and local coordinate systems. The local coordinate system represents the TBM position relative to the longitudinal axis of the designed tunnel alignment. Any deviation of the current TBM position to the designed tunnel alignment can be tracked and continuously presented on the control panel for steering. Operators must steer the TBM as near as possible along the given alignment.

Typically, deviation tolerance to the designed centerline is less than 50 mm. In thrust control TBM, the operators can steer the direction by selecting appropriate thrust jacks (in circular arrangement) to be activated. In pressure control TBM, the operator can set different pressure levels in each thrust jack or group of thrust jacks.

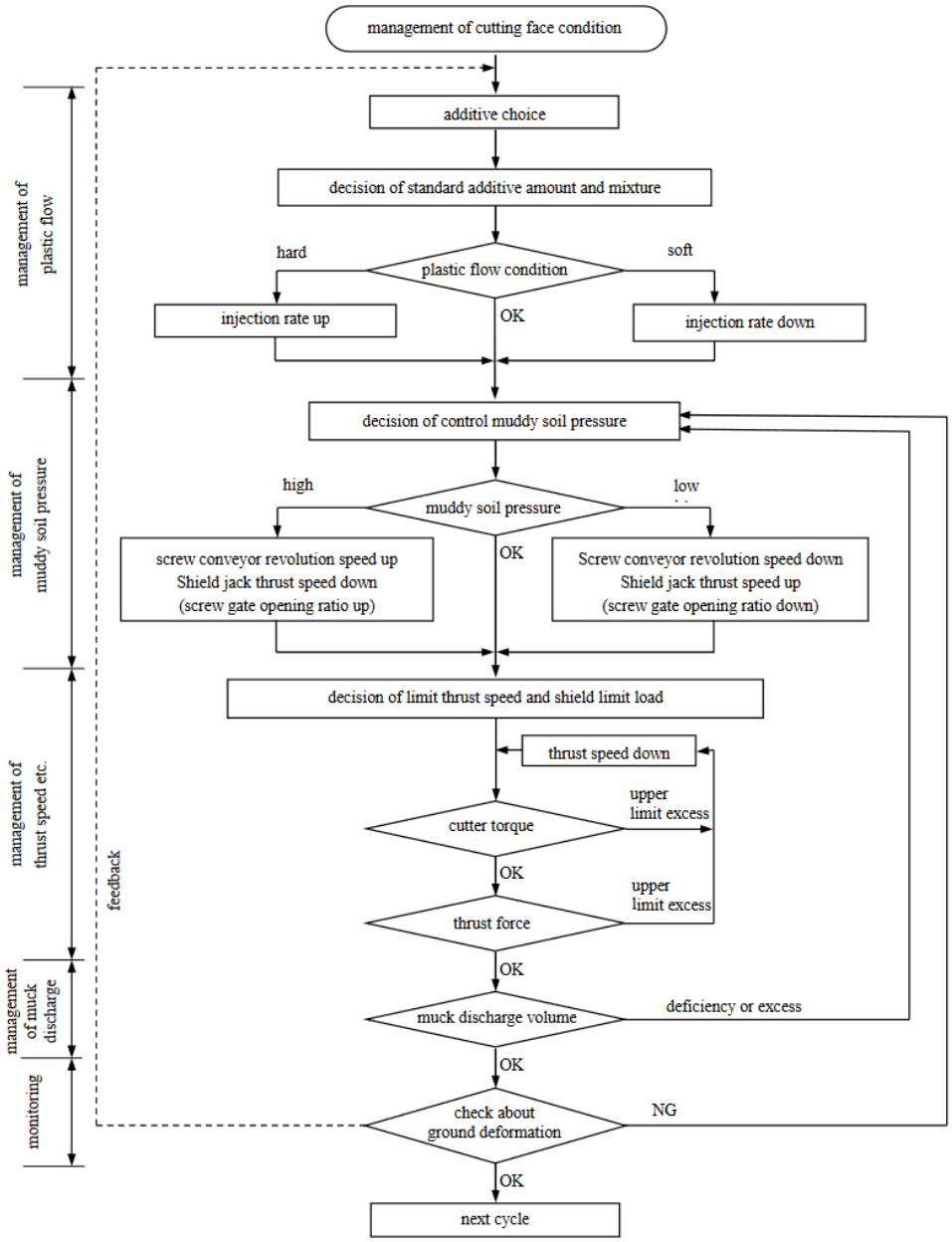

Figure 2.6: Excavation process control (or management) of cutting face conditions in EPB TBM tunneling (JSCE, 2016)

2.5.2.2 Excavation Control

Theoretically, a tunneling excavation process can be controlled based on the quantity of the excavated materials (muck). In this case, the discharged muck volume should balance the theoretical excavation volume. The weight and volume of the discharged muck can be measured on the belt conveyor by sensors such as belt scales and laser scanning. However, the properties of the discharged muck can be different from the unexcavated soils due to the injected ground conditioners and the change of the soil states (from solid to viscous fluid). Therefore, the excavation process cannot be controlled only by the excavated material quantities.

Plastic flow is another measure to control the excavation process. A favorable plastic flow condition of the excavated materials is critical in EPB TBM tunneling to ensure efficient excavation and pressure control. The process starts with setting additive materials and their quantities that are suitable for the current ground conditions. This task involves empirical estimation based on pre-tunneling soil laboratory testing and performance records from the previous tunneling cycles. The plastic flow conditions can be evaluated using several indicators, i.e., (i) muck properties by slump tests, (ii) muck efficiency by comparison between the estimated excavated muck from the advance rate and the screw rotation speed, and (iii) shield load by estimating the time-dependent change of the cutter torques and the screw conveyor. EPB TBM operators need to adjust the specified setting to ensure favorable muck plastic flow conditions. This process involves judgments from the operators and is often achieved by a trial-and-error process.

2.5.2.3 Earth Pressure Balance Control

Excavation chamber pressures should be maintained to balance the earth pressures. The upper and lower pressure limits are typically determined based on the pressure gradient. The earth pressure coefficient is estimated according to the ground movement. Injecting appropriate ground conditioners is also important to obtain favorable muck properties (such as adequate fluidity and low permeability) to facilitate a suitable pressure. The chamber pressure is mainly regulated by adjusting the thrust jack speed and the screw rotation speed.

The chamber pressure can be increased by raising the thrust jack speed or reducing the screw conveyor rotation speed. In contrast, the chamber pressure can be decreased by reducing the thrust jack speed or raising the screw conveyor rotation speed. Figure 2.7 shows a simplified scheme to maintain the chamber pressure at a specified target pressure. This scheme enables automated EPB control in typical tunneling conditions. However, the tunneling excavation process involves uncertainties and often cannot be represented in a theoretical model. Therefore, the experience and judgment of the operators are critical in performing this task (Maidl et al., 2013).

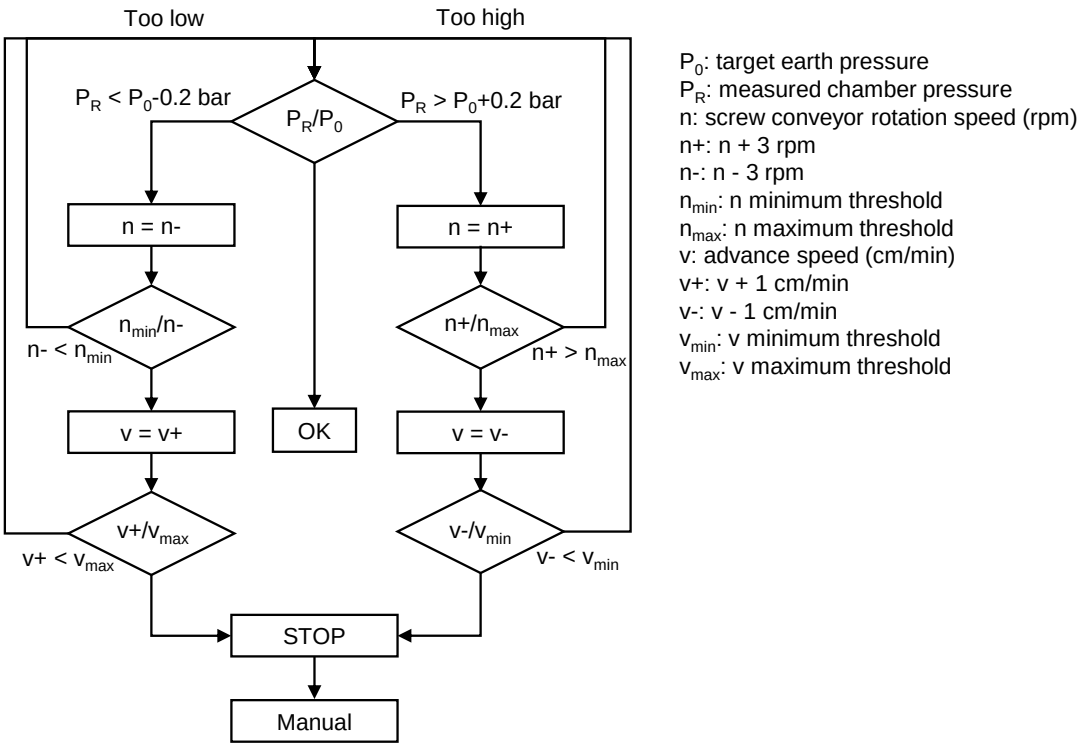

Figure 2.7: Example of EPB regulation flow diagram, adopted from Maidl et al. (2013).

2.5.2.4 Ground Movement Control

Tunneling-induced ground movements are governed by various factors such as tunnel spatial geometries (e.g., depth, diameter), the geologic conditions, and the tunneling process (i.e., the TBM behaviors). Adopting JSCE (2016), in the longitudinal direction, tunneling-induced ground movements can be categorized into (i) pre-tunneling, (ii) excavation face, (iii) shield passes, (iv) shield tail, and (v) post-tunneling phases. These phases are briefly discussed in the following.

- **Pre-Tunneling**. Pre-tunneling ground movements occur before the TBM passes. In cohesionless grounds, the movements can be caused by the decrease in the groundwater table. In soft cohesive grounds, the movements can be caused by ground flows at the excavation face (instability).
- **Excavation Face**. Ground movements in the proximity of the excavation face may occur due to the unbalanced pressure between the earth pressure and the pressure in the excavation chamber. Typically, ground settlements can occur if the earth pressure

exceeds the chamber pressure. In contrast, ground heaving can occur when the earth pressure is smaller than the chamber pressure.

- **Shield Passes**. Ground movements occur during the TBM passes due to interactions between the shield and the ground. Careless TBM driving may disturb the soils and aggravate the induced ground movements. Furthermore, over-excavation for alignment control can also loosen the soils and produce additional ground movements.

- **Shield Tail**. Ground movements occur immediately after the TBM has passed. Ground settlements can occur due to stress reliefs in the tail void. In contrast, ground heaving can occur due to excessive grouting pressures when filling the tail void.

- **Post-Tunneling**. Long-term ground movements may also occur in the post-tunneling phase. The settlements can be due to deformed tunnel lining or soil consolidation (in cohesive soils).

2.5.3 Data Management

The tunneling process produces massive data. Generally, the data include (i) geological or geotechnical data, (ii) geodetical data, (iii) TBM operation data, and (iv) ground monitoring data. These data sets are briefly discussed in the following.

- **Geotechnical data**. Geotechnical data include data from the pre-tunneling stage, i.e., boreholes and all other geotechnical investigations carried out in the design stage. The data may also include data produced during tunneling, e.g., muck properties obtained from soil testing. These data sets provide essential information on the geologic conditions along the tunnel alignment. However, the data is typically discrete, sparse, and limited.

- **Geodetic data**. Similar to the geotechnical data, geodetic data include data from the pre-tunneling stage, i.e., tunnel design alignment coordinates, and data produced during tunneling, i.e., actual TBM position and coordinates. In current tunneling practices, geodetic measurements are typically done automatically in digital formats.

- **TBM operation data**. Modern TBMs are typically equipped with various sensors that generate massive operation data during tunneling. Depending on the design, it can produce hundreds of variables (in small to medium-diameter TBMs) to thousands of variables (in large-diameter TBMs). Depending on the system, the data can be produced every second. Depending on the tunneling length and construction time, the amount of the data can easily reach Terabytes.

- **Ground monitoring data**. Tunneling also involves extensive instrumentation installation for ground movement monitoring. The type and quantity of the instruments depend mainly on the risk involved in the tunneling project. More instruments should

be deployed in critical tunneling projects, such as soft ground in densely populated regions. Typical ground instruments include building and ground surface level measurements, extensometers, inclinometers, and piezometers (Ning et al., 2019; Wan et al., 2017a, 2017b, 2019).

The above discussion highlights massive data produced in tunneling. This reveals the importance of computer-based data management systems in tunneling process control. Manual process control management may not be applicable since humans cannot make real-time decisions based on streaming data produced by numerous sensors. Implementing AI systems to support humans in making decisions based on this massive data remains an open research question (Garcia et al., 2021).

Chapter 3

Control System for EPB TBM Tunneling

3.1 Introduction

Tunnel boring machine (TBM) tunneling can be seen as a dynamical system with a controller that adjusts the input. This controller plays a vital role to regulates the TBM behaviors to achieve the expected tunneling performance. In manual control, this task is performed by humans (i.e., TBM operators). In an autonomous (or intelligent) control, this task is performed by artificial intelligence (AI) systems. This chapter briefly introduces the fundamentals of dynamical systems and control. Subsequently, this chapter provides a review of past studies related to the implementation of intelligent control systems for TBM tunneling. Finally, this chapter discusses the proposed AI-based framework for earth pressure balance (EPB) TBM tunneling. This framework is the basis of the AI systems used in the following chapters of this dissertation.

3.2 Dynamical System

The dynamical system is an input and output process with changing behavior over time (Figure 3.1), where the change can be due to external stimulation or disturbance (Åström & Murray, 2008). In this definition, tunneling can be seen as a dynamical system. In tunneling, the input is the control parameter, while the output is the TBM behaviors and tunneling performance. The input-output processes change depending on the ground conditions as the external stimulation (Figure 3.2). In other words, the same control parameters can result in different TBM behaviors and performance when performed in different ground conditions. Therefore, adjusting the control parameters according to the ground conditions is critical to achieving the desired tunneling performance.

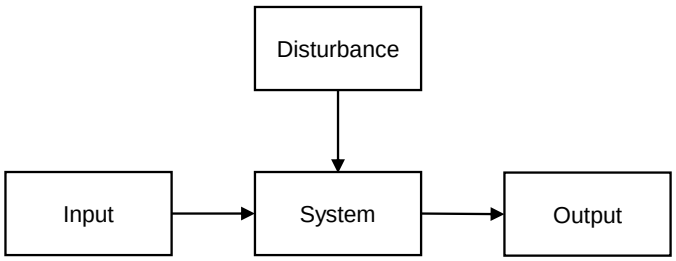

Figure 3.1: Fundamental scheme of a dynamical system.

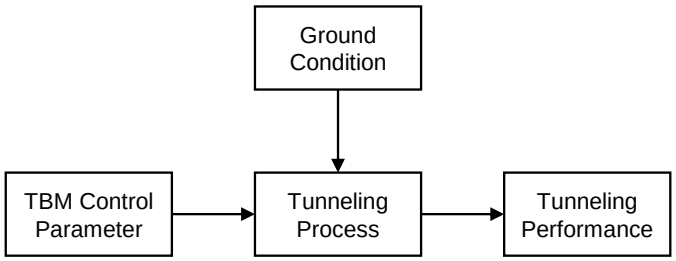

Figure 3.2: Tunneling as a dynamical system.

3.3 Feedback Control System

In a dynamical system, controlling the system input is essential to driving the system to achieve a required state. This task can be done by implementing a control system. Fundamentally, there are two types of control system loops: (i) open loop and (ii) closed (or feedback) loop control system (Figure 3.3). In an open loop system, the control action is independent of the output. In contrast, in a closed loop system, the control action depends on feedback from the system process. In this condition, a closed loop system ensures the controller adjusts the input to achieve a predetermined reference point based on the difference between the reference point and the actual measured outputs.

Tunneling is essentially a feedback control system (Figure 3.4). EPB TBM tunneling has several key performance indicators to be specified as reference points, e.g., the tunneling trajectory, excavation advance rate, and chamber pressure to balance the earth pressure. The TBM operators decide the control parameters and adjust the actuators, e.g., the jack

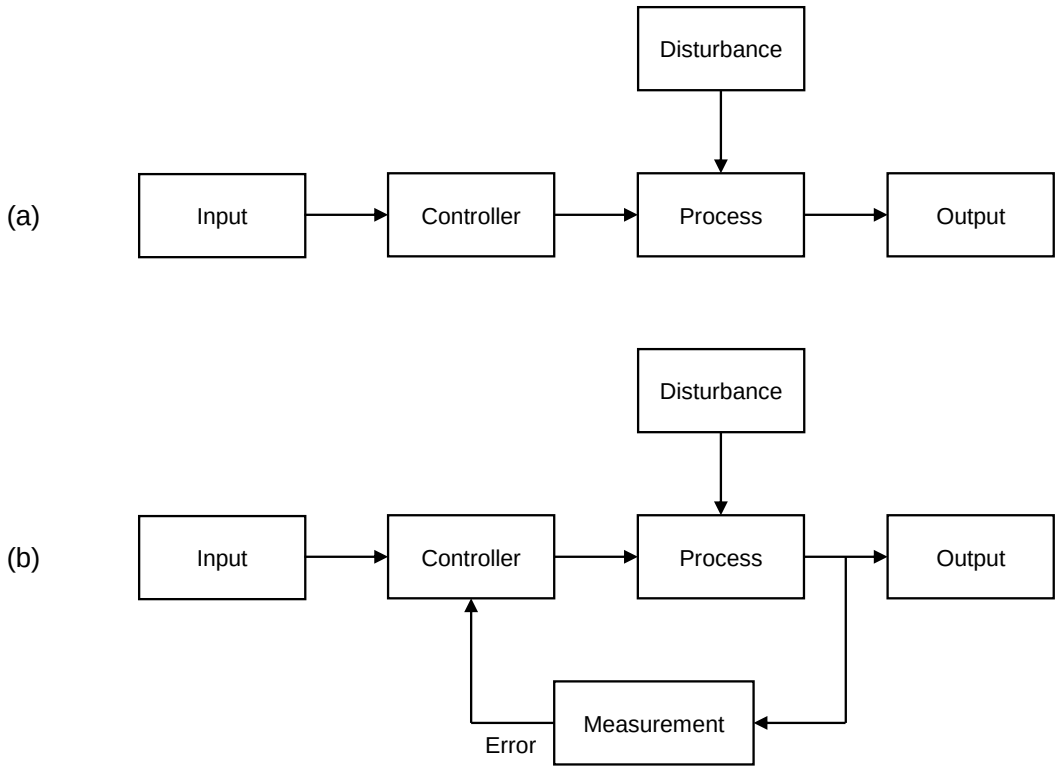

Figure 3.3: Control systems: (a) open loop, and (b) closed (feedback) loop.

selection, thrust jack speed, cutter rotation speed, ground conditioning injection, and screw rotation speed. The decisions govern the TBM behaviors and the actual tunneling performance, which are measured by numerous sensors during the tunneling. This operation data is presented back to the operators to make decisions for the next tunneling steps.

3.4 Controller

A feedback control system requires a controller to make control decisions. Manually, this role is covered by human decisions. In an autonomous (or intelligent) control system, this task is performed by an intelligent controller (e.g., computer-based models or algorithms). Many controller types have been developed in the domain of control theory. Generally, it can be categorized into linear control and model-based control.

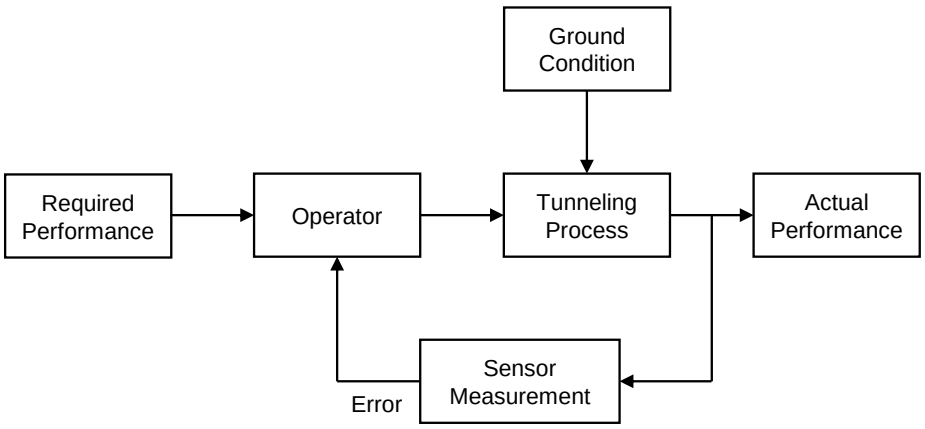

Figure 3.4: EPB TBM tunneling in a feedback control system framework.

3.4.1 Linear Control

The Proportional-Integral-Derivative (PID) controller is a linear control system that is widely used in many real-world control problems. In principle, a PID controller continuously calculates errors as the difference between measured outputs and the reference point. Subsequently, the controller applies a correction based on proportional, integral, and derivative terms. These three terms can cover the treatment of both transient and steady-state responses and therefore provide straightforward and efficient solutions (Ang et al., 2005).

Despite its efficiency, the PID controllers have limitations. To achieve optimal performance, a PID controller needs to adjust its parameters when the dynamical process is altered. This can be challenging if there are time delays or changes in the system process (Efheij et al., 2019). This is where model-based controllers have an advantage, as they can anticipate future steps. Note that tunneling involves many processes with inherent complexity due to the changing ground conditions and the causal-effect relationships of ground, machine, and human control decisions.

3.4.2 Model-Based Control

A more advanced control system is required to build controllers in complex dynamical systems, such as the tunneling process. Model-based predictive control (MPC) uses models (i.e., governing equations or algorithms) that consider various process constraints and disturbances to predict the next step and adjust the control parameters (Schwenzer et al., 2021) (Figure 3.5). The effectiveness of this controller strongly depends on the predictive

model. There are two options for building a model-based controller: (i) theory-based and (ii) machine learning-based models.

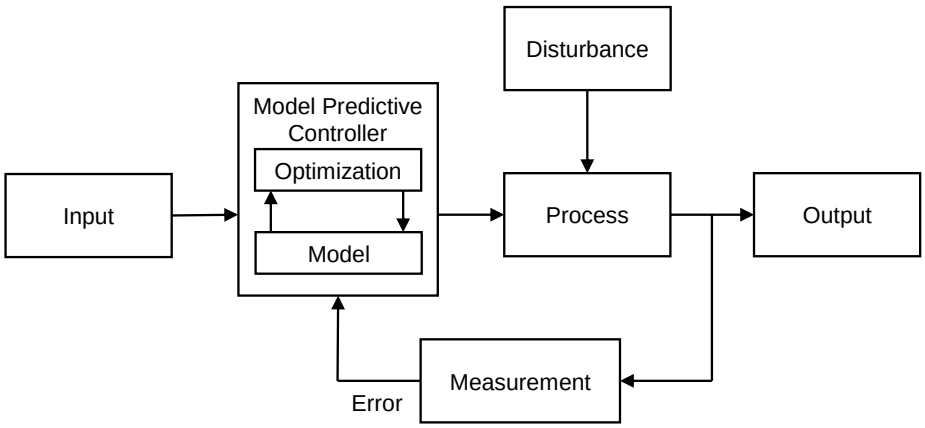

Figure 3.5: Model-based predictive control.

3.4.2.1 Theory-Based Model

Theory-based predictive models are built by implementing physical laws as the governing equations. The model is initially constructed based on detailed physical descriptions of the system. Subsequently, the model is updated according to the measured data. Inference can be made based on the changes.

The physical law-based model is relatively interpretable and generalizable. However, the model may have some limitations:

- The model is typically constructed with an idealization (i.e., simplification) due to modeling complexity and computational cost. Therefore, it cannot represent all details in the real object and may lose much information.
- The model may not be able to represent the system accurately due to uncertainty and unknown process.

This can be a significant issue in modeling a complex process such as TBM tunneling. Not all processes in tunneling can be modeled deterministically using a governing equation.

3.4.2.2 Data-Driven Model

Data-driven models are built purely using data reconstruction and fitting techniques. Unlike theory-based models, data-driven models require limited or even no predefined assumptions. Machine learning algorithms are commonly used for this approach. The model is constructed using available data and continuously developed as the data streams are generated. Inference can be made by identifying patterns in the data.

The model should be able to represent complex reality as it is built by measurements of the real object. No idealization, simplification, or assumptions are required. However, the model may also have some limitations:

- The model strongly depends on the data acquisition design, i.e., the availability and quality of the data.
- The model is less interpretable as many machine learning models are seen as "black box" models.
- In the supervised learning approach, the model will be biased toward the training labels.

Machine learning models have become more popular to be used as intelligent controllers for complex dynamical systems, especially where the existing theoretical model cannot accurately describe the actual process. This is further boosted by the explosion of data availability, the increase in computing power, and the advancement of machine learning algorithms in the past decade.

3.5 Intelligent Control System for TBM Tunneling

The development of feedback system framework and intelligent predictive control systems have triggered the advancement of autonomous technology in various applications, such as Unmanned Aerial Vehicles (UAV) (Goerzen et al., 2009) and self-driving cars (Badue et al., 2021; Paden et al., 2016). However, these advancements cannot be directly implemented in autonomous tunneling systems. TBM tunneling has features and complexity that differ from other autonomous objects. Adopting Hu et al. (2022), some challenges in developing autonomous TBM are listed below.

1. TBMs maneuver within an unconstrained trajectory in three-dimensional (3D) space. The shield TBM movement should also be treated as a 3D segment rather than a single-point coordinate. This is because the head, articulation point, and tail movements of a shield system are interrelated.
2. The encountered ground conditions strongly govern the behaviors of TBMs. However, there is no accurate geologic information available during the tunneling process. The excavated ground condition cannot be seen and constantly changing.

3. TBM tunneling has several performance indicators to be achieved simultaneously. For example, the operators must maintain the trajectory within a tolerable deviation along the tunnel design alignment. At the same time, they must regulate the chamber pressure to be equivalent to an estimated earth pressure to achieve excavation stability and limit ground movements. They also require completing the tunneling process within a specified schedule.

4. TBMs have many components and control parameters. Each component has independent actuators that must be controlled simultaneously with the others. For example, the operators must select the thrust jack configuration and pressures to maneuver the shield along the tunnel design alignment. At the same time, they need to adjust the thrust jack speed and screw rotation speed to regulate the chamber pressure. They must also decide on the ground conditioning injection, cutter rotation speed, and thrust jack speed to achieve the target advance rate. Decisions on each control parameter may interact and affect the others.

Despite these challenges, some efforts have been made to develop autonomous TBM systems in the past three decades. The early development focused on providing intelligent control for a single tunneling task, such as a direction control system and a chamber pressure regulation system. The latest development attempted to build an integrated control system and implement it in real tunneling projects.

3.5.1 Direction Control

Uematsu et al. (1996) reported an early development of automated steering control systems using automatic control direction flow. In principle, the rule-based system was performed by setting a target line (i.e., the tunnel design alignment) and a tolerable deviation limit. Steering operations were executed by specifying the thrust jack stroke difference if the front shield of the TBM exceeded the deviation limit. The applied thrust jack stroke difference was determined based on its statistical relationship to the deviations from past data. This guidance system was implemented for hard rock tunneling in the Tsukui headrace tunnel project in Japan by a Joint venture of Kumagai, Goyo, and Dainippon.

In the decades of the 2000s and 2010s, the development of TBM direction control systems was dominated by numerical and experimental simulations of feedback control systems with various controllers and settings, such as PID controller (Huayong et al., 2009), two closed-loop control structures (Yue et al., 2012), cascade strategy with fixed-value feedforward and variable PID controllers (Xie et al., 2012), multi-cylinder control system (L. Wang et al., 2018), and fuzzy PID controller (Wu et al., 2022). Besides, some studies performed numerical simulations to investigate the interaction between TBM movements and ground conditions. These studies were commonly done using the Finite Element Method (FEM) (Kasper & Meschke, 2004; Komiya et al., 1999). The latest development includes detailed steering process modeling (Alsahly et al., 2019; Alsahly et al., 2016).

The development of autonomous TBM direction control systems emerged recently in the 2020s. Many of the research and developments were carried out by the industry. This allows the implementation of the developed system in real tunneling cases. Hoshino et al. (2020) developed a directional control learning system for shield TBMs. The system worked for force control shield operation and has been implemented by Obayashi Corp. at multiple sites in Japan. It appears that no machine learning model was implemented in the system, as they mentioned that some improvements could be made by linking the system with machine learning.

Wada et al. (2021) developed a directional control guidance system for shield TBMs using a machine learning algorithm as the predictive model. The developed system used the gradient boosting algorithm, a supervised learning algorithm based on sequentially ensembled decision trees. The model used 14 input features, which consisted of operation data related to the machine coordinates, attitudes, deviations, and advancing features (i.e., thrust force, cutter torque, advance rate). The models predicted the optimal point of the resultant force produced by thrust jacks, which control the maneuver of a shield TBM. The system was implemented in a real tunneling project by Shimizu Corp.

Despite the promising results, the study noted that the performance of the guidance system deteriorated when the tunnel alignment rate changed. Furthermore, the model has generalization problems. As every tunneling project has different geologic conditions and TBMs, the model developed in a tunneling project cannot be implemented in other projects.

3.5.2 Earth Pressure Balance Control

Earth pressures in an excavation chamber depend on the geologic conditions and operational factors, such as ground conditioning and mixing method, advance rate, and screw rotation speed (Anagnostou & Kovári, 1996). H. Yang et al. (2009) proposed a theoretical model for EPB control based on the dynamic equilibrium of intake volume and discharge volume. The mathematical model was developed on the assumption of plasticized soil in a closed soil chamber with two key parameters to achieve earth pressure balance. The thrust jack speed represents the intake volume, and the screw conveyor speed governs the discharge volume. An EPB control system was developed based on this theoretical model. The proposed control system was tested in an experiment with controlled soil conditions.

X. Liu et al. (2011) proposed a model predictive control strategy for EPB. The proposed predictive control system consisted of predictive and optimization models. The Least Squares-Support Vector Machine algorithm (LS-SVM) was used as the predictive model, and the Particle Swarm Optimization algorithm (PSO) was used as the optimization model. The predictive model was used to estimate earth pressures in the excavation chamber based on the screw conveyor rotation speed, advance rate, cutter rotation speed, and thrust force. The optimization model was used to determine the optimal EPB control parameter (i.e., the advance rate and screw conveyor rotation speed) to achieve the required earth pressure.

Many studies work on this problem using the same strategy but with different predictive and optimization models. K. Li and Shao (2015) used the Adaptive Neuro-Fuzzy-based

Inference System algorithm (ANFIS) as the predictive model and the Ant Colony System algorithm (ACS) as the optimization model. X. Liu and Zhang (2019) used LS-SVM as the predictive model and ACS as the optimization model. Gao et al. (2020) used a neural network-based algorithm with a Gated Recurrent Unit (GRU) as the predictive model and the genetic algorithm as the optimization model. X. Liu et al. (2020) used Backpropagation Neural Networks (BPPN) in Dual Heuristic Programming (DHP) as the prediction and optimization model.

The discussed studies show that data-driven model predictive control, consisting of predictive and optimization models, has been a popular option for developing an intelligent EPB control system. The data-driven model enables flexible modeling based on data as the process in the excavation chamber is complex, involves uncertainties, and cannot be accurately represented by a theoretical model.

3.5.3 Integrated Control with Minimal Human Intervention

Xiong et al. (2020) developed autonomous control algorithms to steer and operate shield TBMs. The predictive system inside the controller was not clearly described, but it appears to be some rule-based algorithms embedded in a closed-loop control system. The input of this system is not clearly listed as well, but it contained features such as the shield position and deviations. The feedback system was used to determine the optimal thrust jack pressure assignments to direct the machine toward the design tunnel alignment. The jack pressure assignments were adjusted automatically and continuously in real-time as the data was fed into the system every 3 seconds. They also used the system to regulate the contact force and penetration rate to protect the cutting tools.

The autonomous system was deployed by MMC-Gamuda KVRMT(T) Sdn., Bhd. in the Klang Valley MRT Line 2 tunnel in Kuala Lumpur, Malaysia. The tunneling project employed variable density shield TBM with multi-mode operation capability, i.e., slurry and earth pressure balance modes. The variable density shield TBM was used due to the variation of the geologic conditions along the tunnel alignment, including Karstic Limestone, Kenny Hill, and granite formations. The tunneling was performed with minimal human intervention in a tunneling section of about 5 km. The tunneling results were reported to seamlessly adapt to the changing ground conditions and improve the operation's consistency, safety, and accuracy.

3.5.4 Integrated Control with No Human Intervention

Hu et al. (2022) developed an autonomous EPB TBM control system based on the human performance model (Rasmussen, 1983). Rasmussen categorized human behavior into three levels based on task difficulties.

1. The skill-based level, when humans perform simple actions solely based on feedback signals without requiring substantial attention.

2. The rule-based level, when humans perform actions based on their experience.

3. The knowledge-based level, when humans perform actions on a task that has never been performed before and therefore requires a complex decision-making process.

Similarly, Hu et al. (2022) decomposed tasks in shield tunneling into three levels, namely, (i) the execution level, (ii) the planning level, and (iii) the strategic level. The execution level corresponds to the skill-based level in humans. This task includes adjusting control parameters (e.g., screw rotation speed, grouting system injection) given a target performance (e.g., target chamber pressure, target grouting volume) and can be solved using a basic controller such as PID and fuzzy-based controller. The planning level corresponds to the rule-based level. This task includes setting multiple target performances simultaneously and can be solved using Neural Network-based algorithms (Hu et al., 2022). In typical operations, these targets are determined by tunnel engineers and operators based on engineering knowledge and judgment from the previous tunneling performance. The strategic level corresponds to the knowledge-based level. This task includes setting multiple target performances simultaneously in changing environments, i.e., geologic conditions, and was proposed to be solved using methods such as knowledge graphs (Paulheim, 2017).

The system was implemented in the intercity railway tunnels between Hangzhou and Shaoxing, China, by Shanghai Tunnel Engineering Co., Ltd. (STEC). The tunneling was performed using an articulated EPB TBM called "Zhiyu". The EPB TBM was equipped with intelligent modules to access and control various set points, i.e., cutter rotation, thrust jack speed, jack selection, earth pressure, screw rotation, gate opening, as well as tail grease and grouting quantity. Unfortunately, the details of the intelligent control modules were not clearly described. The tunneling was completed without human intervention in some tunnel sections. The tunneling results were reported to meet the engineering requirements (i.e., trajectory deviations and ground movements) and offer better performance than the manual control results.

Despite the promising results, Hu et al. (2022) admitted that the system needs to be readjusted and retrained in different ground conditions. This means the system was not fully intelligent for learning and adapting to changing ground conditions, which is the main challenge in tunneling. Furthermore, although the system appears to be comprehensive, the hierarchy levels may constrain it to be flexible and adaptable. Note that a TBM can be driven in various ways depending on the driving style and requirements (Hammerer, 2015).

3.6 Proposed Framework for EPB TBM

3.6.1 Cognitive Model

An intelligent controller should behave similarly to the human cognitive system and decision-making process. Biologically, the human brain works on the framework of sensing, perceiving, and acting (O'Shea, 2005). Sensory information such as sight, touch, audition, taste, and

smell flows into the brain. The brain transforms this sensory information into perceptions. The perceptions will be the bases for decisions on future actions (Figure 3.6).

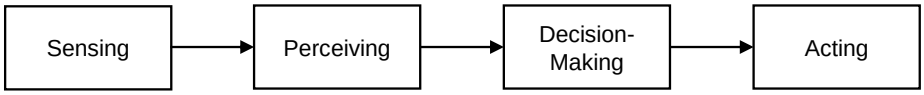

Figure 3.6: Human cognitive model.

This simple framework can be used to build an intelligent system for EPB TBM tunneling (Figure 3.7). The sensing system consists of tunneling data. The key component is the numerous operation data that is generated by numerous sensors every few seconds during tunneling. The other data can be pre-tunneling data (e.g., geotechnical data) and ground monitoring data (e.g., ground instrumentations and remote sensing data). The perceiving system consists of intelligent systems such as machine learning models that can convert the data into useful information for tunneling processes, such as the encountered ground condition interpretation and the induced ground movements. The decision-making system consists of intelligent systems that can model the complexity of interrelated tunneling processes and make optimal decisions for adjusting the control parameters.

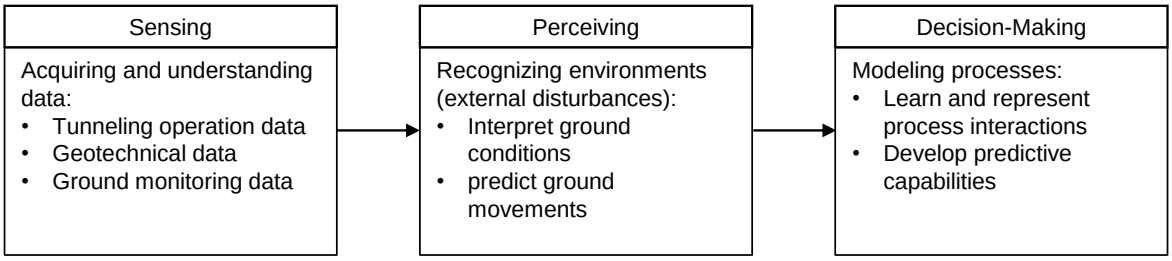

Figure 3.7: Implementation of the cognitive model for intelligent EPB TBM systems

3.6.2 AI-Based System

The proposed framework of AI-based control systems for EPB TBM is developed based on the feedback control system (Figure 3.8). The AI systems play a central role in the controller module and are developed according to the previously discussed cognitive model. The

systems leverage both streams of sensor measurement data are generated at the particular tunneling phase (or time step) and recorded data from the prior phases to (i) perceive the tunneling environments, i.e., the ground conditions and responses, and (ii) model the tunneling process and make decisions on the control parameters. These data-driven systems are built using machine learning algorithms, which enable the task to be completed in real-time during tunneling. More discussion on the feasibility of empowering feedback loop control systems with AI systems can be found in a study by Schöning et al. (2022).

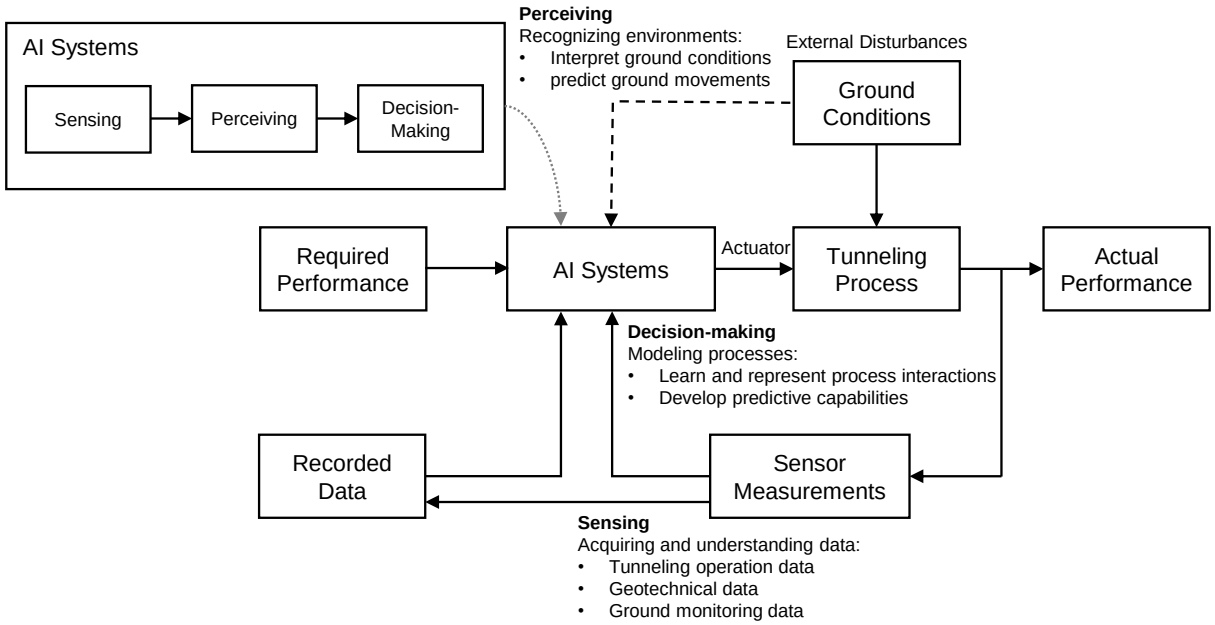

Figure 3.8: Intelligent control system for EPB TBM

Chapter 4

Machine Learning for EPB TBM Tunneling

4.1 Introduction

The availability of massive data, the development of machine learning, and the expansion of computing performance have been the main reasons for the rise of artificial intelligence (AI) in the past decade. Machine learning has emerged rapidly in the past few years and has been successfully implemented to exploit big data in various domain applications. Many studies have also attempted to implement machine learning in tunneling. This can be important since a tunneling operation produces enormous amounts of data that cannot be fully utilized using traditional data processing techniques. This chapter briefly introduces the fundamentals of machine learning, including the general understanding of machine learning concepts and some mathematical basis of the algorithms that are used throughout this dissertation. Subsequently, this chapter also briefly reviews some machine learning implementations in tunneling applications.

4.2 Machine Learning

4.2.1 Overview

Machine learning can be defined as algorithmic systems that are capable of "learning" (or improving their performance) through experience (Jordan & Mitchell, 2015). Fundamentally, machine learning is a set of algorithms that use statistical and computational principles to find patterns in a data set. Different from traditional rule-based methods, machine learning does not need to be explicitly programmed to achieve a specified task. The learning process is done by feeding the algorithm with more data. Machine learning has emerged intensely over the past two decades. It has been implemented in various applications, for example, developing AI systems (e.g., computer vision, natural language processing, robotics), solving

problems in data-intensive industries (e.g., processing internet and commercial data), and providing a novel way for exploring massive experimental data in scientific research.

The two main branches of machine learning are supervised and unsupervised learning, with additional emerging branches such as semi-supervised and reinforcement learning (Figure 4.1). Supervised learning can be seen as a new form of predictive modeling. In this learning approach, a set of answers in the form of training data is provided to a supervised learning algorithm. The algorithm fits the training data to build a prediction model. This approach can be used either for classification or regression tasks. The least squares method (more widely known as linear regression) can be seen as a basic form of supervised learning. More sophisticated supervised learning algorithms do not require any predefined assumptions such as linearity, normality, or even predefined predictor variables.

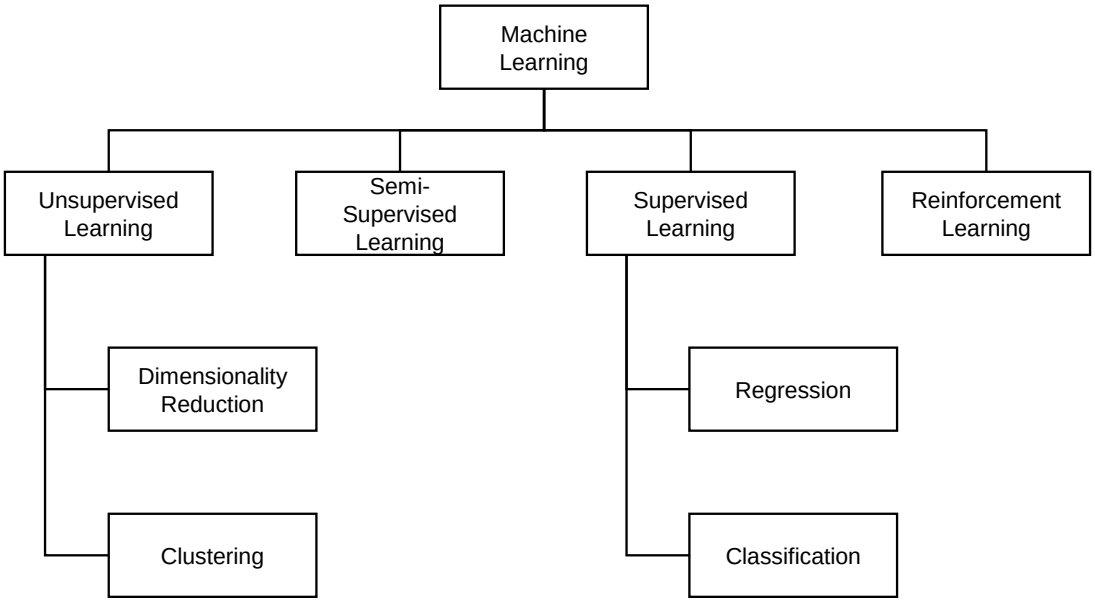

Figure 4.1: General taxonomy of machine learning.

Unsupervised learning can be seen as a new form of descriptive modeling. Unlike supervised learning, unsupervised learning algorithms do not require training data to find patterns and structures in the given data. This learning approach can be used to reduce the dimensionality of the data and cluster the data points. Dimensionality reduction algorithms can find patterns in the data by projecting it into lower dimensional space. This can be useful to extract data with thousands of variables into only two or three variables so that they can be plotted on two or three dimension coordinates without sacrificing too much information within the original data set. Clustering algorithms can identify data points with

similar characteristics and group them together. Both supervised and unsupervised learning approaches were used to develop AI systems for EPB TBM in this study.

4.2.2 Algorithms and Selection

Many machine learning algorithms are available to perform the tasks mentioned above. Some algorithms can be used exclusively for a specific task; for example,

- dimensionality reduction: Principal Component Analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (t-SNE),
- clustering: k-means, Agglomerative Hierarchical Clustering, and Density-based Spatial Clustering of Applications with Noise (DBSCAN),
- regression: Ordinary Least Squares (OLS), and
- classification: Logistic Regression.

Some supervised learning algorithms can also be used for both classification and regression tasks; for example, k-Nearest Neighbor (kNN), tree-based algorithms (such as Random Forest and Gradient Boosting), and Support Vector Machines (SVM). Lastly, neural network-based algorithms can be used for both supervised and unsupervised tasks. These algorithms are popularly known as deep learning.

Since many options are available, selecting a machine learning algorithm that works best for a particular problem is important. Typically, the trade-off between model flexibility and interpretability is one of the primary considerations in the selection. Some algorithms are inflexible, which means that these algorithms are only capable of fitting the data in a presumed pattern. For example, Linear and Logistic Regressions rigidly fit the data in linear relationships. Some other algorithms are more flexible and can fit different patterns in the data. For example, non-parametric algorithms such as Random Forest and SVM can fit nonlinear patterns in the data without any predefined assumptions. Nevertheless, flexible algorithms may also have disadvantages. Machine learning models built using these algorithms are prone to overfitting and more difficult to be interpreted. Figure 4.2 illustrates the trade-off between the flexibility and interpretability of machine learning methods.

Model accuracy is another primary consideration in the selection. To date, model accuracy has been the main focus of machine learning research. The aim is to build a machine learning model with the highest prediction accuracy. This can be one of the main reasons for the popularity of deep learning-based algorithms, such as Multi-Layer Perceptron (MLP), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN). Deep learning models are exceptionally flexible and thus can produce superior accuracy in various tasks and data types, such as image segmentation, text recognition, and natural language processing. Deep learning has become a standard practice in working with unstructured types of data sets.

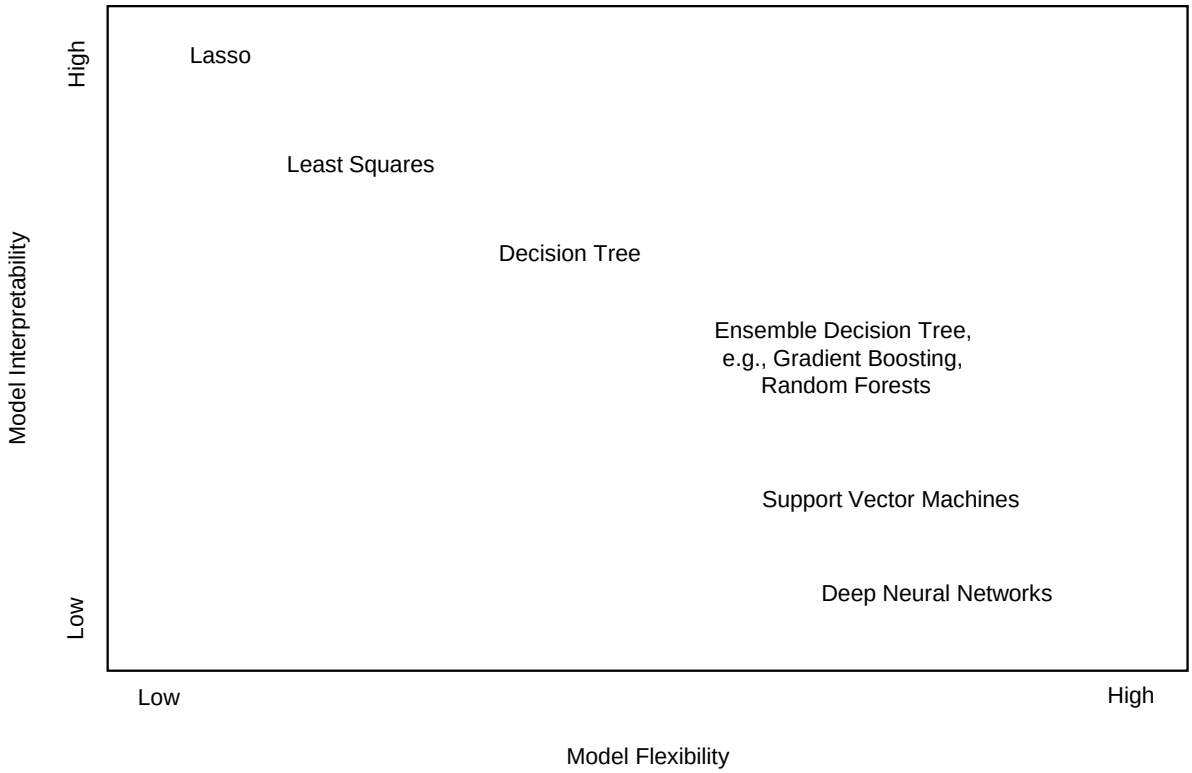

Figure 4.2: Model flexibility vs. interpretability, modified from James et al. (2013).

Despite its superior performance in unstructured data, deep learning has been argued not to be the best option for working with structured (or tabular) data sets. Gorishniy et al. (2021) compared many existing deep learning architectures for tabular data sets on diverse tasks under the same training and tuning protocols. Subsequently, the best deep learning performer was compared to Gradient Boosting. They concluded that there is no superior solution that can be accepted universally. Shwartz-Ziv and Armon (2022) compared the performance of recently proposed deep learning models for tabular data sets to Extreme Gradient Boosting (XGBoost). They concluded that XGBoost outperformed the deep learning models across the data sets and was more practical since it required less tuning.

Grinsztajn et al. (2022) compared various deep learning architectures and tree-based models (i.e., XGBoost and Random Forest) across different tabular data sets and hyperparameter combinations. Their results show that, for medium-sized data (10K samples), tree-based models produced better prediction performance with less computational costs. Their further investigation reveals that deep learning models are biased toward overly smooth solutions

and can be more affected by uninformative features. The above recent studies indicate that deep learning may not be the best model when working with structured data. In this type of data, the ensemble decision tree-based algorithms, such as Gradient Boosting and Random Forest, can deliver better prediction performance with more practical hyperparameter tuning and less computational cost.

4.2.3 Mathematical Basis

This dissertation consists of various tasks that involve different types of machine learning models. The summary of the tasks and the employed machine learning algorithms is presented in Table 4.1. The fundamental bases of algorithms that are used in multiple chapters are discussed in the following sections. The fundamental bases of some other algorithms, such as PCA and Bayesian network structure learning (BNSL), are discussed in the corresponding chapters.

Table 4.1: Summary of machine learning tasks and algorithms in this dissertation

Chapter	Modeling objective	Task	Algorithm
Chapter 5	Evaluating the effects of data aggregation	Regression	OLS, Random Forest
Chapter 6	Evaluating the effects of feature selection	Regression	OLS, Ridge regression, Lasso, Random Forest
Chapter 7	Interpreting geologic conditions	Classification and regression	OLS, Random Forest
Chapter 8	Detecting geologic anomalies	Dimensionality reduction (embedding)	PCA
Chapter 9	Estimating ground responses	Regression	OLS, Random Forest
Chapter 10	Modeling excavation process	Probabilistic modeling	BNSL
Chapter 11	Modeling steering decisions	Multioutput regression	Multivariate Output, Random Forest

4.2.3.1 Ordinary Least Squares

The OLS was selected to represent linear and parametric prediction models. In principle, the linear model defines a vector of a response variable, y_i , as a linear function of the predictor

variables, $x_1, x_2, \ldots, x_p$,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i, \tag{4.1}$$

where β is the regression coefficients and ε is the residual error. In OLS, the coefficients can be estimated by minimizing the residual sum of squares, RSS,

$$\begin{aligned} RSS &= \sum_i \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \hat{y_i})^2 \\ &= \sum_{i=1}^{n} \left(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \cdots + \hat{\beta}_p x_{ip}) \right)^2 \\ &= \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{p} \hat{\beta}_j x_{ij} \right)^2 \end{aligned} \tag{4.2}$$

This minimization can be solved efficiently in matrix algebra as

$$\hat{\beta}_{OLS} = \min RSS = (X^T X)^{-1} X^T y, \tag{4.3}$$

where X is the matrix of predictors and y is the vector of the response variable (James et al., 2013).

4.2.3.2 Ridge Regression

The Ridge regression is a type of regularized linear prediction model. The regularization can reduce the variance making it more robust to overfitting, thus leading to better generalization. Besides, Ridge regression can deal with data sets with multicollinearity problems. Ridge regression has the same RSS as the OLS. But the coefficient $\hat{\beta}^R$ is estimated by minimizing RSS with a constraint

$$RSS + \lambda \sum_{j=1}^{p} \beta_j^2 = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{p} \hat{\beta}_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2, \tag{4.4}$$

Where $\lambda \geq 0$ is a tuning parameter. The tuning parameter controls the relative impact of the shrinkage penalty (i.e., the $\lambda \sum_{j=1}^{p} \beta_j^2$ term) on the regression coefficient estimates. The shrinkage penalty term will not affect the RSS term when the tuning parameter is equal to zero, giving the regression to be equivalent to OLS. In contrast, the shrinkage penalty will reduce the regression coefficient estimates to approach zero when the tuning parameter is high. Selecting an appropriate value of the tuning parameter is critical and can be done by performing cross-validation.

4.2.3.3 The Least Absolute Shrinkage and Selection Operator (Lasso)

The Least Absolute Shrinkage and Selection Operator (Lasso) is another type of regularized linear prediction model. In the Lasso, the coefficient $\hat{\beta}^L$ is estimated by minimizing RSS with an absolute constraint

$$RSS + \lambda \sum_{j=1}^{p} |\beta_j| = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{p} \hat{\beta}_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j|. \tag{4.5}$$

Unlike Ridge regression which keeps all predictors in the prediction results, the Lasso may shrink some of the predictors to zero. This shrinkage capability leads to better interpretability of the prediction results and makes the Lasso an embedded feature selection method.

4.2.3.4 Random Forests

The Random Forest (Breiman, 2001) was selected to represent nonlinear and non-parametric machine learning models. As discussed, compared to more advanced models (e.g., deep learning), the ensemble decision tree-based models (e.g., Gradient Boosting and Random Forest) have been shown to result in better prediction performance with less computational costs in structured data sets (Gorishniy et al., 2021; Grinsztajn et al., 2022; Shwartz-Ziv & Armon, 2022). Random Forest was preferred over Gradient Boosting as it requires more straightforward hyperparameter tuning and is more robust to overfitting problems.

Random Forest is an ensemble supervised learning algorithm that aggregates a large number of decision trees as the base learners to produce the predictions (Breiman, 2001). Briefly, the algorithm is performed by the following steps (Hastie et al., 2009; James et al., 2013):

1. Generate B bootstrapped training data sets by drawing samples from the original training data set with replacement.
2. Construct decision trees, $\hat{f}^{*b}(x)$, using each of the bth bootstrapped training data sets. A single decision tree is constructed using binary recursive partitioning into distinct subsets so that one parent node leaves two child nodes (Therneau & Atkinson, 1997).
   a) Randomly select m predictors from the p predictors.
   b) Choose the best-split point among the m.
   c) Split the node into two child nodes.

   In greedy recursive partitioning, the objective is to find regions, R_1 and R_2, which minimizes the RSS.

$$R_1(j, s) = \{X | X_j < s\}, \qquad R_2(j, s) = \{X | X_j \geq s\}, \tag{4.6}$$

where j is the selected predictor to split, and s is the cutpoint. The value of j and s can be computed by minimizing the RSS, defined as

$$RSS = \sum_{i:x_i \in R_1(j,s)} (y_i - \hat{y}_{R_1})^2 + \sum_{i:x_i \in R_2(j,s)} (y_i - \hat{y}_{R_2})^2 \tag{4.7}$$

The recursive partitioning is done until the specified minimum node size is reached.

3. Obtain the final result by aggregating (averaging) all the constructed decision tree results,

   a) For regression:

   $$\hat{f}_{RF}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}^{*b}(x). \tag{4.8}$$

   b) For classification:

   $$\hat{C}^B_{RF}(x) = majority\, vote\{\hat{C}_b(x)\}_1^B \tag{4.9}$$

   Where $\hat{C}^B_{RF}(x)$ is the class prediction.

As a non-parametric model, Random Forest does not require the data to meet certain assumptions or parameters (Malley et al., 2012). Thus, it is suitable for data in which the distribution cannot be predetermined. Random Forest has become popular in classification and regression problems due to its strong predictive performance in high-dimensional data, versatility in various feature types and scales, robustness to missing data, outliers, and noises, as well as its ability to measure the importance of the features (Biau & Scornet, 2016; Qi, 2012).

4.2.3.5 Multivariate Output Random Forests

The Multivariate Splitting-rule Random Forest (Ishwaran et al., 2021) was selected to perform predictions with multivariate outputs. In principle, Random Forest regression can be extended to multivariate output cases by applying the RSS split statistic to each coordinate separately. Let the responses $Y_i = (Y_{i,1}, \ldots, Y_{i,q})^T$, where q denotes the multivariate outputs an $i = 1, \ldots, n$. The multivariate regression split rule can be written as

$$RSS = \sum_{k=1}^{q} \left\{ \sum_{i \in R_1(j,s)} \left(Y_{i,k} - \hat{Y}_{R_{1_k}}\right)^2 + \sum_{i \in R_2(j,s)} \left(Y_{i,k} - \hat{Y}_{R_{2_k}}\right)^2 \right\}. \tag{4.10}$$

The best split can be achieved by minimizing the RSS.

Equation 4.10 is a composite mean-squared error rule and does not consider any correlations among the multivariate outputs. Ishwaran et al. (2021) used the Mahalanobis distance to include the correlations. The Mahalanobis distance of a random element Z is defined as

$$D_M(Z) = (Z - \mu_Z)^T \sum_{Z}^{-1} (Z - \mu_Z), \tag{4.11}$$

Where μ_Z is the mean of$Z$.

Note that the covariance matrix can be singular when constructing a three where the number of observations decreases rapidly. Thus Ishwaran et al. (2021) proposed to solve this problem by employing the Moore-Penrose generalized inverse matrix, where for any matrix $A_{x\times p}$, the generalized inverse is the unique matrix $A^+_{p\times n}$. In non-singular A, then $A^+ = A^{-1}$.

The generalized inverse matrix can be solved using singular value decomposition (SVD). The SVD of $A_{n\times p}$ is $A = UDV^T$, where $U_{n\times p}$ is an orthonormal matrix, $V_{p\times p}$ is an orthogonal matrix, and $D_{p\times p}$ is the diagonal matrix. Then, the generalized inverse for A, given $n \geq p$, can be written as

$$A^+_{p\times n} = VD^+U^T. \tag{4.12}$$

Using the generalized inverse matrix, an efficient multivariate splitting rule based on Mahalanobis distance can be defined as

$$D_{M,j}(R_1, R_2) = \frac{n_1}{n}\sum_{i\in R_1}(Y_i - \hat{Y}_{R_1})^T(Q^*_j)^+ + (Y_i - \hat{Y}_{R_1})^T + \frac{n_2}{n}\sum_{i\in R_2}(Y_i - \hat{Y}_{R_2})^T(Q^*_j)^+ + (Y_i - \hat{Y}_{R_2})^T. \tag{4.13}$$

The best split j can be obtained by minimizing $D_{M,j}(R_1, R_2)$.

4.2.4 Modeling Consideration

The goal of predictive modeling, whether for classification or regression problems, is to build a model with the best possible prediction and generalization performance. This model can produce accurate predictions not only for the training data but also for new data. A generalizable model can be achieved by preventing the model from being underfitting or overfitting. Figure 4.3 illustrates the concept of model underfitting and overfitting. Underfitting is when a model fails to fit the training data points sufficiently, producing low prediction performance even for the training data. The training error reflects the bias of the model. High training error means the model is highly biased. In contrast, overfitting is when a model can perfectly fit the training data points but fails to capture the general pattern of the data, causing low prediction performance in new data (i.e., testing data). The test error reflects the variance of the model. Poor testing performance indicates a high variance in the model.

The concept of bias-variance trade-off is critical in developing a machine learning model. The best model is typically achieved by finding an optimal point between the two extreme conditions. This can be done by developing a model using a proper machine learning workflow (Figure 4.4), which includes a k-fold cross-validation (CV) (Figure 4.5). Furthermore, the risk of overfitting can also be reduced by selecting a machine learning algorithm that is flexible but less prone to overfitting, such as the ensemble models.

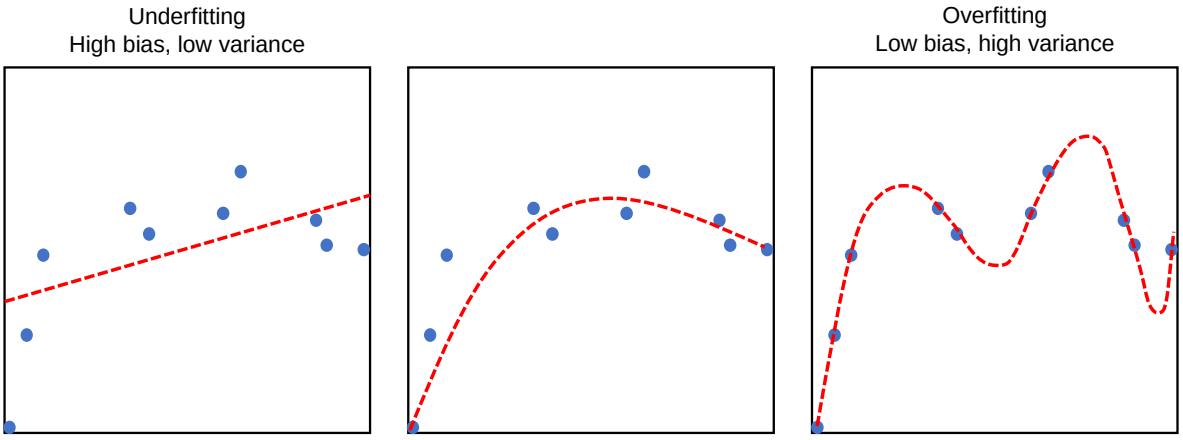

Figure 4.3: Concept of model underfitting and overfitting.

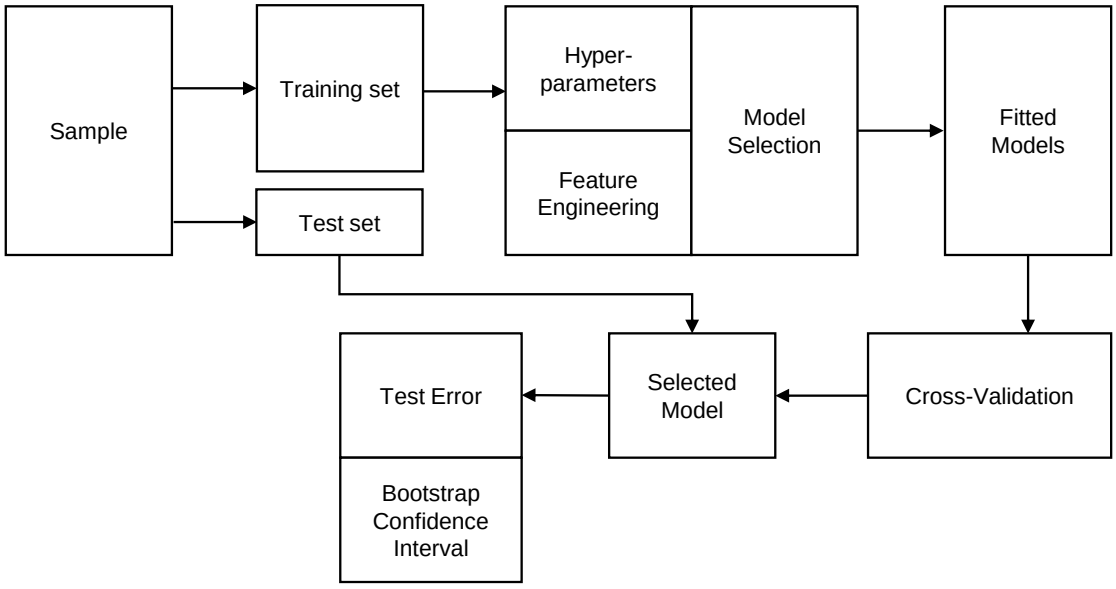

Figure 4.4: Example of robust machine learning workflow.

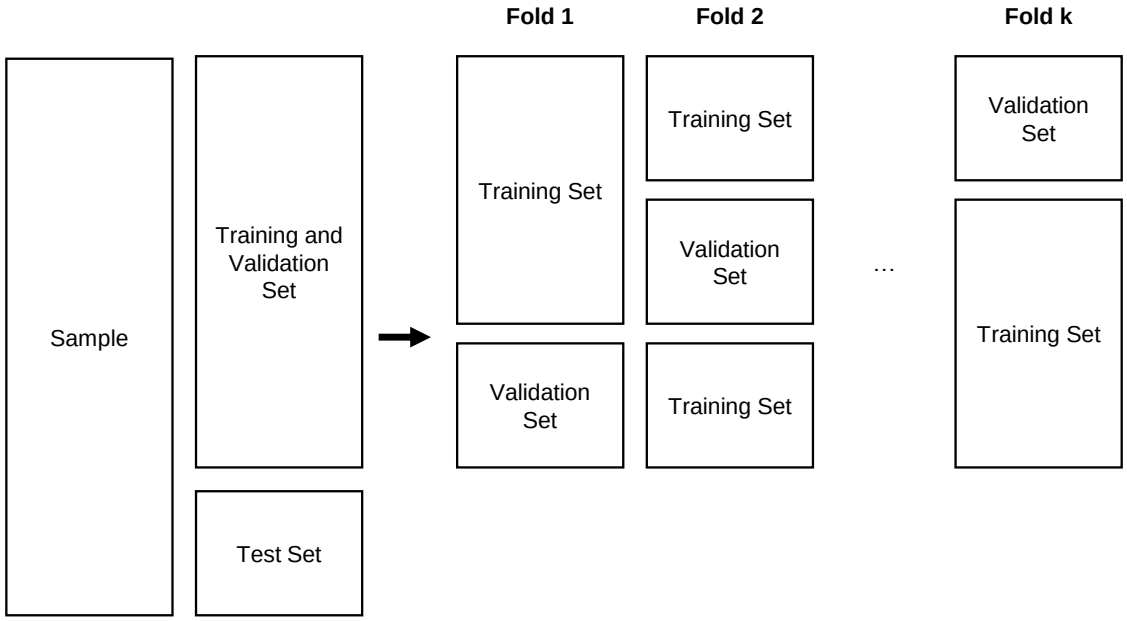

Figure 4.5: Data splitting with cross-validation.

4.3 Machine Learning for EPB TBM Tunneling Applications

Many studies have been conducted to leverage machine learning for tunneling applications. Most of the studies used machine learning to build data-driven predictive models for informing tunneling operations. This includes interpreting the geologic conditions, estimating the induced ground movements, and modeling the tunneling excavation process. These implementations are briefly discussed as follows. More discussions on these past developments can be found in the corresponding chapters of this dissertation.

4.3.1 Geologic Interpretation

The implementation of machine learning for developing geologic interpretation systems in TBM tunneling has emerged in the past few years. Various machine learning algorithms have been used as prediction models and many comparative studies have been conducted. Q. Zhang et al. (2019) proposed a method to transform TBM operation data to rock mass types using three machine learning algorithms, i.e., SVM, Random Forest, and kNN. Zhao et al. (2019) evaluated the performance of eight machine learning algorithms to classify geologic types, i.e., kNN, Bayesian Linear Regression (BLR), SVM, Decision Tree, Random Forest,

CatBoost, XGBoost, and artificial neural network (ANN). Recently, the development has been extended to aim for a real-time geologic interpretation system. Erharter et al. (2020) proposed a real-time (or online) rock classification system during tunneling using an ensemble Long Short-Term Memory (LSTM) model. H. Yu (2021) investigated methods to characterize the encountered ground condition based on EPB TBM data during tunneling by employing two machine learning frameworks, i.e., (i) supervised learning using multinomial logistic regression and (ii) semi-supervised learning using a similarity matrix and label propagation method.

These studies have shown the possibility of systematically and quantitatively interpreting the geologic conditions based on TBM data in real-time during tunneling. Machine learning algorithms have been shown to be capable of finding patterns in the input features (i.e., the TBM operation data) that contain the geologic information.

4.3.2 Ground Movement Estimation

Shi et al. (1998) conducted the early implementation of machine learning to estimate tunneling-induced ground movements. They implemented ANN to estimate the tunneling-induced ground movements of the Brasilia Tunnel project in Brazil. Subsequently, Suwansawat and Einstein (2006) examined the performance of various ANN architectures in estimating the maximum surface settlements of the Bangkok MRTA project in Thailand. Boubou et al. (2010) proposed a method to estimate the cross-section profile of tunneling-induced ground movements of the subway line B tunnel project in Toulouse, France.

Different machine learning algorithms have been used as prediction models in the past decade. For example, Bouayad and Emeriault (2017) implemented the adaptive neuro-fuzzy-based inference system algorithm (ANFIS), and Goh et al. (2018) implemented multivariate adaptive regression splines (MARS). Some studies have been conducted to compare and evaluate the performance of different machine learning algorithms in estimating tunneling-induced ground movements. These studies include Mahmoodzadeh et al. (2020), Ocak and Seker (2013), Tang and Na (2021), and W. G. Zhang et al. (2021). These studies involved various machine learning algorithms, from regression-based and decision trees-based algorithms to neural network-based algorithms.

The previously discussed studies focused mainly on connecting the maximum surface settlements and the TBM behaviors. F. Wang et al. (2013) reported an early development of the study that connected TBM operation parameters to the longitudinal ground movements. They used a wavelet smooth relevance vector machine algorithm to model longitudinal ground movement progression during EPB TBM tunneling. P. Zhang, Wu, Chen, Dai, et al. (2020) investigated interactions between a shield TBM and the ground in cross-sections and longitudinal sections. They used two machine learning algorithms to develop the prediction models, i.e., Random Forest and LSTM. The performance of the two models was compared using a data set from the Changsha metro project.

This review demonstrates the opportunity of utilizing data-driven methods for estimating tunneling-induced ground movements in both the cross and longitudinal sections. Machine

learning algorithms have been shown to be capable of finding complex interactions between the input features (e.g., the TBM operation) and the response (i.e., the estimated ground movement) without predefined assumptions. This is important in estimating tunneling-induced ground movements since they are affected by numerous factors and may not always follow the typical Gaussian shape assumption.

4.3.3 Tunneling Process Modeling

In most previous studies, tunneling processes were modeled separately as an individual process, for example, to model the earth pressure balancing process. In this case, machine learning was used as a prediction model embedded in an intelligent control system. Machine learning algorithms such as the Least Squares-SVM (LS-SVM) have been used to develop a model predictive control strategy for EPB TBM (X. Liu et al., 2011; X. Liu & Zhang, 2019). The predictive model was used to estimate earth pressures in the excavation chamber based on the screw conveyor rotation speed, advance rate, cutter rotation speed, and thrust force. Some studies used neural network-based algorithms to develop the EPB control system. Gao et al. (2020) used a neural network-based algorithm with a Gated Recurrent Unit (GRU) as the predictive model. X. Liu et al. (2020) used Backpropagation Neural Networks (BPPN) in dual heuristic programming (DHP) as the prediction and optimization model.

These studies show that implementing machine learning algorithms to build a predictive model for tunneling processes has become more prevalent. Machine learning enables flexible modeling without any predefined assumption. This can be beneficial for tunneling where the processes are governed by complex causal effects interactions of ground, machine, and human decisions.

Chapter 5

Effects of Data Aggregation on Prediction Models

5.1 Introduction

5.1.1 Background

Tunnel boring machine (TBM) operation data have been utilized to build various data-driven models, such as for classifying the geologic conditions (Cao et al., 2021; Erharter et al., 2020; Sousa & Einstein, 2012; Q. Zhang et al., 2019; Zhao et al., 2019), estimating the tunneling-induced ground movements (R.-P. Chen et al., 2019; R. Chen et al., 2019; Kim et al., 2022; Suwansawat & Einstein, 2006; W. G. Zhang et al., 2021), and predicting the tunneling performance (e.g., advance rate) (Mokhtari & Mooney, 2020). The models have also been developed using various machine learning algorithms, for example, linear regression-based methods (Mokhtari et al., 2020), support vector regression methods (Mokhtari & Mooney, 2020; Zhou et al., 2021), ensemble decision tree-based methods (Kong et al., 2022; P. Zhang et al., 2019), and the artificial neural network-based methods (Boubou et al., 2010; Erharter et al., 2020; Suwansawat & Einstein, 2006).

Despite this development, less effort has been given to understanding the characteristics of TBM data and the effects of data preparation on the models. There is still no agreement on how TBM data should be prepared (Marcher et al., 2020; Sheil, Suryasentana, Mooney, et al., 2020). Every proposed model was built using different data aggregation levels. For example, to develop geologic prediction models, Sousa and Einstein (2012) used ring aggregate data points, Q. Zhang et al. (2019) compressed the observation data points using the BIRCH algorithm (T. Zhang et al., 1996), Erharter and Marcher (2020) segmented the observation data points according to the linear trends using sliding windows, Cao et al. (2021) aggregated the observation data points to a period of 3-min intervals and then applied a median filter to smooth the time series. Another example, to develop TBM parameters prediction models, Mokhtari and Mooney (2019) aggregated the 5-sec observation data points resolution in one cutter head rotation, and Xu et al. (2021) smoothed 1-sec observation data points using a

denoising algorithm.

The above examples demonstrate that TBM data-driven models have been developed using different data aggregation levels. Figure 5.1 shows an example of a TBM variable feature, i.e., the cutter torque, over a segment of chainages from a tunneling data set. The original observation data points were produced every 5 seconds during the tunneling. In this figure, the data points were plotted in various data aggregation levels, following different data preparation methods from several previous studies, i.e., one cutter rotation, 20 mm, 3 min, and ring length aggregates. The figure shows that each aggregation level produces different longitudinal patterns. This may lead to some questions, how this difference alters the statistical characteristics of the data? Would this affect the prediction performance and behaviors of the developed data-driven models?

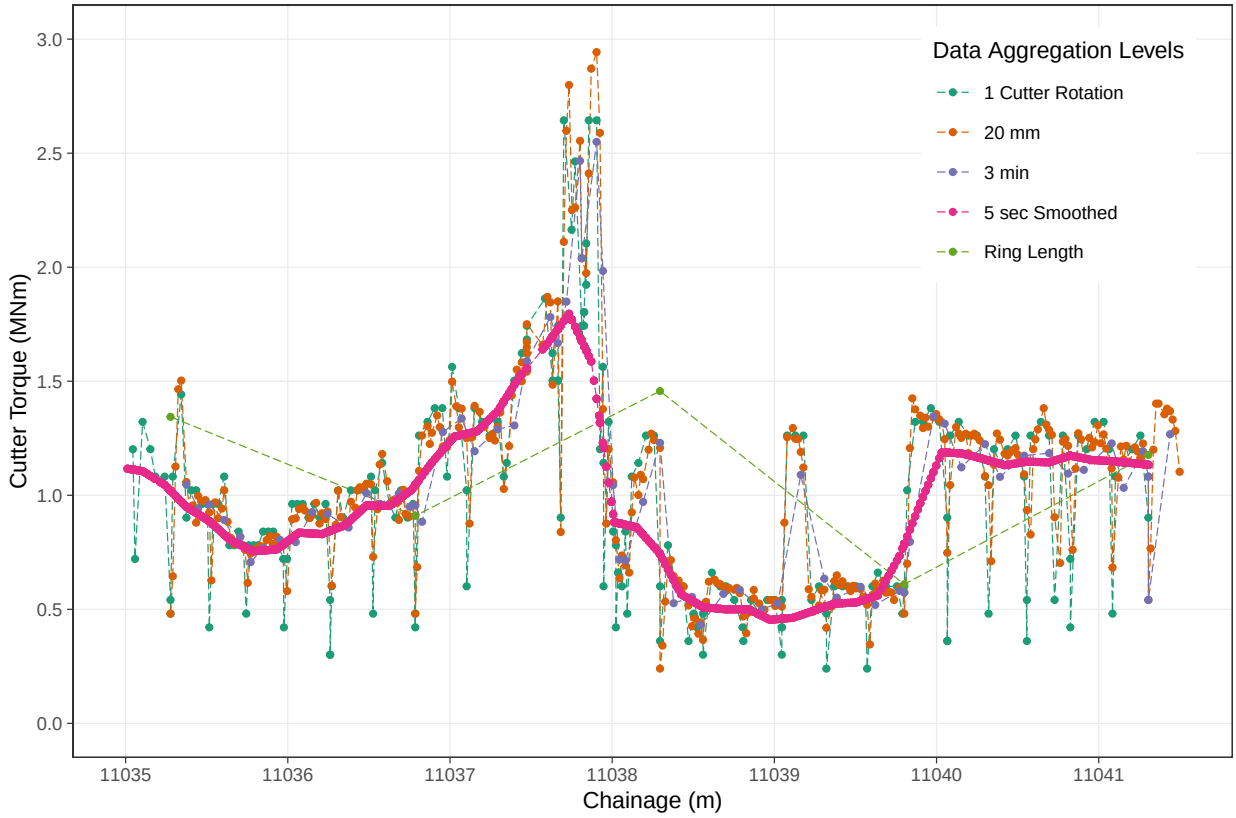

Figure 5.1: Example of a feature of TBM operation data plotted in different data aggregation levels, following different data preparation methods from several previous studies.

5.1.2 Related Works

Data aggregation has been studied over the decades in various problem domains using various data sets. Early studies emerged in the domain of statistics under the terminology of ecological correlations and regressions. This term describes the correlation or regression between variables as a group of individual data. Using the US census data, Robinson (1950) demonstrated that the correlation for the same two variables could be different at the individual and the group levels. Using the LA household data, Clark and Avery (1976) studied that the correlation between the independent and dependent variables changed with increasing aggregation levels and can be related to the changes in the regression coefficients. Amrhein (1995) performed controlled numerical simulations to explore the effects of aggregation in areal unit problems and concluded that means and variances were resistant to aggregation effects. However, dramatic effects were exhibited on the correlations and regression coefficients. Freedman (1999) highlighted that naïve inference from aggregated data may lead to ecological fallacies, which believe that relationships observed in groups are the same as in the individuals. These studies suggest that the effects of data aggregation are real, and inferences from aggregated data may produce valuable information but should be done cautiously due to the presence of aggregation bias and confounding problems (Freedman, 1999; Subramanian et al., 2009).

Studies on data aggregation have also emerged in engineering, especially in the application of wireless sensor networks (WSN). In this domain, data aggregation is critical to reducing the data size due to limited bandwidth and energy (Heidemann et al., 2001). It is also essential to increase the reliability of data due to erroneous individual sensor readings (Shrivastava et al., 2004). In the early development, Madden et al. (2002) implemented database-style aggregates (e.g., count, min, max, sum, average) to sensor readings flowing through sensor networks. Shrivastava et al. (2004) then extended the aggregation class queries into the quantiles (e.g., median). The queries were more challenging to be implemented in a streaming data flow, but they offered better approximation quality.

Various methods have been proposed to minimize the information loss due to aggregation in WSN data: e.g., Z. Ye et al. (2009) proposed optimal stochastic policies for distributed data aggregation in WSN. The optimal aggregation method became critical due to the presence of aggregation delay. Jiang et al. (2010) proposed a parameter-based data aggregation method using the expectation-minimization (EM) algorithm to extract statistical information from the data. Khedo et al. (2010) proposed an aggregation method by eliminating data redundancy, whereas S. Li et al. (2013) implemented the compressed sensing (CS) method to compress and reconstruct the data. While data aggregation can be useful, these studies suggest that some information contained in the original data may be eliminated due to the aggregation.

More recently, data aggregation studies have been growing in machine learning. The challenge is to train machine learning models using aggregated data instead of individual data (S. Chen et al., 2009; Musicant et al., 2007). This is motivated by, for example, limitations of class labels, communication, storage, and privacy preservation (Bhowmik,

2018; B.-C. Chen et al., 2006). S. Chen et al. (2009) proposed a K-means-based framework for classification problems with aggregate outputs as the training labels. In contrast, Y. Zhang et al. (2020) extended the multiple instances learning to multiple classifications and regression problems and presented a probabilistic framework that accommodates variation in aggregate observations. Recently, Kammerer et al. (2020) showed that reducing training data by aggregating through k-means and random sampling might produce a minimal loss in test accuracy. These studies indicate that the effects of data aggregation on machine learning prediction models are not clearly understood and are still an active research area. More studies on this problem domain are expected to be carried out due to the growing number of data sets, machine learning algorithms, and computing power.

5.1.3 Objectives

TBM data may have different characteristics from other data in the previous studies (in the other domains). The data are continuously generated during tunneling. The statistical characteristics of the data are also expected to change continuously, depending on complex causal effect interactions of the operator control decisions, the TBM behaviors, and the ground conditions. To date, no study has exclusively discussed the effects of data aggregation on TBM data sets. Therefore, the effects of TBM data aggregation on the prediction models have not been identified.

This study aims to investigate the effects of data aggregation on (i) the statistical characteristics of TBM operation data and (ii) the behaviors of the prediction models. The statistical characteristics of the data were represented in terms of central tendency, variation (or dispersion), and distribution. The effects of data aggregation on the models were investigated in the framework of the models' predictability, computability, and stability (B. Yu & Kumbier, 2020).

The advancement of data-driven modeling is key to developing more systematic and automated tunneling systems. Understanding the characteristics of the data is critical to understanding the results of machine learning predictions (Anik & Bunt, 2021; Westermann et al., 2021). This study is expected to provide a better understanding of the characteristic of tunneling operation data and the effects of data preparation (i.e., data aggregation) on TBM data-driven models, which can be a foundation to facilitate a unified method of TBM data preparation.

5.2 Data

5.2.1 Tunneling Cases

This study was performed using three earth pressure balance (EPB) TBM data sets from two tunneling cases: (i) the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA, and (ii) the North-South Line Phase-1 (NSLP1) Mass Rapid Transit (MRT) railway

twin tunnels in Jakarta, Indonesia. These data sets were selected to allow investigation on the model generalization since these two tunneling cases were excavated using different TBM sizes, controlled by different TBM operators and engineers, and carried out in significantly different geologic conditions. The SR99 is a highway tunnel with a length of 2830 m (1.756 mi) and a maximum depth of 65.5 m (215 ft) below the ground surface. The geologic conditions along the tunnel alignment are dominated by over-consolidated glacial and non-glacial pre-Vashon geologic units (WSDOT, 2010a, 2010b, 2010c). The tunnel was constructed using a 17.5 m (57.5 ft) diameter EPB TBM.

The NSLP1 is a double-track MRT railway system consisting of the Northbound (NB) and Southbound (SB) lines. This study uses the data from the underground sections Contract Package (CP) 104-105, where each line consists of four tunnels that connect four underground stations. The total length of each of the lines is 2612 m, and the maximum depth of the tunnels is 15.6 m below the ground surface. The geologic conditions along the tunnel alignments are dominated by alluvial deposits of bedded fine and sandy tuff, and Pleistocene marine and non-marine deposits to a significant depth (SOWJ-JV, 2013a, 2013b; Turkandi et al., 1992). The tunnels were constructed using two EPB TBMs with a diameter of 6.8 m.

5.2.2 Data Description and Preparation

The original EPB TBM data is a massive data set that consist of hundreds or even thousands variable features. The features are the tunneling operation measurement records generated by numerous sensors installed on the TBM. The SR99 original data set consists of more than 5000 features, while each NSLP1 original data set consists of more than 400 features. The investigation was performed on seven selected EPB TBM features, i.e., the advance rate, thrust force, cutter torque, foam volume, chamber pressure, screw rotation speed, and grout volume. These features were in various physical measurement units and chosen to represent key features for each EPB TBM process, based on JSCE (2016) (JSCE, 2016). More descriptions of the selected features are presented in Table 5.1.

This selection is essential to limit the complexity and enable more interpretability of the models. Note that the selection might exclude some other important features, such as the cutter rotation speed, screw pressure, and other types of ground conditioners (e.g., polymer, bentonite). It might also restrict the contained information in the prediction models. However, it can be considered adequate since the focus of this study is to investigate the effects of data aggregation, not to develop the most accurate model.

The TBM data is typically generated every few seconds. However, the observation data points in this study were represented in spatial series. This means an observation data point represents all measured values within one spatial unit. Observation data points for gauge metrics (e.g., features related to pressures, forces, torques, and speeds) were the average measured values within a spatial unit. Observation data points for cumulative increment metrics (e.g., features related to volume) were the final measured value within a spatial unit. The effects of data aggregation were investigated by preparing each data set in two levels of spatial aggregation as the observation data points, i.e.,

(i) fine aggregate data, represented by 20 mm chainage-length data points, and

(ii) coarse aggregate data, represented by ring-length data points.

The fine aggregate data was used as the reference, while the coarse aggregate data was used to evaluate the aggregation effects. Note that the average ring length of the SR99 and the NSLP1 tunnels are approximately 2000 mm and 1500 mm, respectively. Figure 5.4 presents the longitudinal plot of the SR99 (a), NSLP1-SB (b), and NSLP1-NB (c). These figures demonstrate substantial differences between the two aggregation levels.

Following the preparation, the three data sets were cleaned according to the following steps.

(i) **Removal of non-excavation phase**. The investigation was performed on observation data points during the excavation phase. Thus, observation data points during the non-excavation phase, such as during segment erection, were not considered. This was done by removing observation data points with an advance rate of 0 mm/min and cutter torque of 0 MNm.

(ii) **Volume normalization**. Features of cumulative measurements, such as volumes, were normalized since there is some variation in ring length (i.e., thrust stroke).

(iii) **Removal of erroneous records**. The data were cleaned by removing erroneous records, such as duplicated chainages and observations with missing values. This reduces the amount of observation data points but avoids further complexities from imputing missing data with assumptions.

(iv) **Removal of outliers**. The data were also cleaned by removing the outliers. The outliers were defined as any values higher than five times the standard deviation of a feature. According to Chebyshev inequality, this ensures to remain at least 96% of the data (Seo, 2006).

The number of observation data points returned from these data preparation are summarized in Table 5.2. The longitudinal plot of the data are shown in Figure 5.2 to 5.4.

5.3 Methods

This study aims to investigate the effects of data aggregation on the statistical characteristics of EPB TBM data and the prediction models. This was achieved by conducting three main tasks, as listed below.

1. **Evaluating the dynamic behaviors of TBM data**. This task was performed to understand how the statistical parameters of TBM data evolve during tunneling. This was achieved by investigating the change of statistical parameters of the reference data sets (i.e., fine aggregate observation data points) along the chainages.

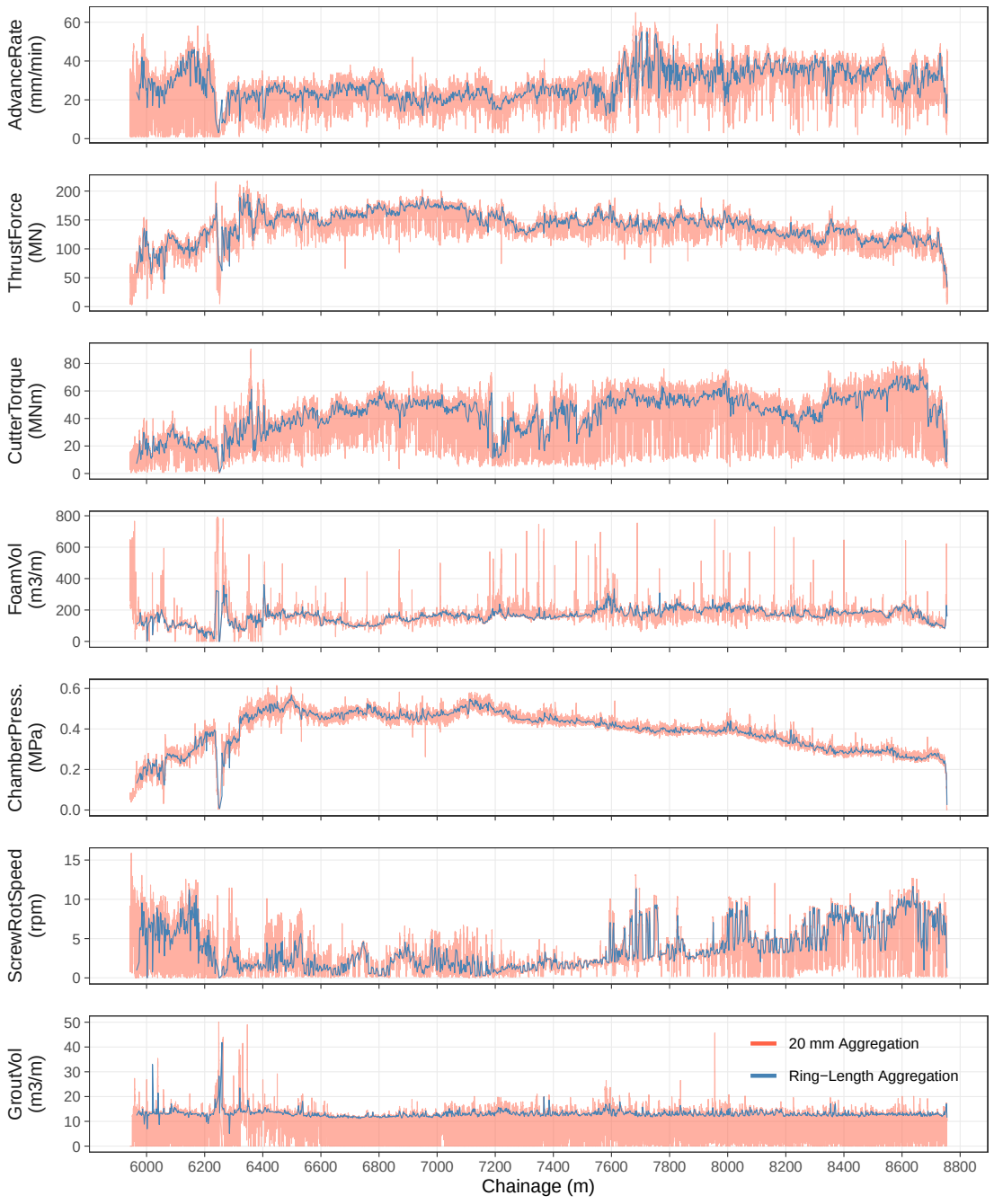

Figure 5.2: Longitudinal plot of the SR99 data sets in both fine and coarse data aggregation levels.

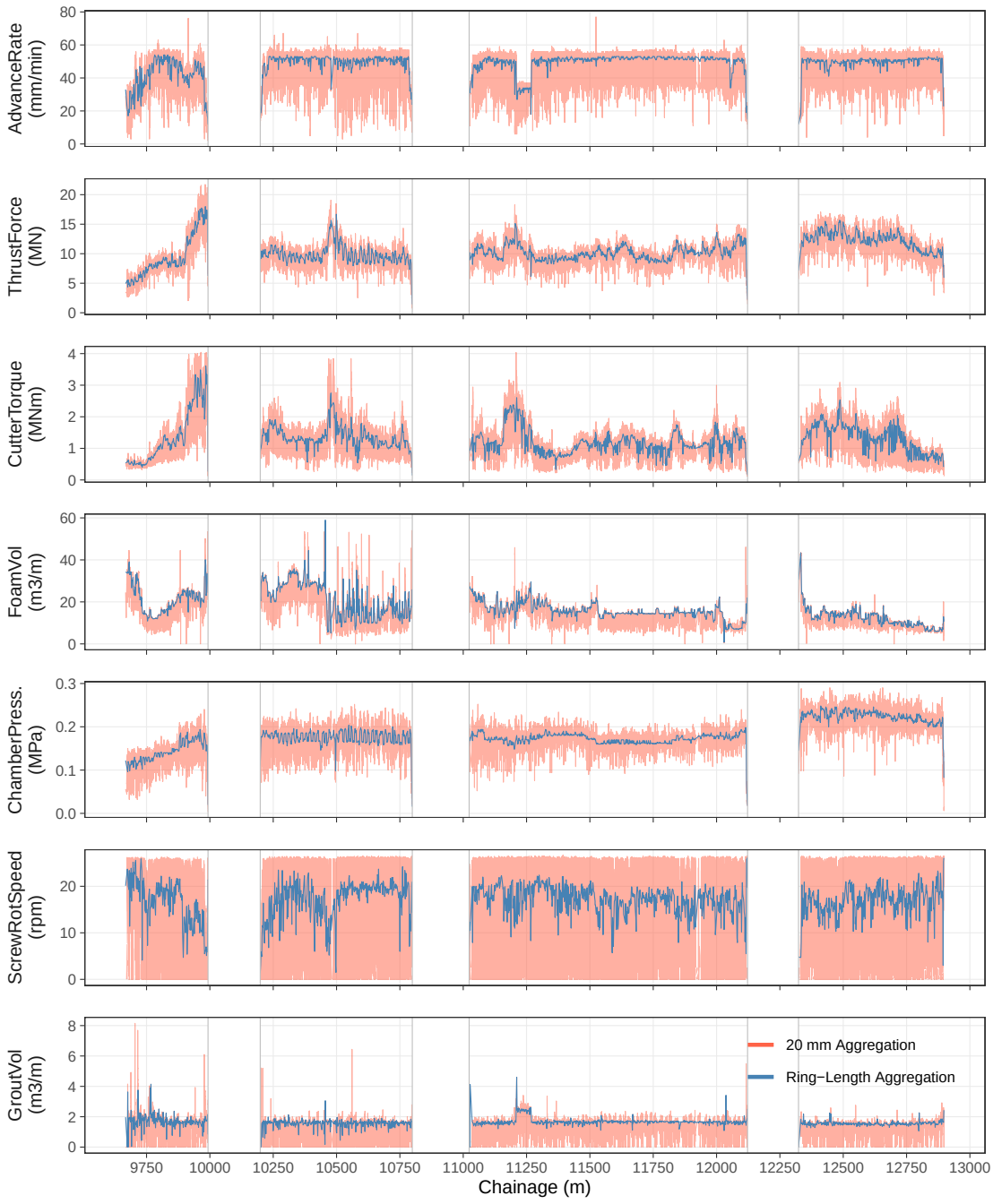

Figure 5.3: Longitudinal plot of the NSLP1-SB data sets in both fine and coarse data aggregation levels.

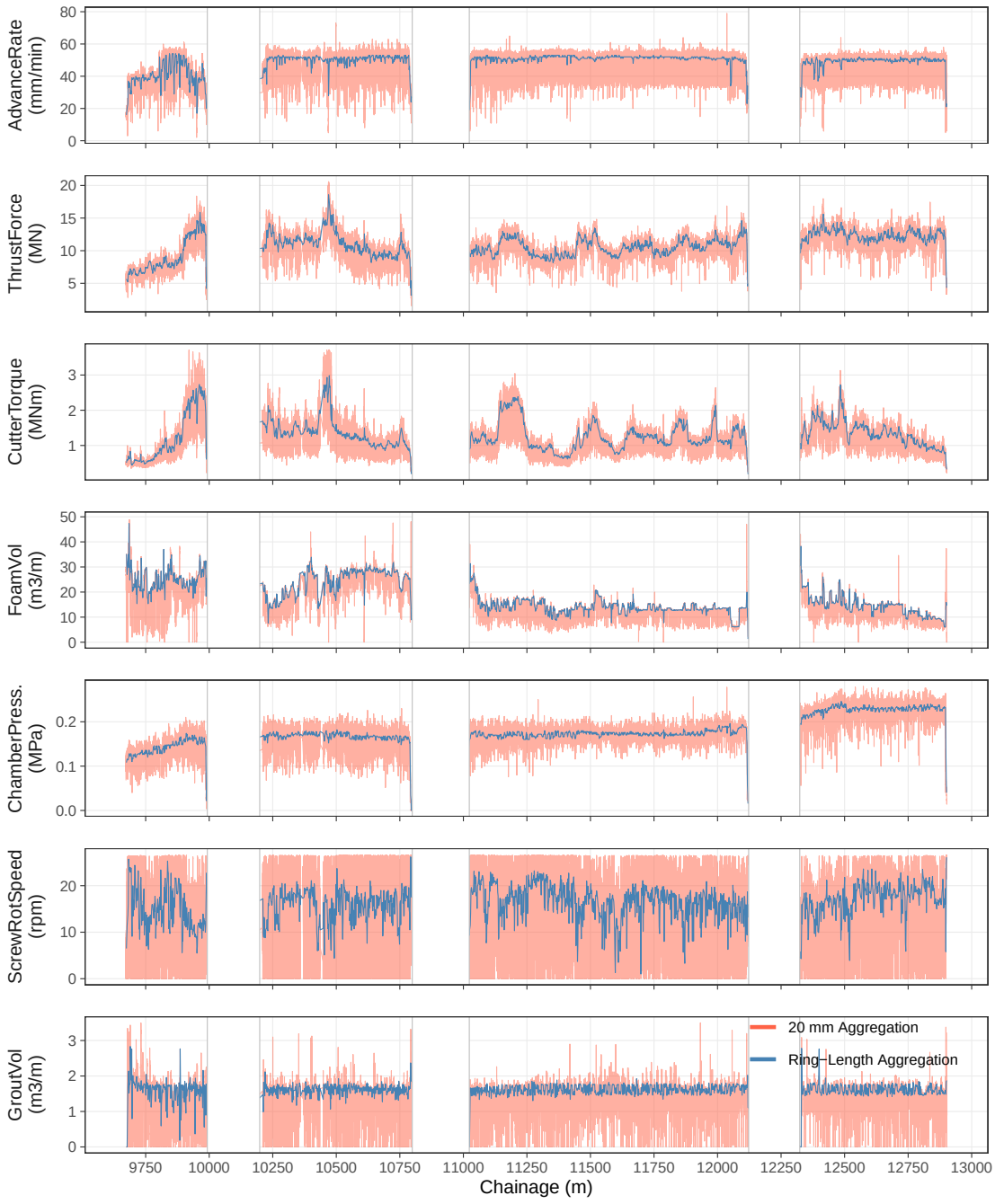

Figure 5.4: Longitudinal plot of the NSLP1-NB data sets in both fine and coarse data aggregation levels.

Table 5.1: List of EPB TBM features selected for the data aggregation analysis.

No.	Feature	Unit	Group	Remark
1	Advance rate	mm/min	Advancing	Average speed of cylindrical thrusts
2	Thrust force	MN	Advancing	Total force of cylindrical thrusts
3	Cutter torque	MNm	Excavation	Torque of cutter head
4	Foam volume	m^3/m	Ground conditioning	Total volume of injected foam normalized by the injection distance (i.e., thrust stroke)
5	Chamber pressure	MPa	Earth pressure balancing	Average pressure in the chamber
6	Screw rotation speed	rpm	Muck discharging	Average rotation speed of screw conveyor
7	Grout volume	m^3/m	Backfill grouting	Total volume of injected grout normalized by the injection distance (i.e., thrust stroke)

Table 5.2: Summary of the observation data points.

Aggregation Level	SR99	NSLP1-SB	BSLP1-NB
20 mm-length	140377	130499	132039
Ring-length*	1415	1737	744

*) slightly varies with approx. 1.5 m

2. **Investigating the effects of data aggregation on the statistical parameters**. This task was performed to understand how data aggregation changes the statistical parameters of TBM data. This was achieved by investigating the ratio of statistical parameters in different data aggregation levels along the tunnel chainages.

3. **Investigating the effects of data aggregation on the prediction models**. This task was performed to understand what aspects of prediction models are distorted by data aggregation. This was achieved by developing models using two different data aggregation levels (i.e., fine and coarse aggregates) and investigating the models' predictability, computability, and stability.

More descriptions of the methods can be found in the following sections.

5.3.1 Statistical Parameters

The statistical characteristics of the data were investigated in terms of (i) the mean, which represents data central tendency, (ii) the standard deviation, which represents the data variation (or dispersion), and (iii) Kolmogorov-Smirnov (KS) D-statistic, which represents the difference of the data distribution to a benchmark distribution, i.e., the normal distribution.

The arithmetic mean, μ, is defined as

$$\mu = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right), \tag{5.1}$$

where x_i is the value of x variable feature at an observation data point i, and n is the total number of the observation data points.

The standard deviation, σ, is defined as

$$\sigma = \sqrt{\frac{1}{n}(x_i - \mu)^2}. \tag{5.2}$$

The KS test is a nonparametric method to quantify the difference between two data distributions. The test measures differences between two cumulative frequency distributions and take the maximum absolute difference, D, defined as

$$D = \max|F_o(x) - F_{ref}(x)|, \tag{5.3}$$

where $F_o(x)$ is the observed cumulative frequency distribution of a variable and $F_{ref}(x)$ is the reference cumulative frequency distribution. In the KS test, a larger D indicates a larger difference between the two data distributions.

These statistical parameters were computed over the spatial sequence to investigate the change in the data characteristics during tunneling. This means each parameter was computed at every observation data point (chainage points), using all the previous observation data points from the beginning of the tunneling. This was applied to all the selected features and all the tunneling cases.

5.3.2 Aggregation Ratio

The effects of aggregation on the statistical parameters were represented by an aggregation ratio, R, defined as

$$R = \frac{\phi_{agg}}{\phi_{ref}} \tag{5.4}$$

where ϕ_{agg} and ϕ_{ref} are statistical parameters from the aggregated data (i.e., the coarse aggregate) and referenced data (i.e., the fine aggregate), respectively. An aggregation ratio of 1.0 indicates that the aggregation does not change the statistical parameter of the data, i.e., both aggregation levels exhibit an equivalent parameter. In contrast, high or low aggregation ratios indicate strong aggregation effects on the parameter. The aggregation ratio was also computed over the spatial sequence to investigate the change of the aggregation effects along the tunnel chainages.

5.3.3 Model Setup

The effects of data aggregation on prediction models were investigated by developing models using two different data aggregation levels, i.e., the fine aggregation (20 mm chainage-length data points) and coarse aggregation (ring-length data points). The schematic diagram of the prediction model is presented in Figure 5.5. This model considers a TBM feature as a response variable and the other six as predictor variables. To limit the complexity, this model does not consider correlations between observation data points. This means a response variable at an observation data point of i is predicted using predictor variables at the same observation data point of i, without considering other predictor variables from the previous observation data points (Zeger & Liang, 1992).

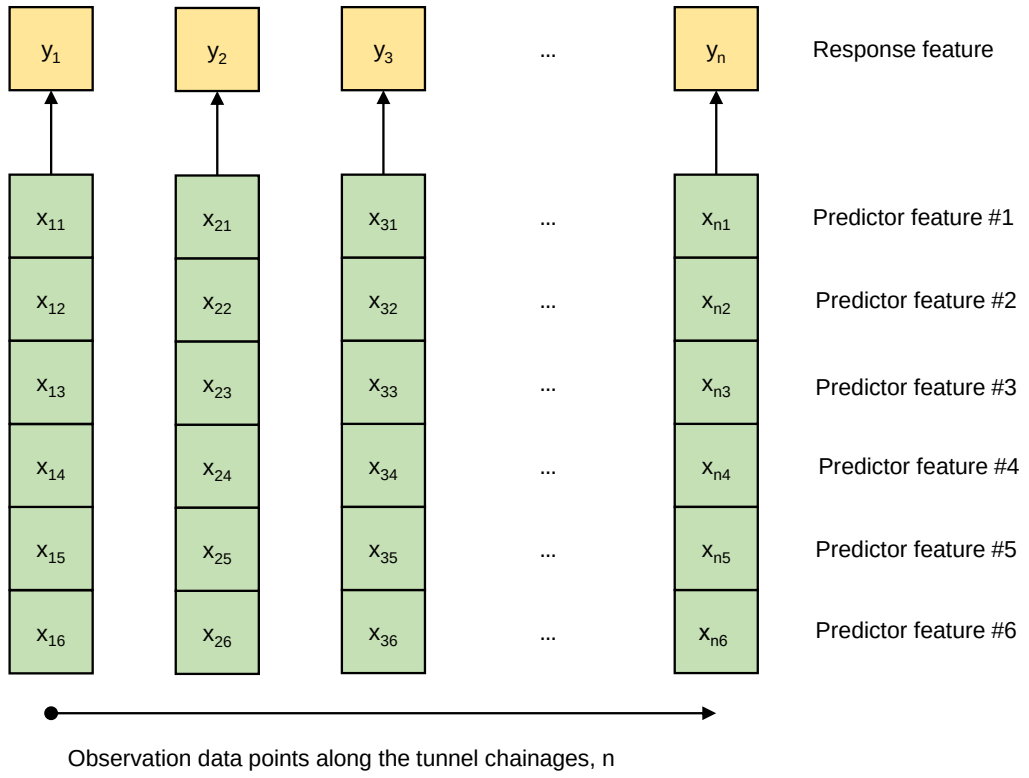

Figure 5.5: Schematic diagram of the prediction model.

The models were developed in the dynamic sequential learning scheme to represent the real tunneling sequence and data acquisition process. The schematic diagram of the dynamic sequential learning scheme is presented in Figure 5.6. In this scheme, the predictions are performed in sequential order at every observation data point (as testing data) based on

past observations (as training data). In other words, to predict a response variable at an observation data point, the prediction model is trained using all previous observations before that point. The prediction model is continuously retrained to predict the following observation data point. This means the model learns incrementally during the tunneling process.

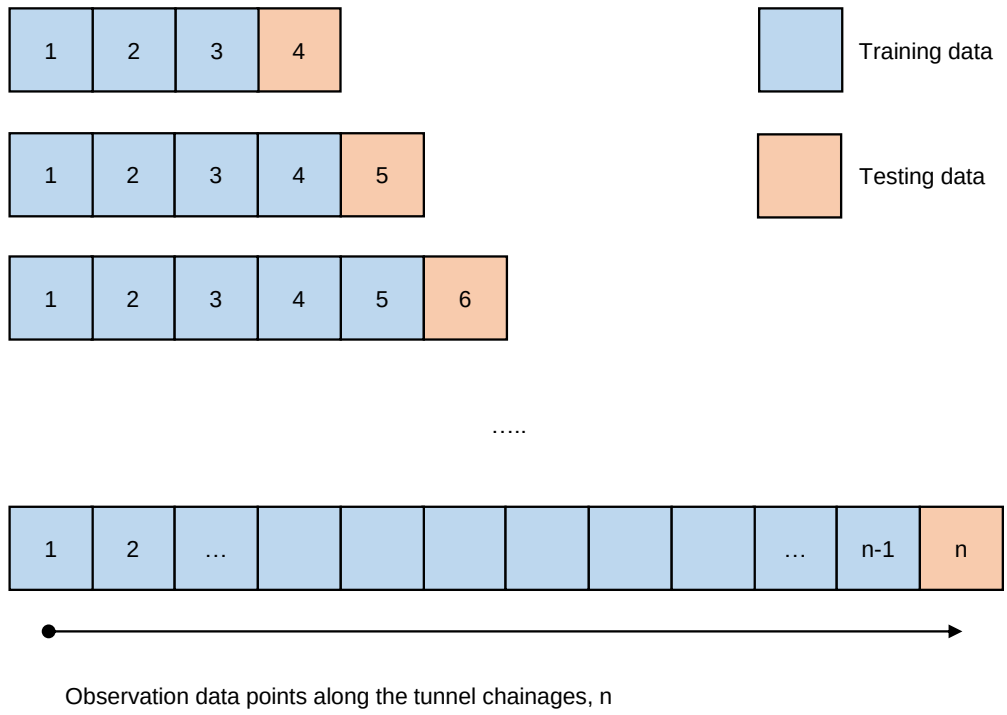

Figure 5.6: Schematic diagram of dynamic sequential data splitting for model training and testing.

5.3.4 Prediction Methods

To investigate the effects of data aggregation, the models were also developed using different prediction methods, i.e., the Ordinary Least Squares (OLS) and Random Forests (RF).

- OLS is a widely used linear regression method that uses regression coefficients to define a response variable, y_i, as a linear function of the predictor variables, $x_1, x_2, \ldots, x_p$.

The regression coefficients can be estimated by minimizing the residual sum of squares (RSS). This method was selected to represent parametric linear prediction methods.

- RF is an ensemble-supervised learning algorithm that aggregates a large number of decision trees to produce predictions (Breiman, 2001). RF was selected to represent nonparametric nonlinear machine learning methods. This study used the RF fast implementation in C++ and R (Wright & Ziegler, 2017).

RF was selected since this method has straightforward hyperparameter tuning and performs excellently in tabular data (Grinsztajn et al., 2022). The method is also robust to noises and less prone to overfit (Y. Liu et al., 2012). Furthermore, as a nonparametric method, RF does not require the data to meet certain assumptions or parameters (Malley et al., 2012). More discussion of these prediction methods and their mathematical bases can be found in Chapter 4.

5.3.5 Model Hyperparameters

RF regression models can be tuned by specifying the hyperparameters, such as, the number of trees (ntrees), the possible number of features that are randomly selected to split at each node (mtry), the minimum node size, and the splitting rule. However, various studies have shown that, as long as the number of trees is adequate, variation of the other hyperparameters will not significantly affect the RF model performance (Apoji et al., 2022b; Bernard et al., 2009; Probst et al., 2019). In this study, the RF hyperparameters were selected based on recommendations from previous studies (Apoji et al., 2022b; Probst et al., 2019), i.e., ntrees = 500, mtry = p/3, min node size = 5.

5.3.6 Model Evaluation and Feature Importance

This study used the normalized mean absolute error (NMAE) as the performance metric. This metric was selected to accommodate comparisons of errors across features with different units. The mean absolute error (MAE) is defined as

$$MAE = \frac{\Sigma_i^n (y_i - \hat{y}_i)^2}{n}. \tag{5.5}$$

Then, the NMAE can be obtained, which is defined as

$$NMAE(y, \hat{y}) = \frac{MAE(y, \hat{y})}{\frac{1}{n}\Sigma_i^n |y_i|} \tag{5.6}$$

The stability of the prediction models was represented as the sensitivity of the feature importance ranks under various conditions. This means a model can be considered stable (or not sensitive to data aggregation) if it has comparable feature importance ranks in various model settings and input, i.e., different prediction methods, data aggregation levels, and tunneling cases.

- The feature importance rank of OLS models was obtained from the absolute value of OLS coefficients. The higher the OLS coefficients, the stronger that feature as a predictor. Note that the predictor variables were in different scales. Therefore, these predictor variables were standardized to allow comparisons of OLS coefficients.

- The feature importance rank of RF models was obtained from the permutation importance scores. The permutation-based feature importance analysis measures the difference between the prediction accuracy of the model developed using the original predictors and the prediction accuracy of the model developed using a predictor being permuted (Gregorutti et al., 2017; Nicodemus et al., 2010). Higher permutation importance scores of a feature mean higher prediction errors when it is permuted, indicating the importance of the feature as a predictor in the model.

5.4 Results and Discussion

5.4.1 Statistical Parameter Changes Over Dynamic Sequence

Figure 5.7 presents statistical parameters of all the considered EPB TBM features from the fine aggregate data sets. Since the features are in different units, these parameters were normalized to their end values for the presentation purpose, e.g., $\frac{\mu_i}{\mu_n}$, where μ_i is the mean of a feature until a particular chainage location i, and μ_n is the mean of the feature until the end of the tunneling. The figure shows that every tunneling case produced different patterns, even for the NSLP1 twin tunnels that were carried out in the same geologic conditions. This demonstrates the complexity and uniqueness of every TBM data.

The figure shows that the mean and the standard deviation of all the features and from all the tunneling cases were unstable during the early part of the tunnel chainages when the observation data points were limited. These parameters became more stable in larger numbers of observation data points since a single new observation had a smaller impact on shifting the parameters. This is also related to the Law of Large Numbers (LLN), which postulates that the distribution of an independent and identically distributed random variable will converge to the mean of the population as the sample size increases.

The KS D parameter in this figure represents the similarity between the distribution of a variable feature and the Gaussian (normal) distribution. The lower the D parameter, the closer the data distribution of that feature is to the normal distribution. Similar to the other parameters, the D parameter fluctuated during the early part of the tunnel chainages and then stabilized in larger numbers of observation data points. Overall, the values were above 0 but below 0.3. This means no features had precisely the same distribution as the Gaussian distribution. Nevertheless, approximating the distributions with the Gaussian distribution may still be reasonable, especially for features with a D parameter close to 0.

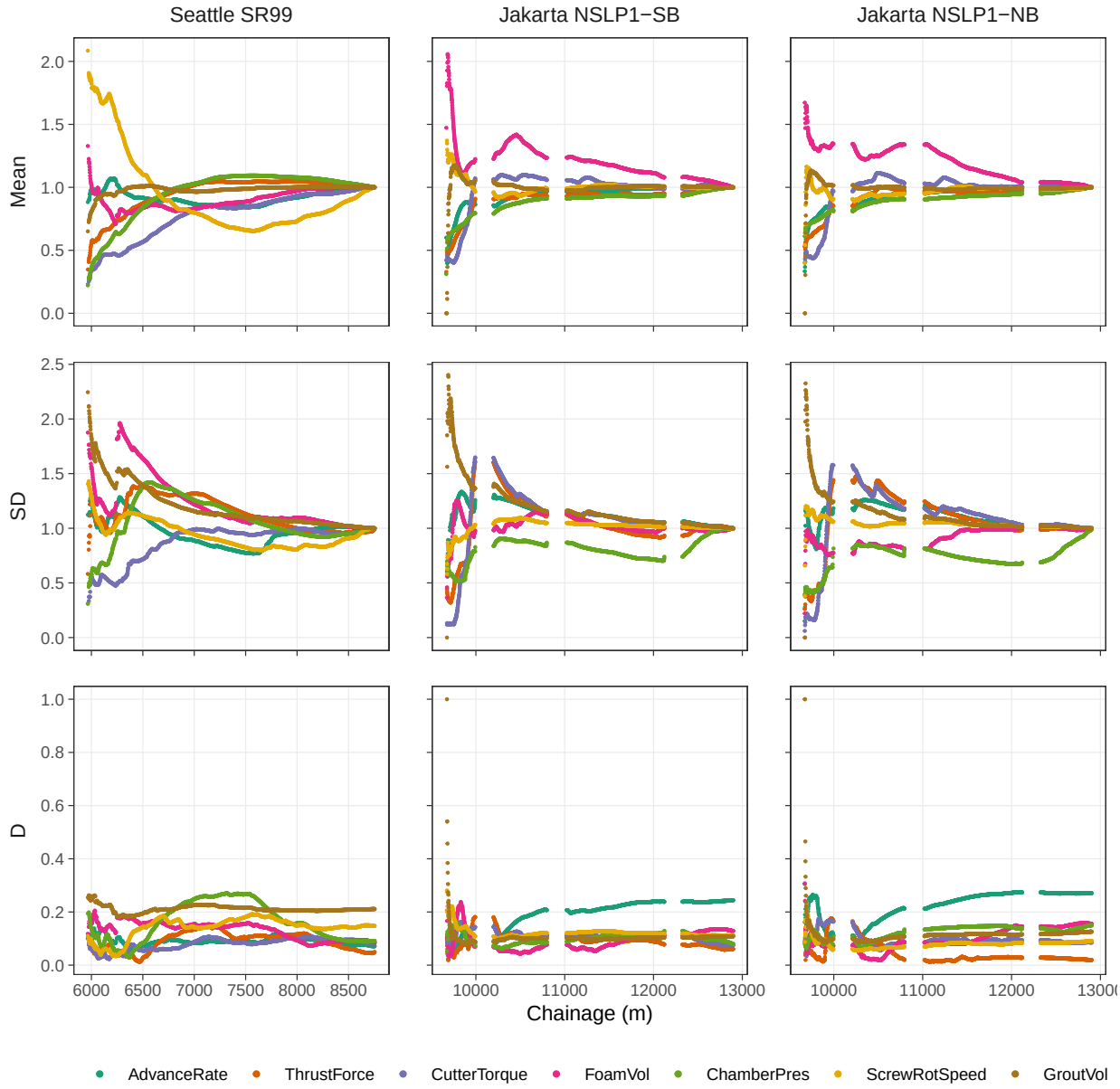

Figure 5.7: Statistical parameters of the considered EPB TBM features from all the tunneling cases.

5.4.2 Effects on Statistical Parameters

Figure 5.8 presents aggregation ratios of the considered EPB TBM features from all the tunneling cases. The figure shows that data aggregation levels did not affect the central tendency of most variable features. The aggregation ratio of the means jittered at the early chainages but converged and stabilized at approximately 1.0, as expected mathematically. In contrast, the data aggregation levels affected the data variation. The aggregation ratio of the standard deviations also jittered at the early chainages.

However, many features converged and stabilized at values lower than 1.0. This means the coarse aggregate data have a lower variance than the fine aggregate data. This demonstrates that TBM data aggregation is essentially the same as data smoothing (or denoising), which attempts to reduce the data variance while preserving the central tendency.

The aggregation ratio of the KS D parameter shows different behaviors. The aggregation ratio continuously fluctuated from the beginning until the end of the tunnel chainages. Furthermore, the effects were more substantial. In many features, the ratios were in the range of 0.5 to more than 2.0. This suggests that the effects of aggregation on TBM data distribution can be more sensitive and less predictable than the other statistical parameters. This can be an important consideration when developing a TBM data-driven model using a prediction method that works under certain data distribution assumptions.

5.4.3 Effects on Model Predictability

Figure 5.9 presents the prediction performance comparison between models trained using the coarse and fine aggregate data. The figure also shows the performance comparison between the OLS and the RF prediction methods. Three features were selected as the response features, i.e., the cutter torque (left panels), foam volume (center panels), and advance rate (right panels). The figure shows that both data aggregation levels and prediction methods produced high errors and intense jitters during the early part of the chainages. This low prediction performance was likely due to limited training data since the errors dropped and stabilized in the later chainages when more training data were available. This indicates that there may be a threshold level where the training data can be considered sufficient to produce stable and reliable predictions.

The figure also shows that the RF models consistently outperform the OLS models at the same data aggregation level. This indicates the presence of complex and nonlinear relationships in TBM features, which are the product of interactions among the operator control decisions, TBM behaviors, and geologic conditions. The OLS rigidly fitted the feature relationships to be linear. In contrast, the RF method flexibly facilitated nonlinear relationships among the features, resulting in better prediction performance. This demonstrates the value of using nonlinear nonparametric machine learning methods to develop TBM prediction models.

Furthermore, the figure shows that RF models with coarse aggregate data produced better prediction performance than those with fine aggregate data. This is due to the lower

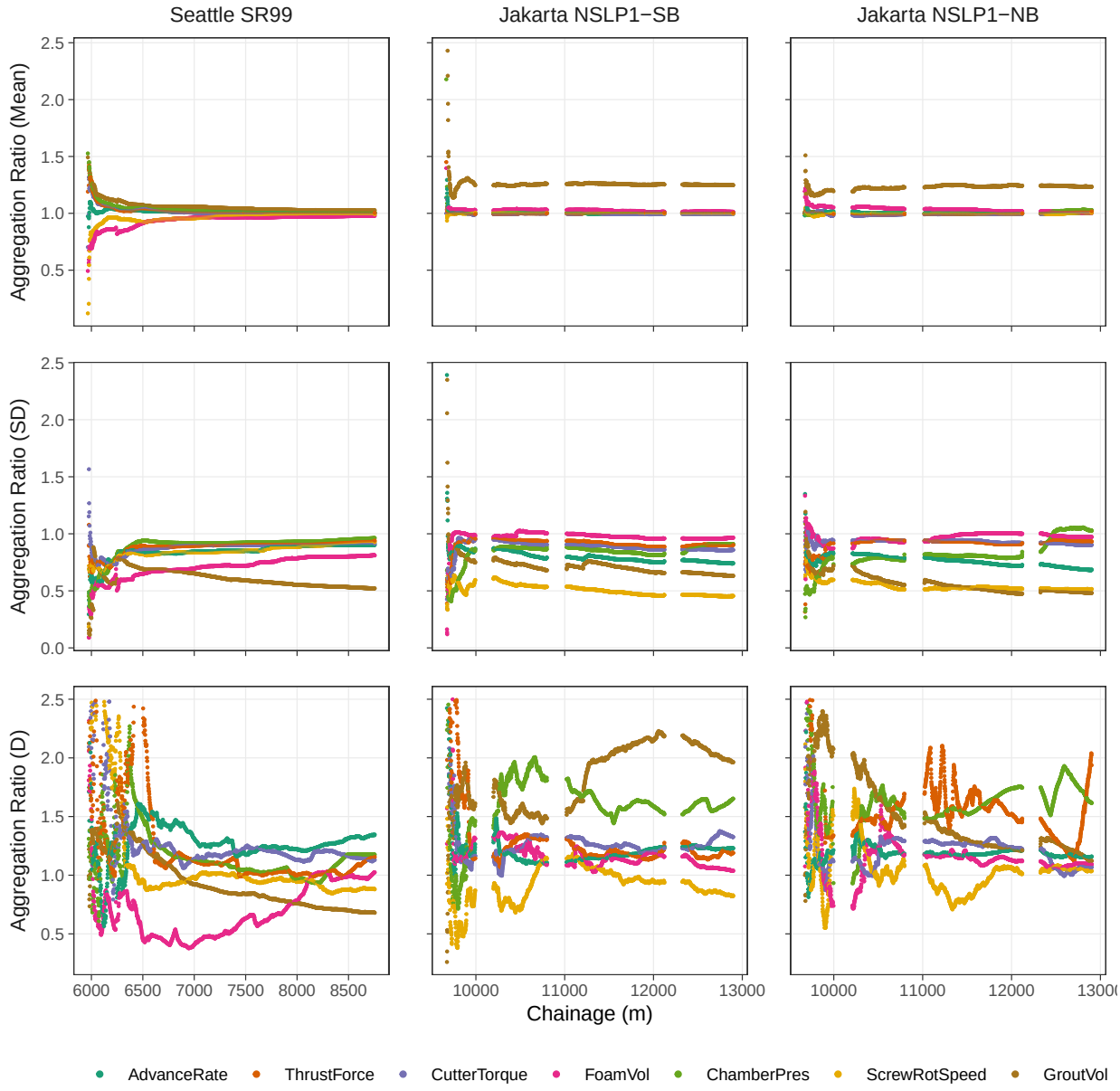

Figure 5.8: Aggregation ratios of the considered EPB TBM features from all the tunneling cases.

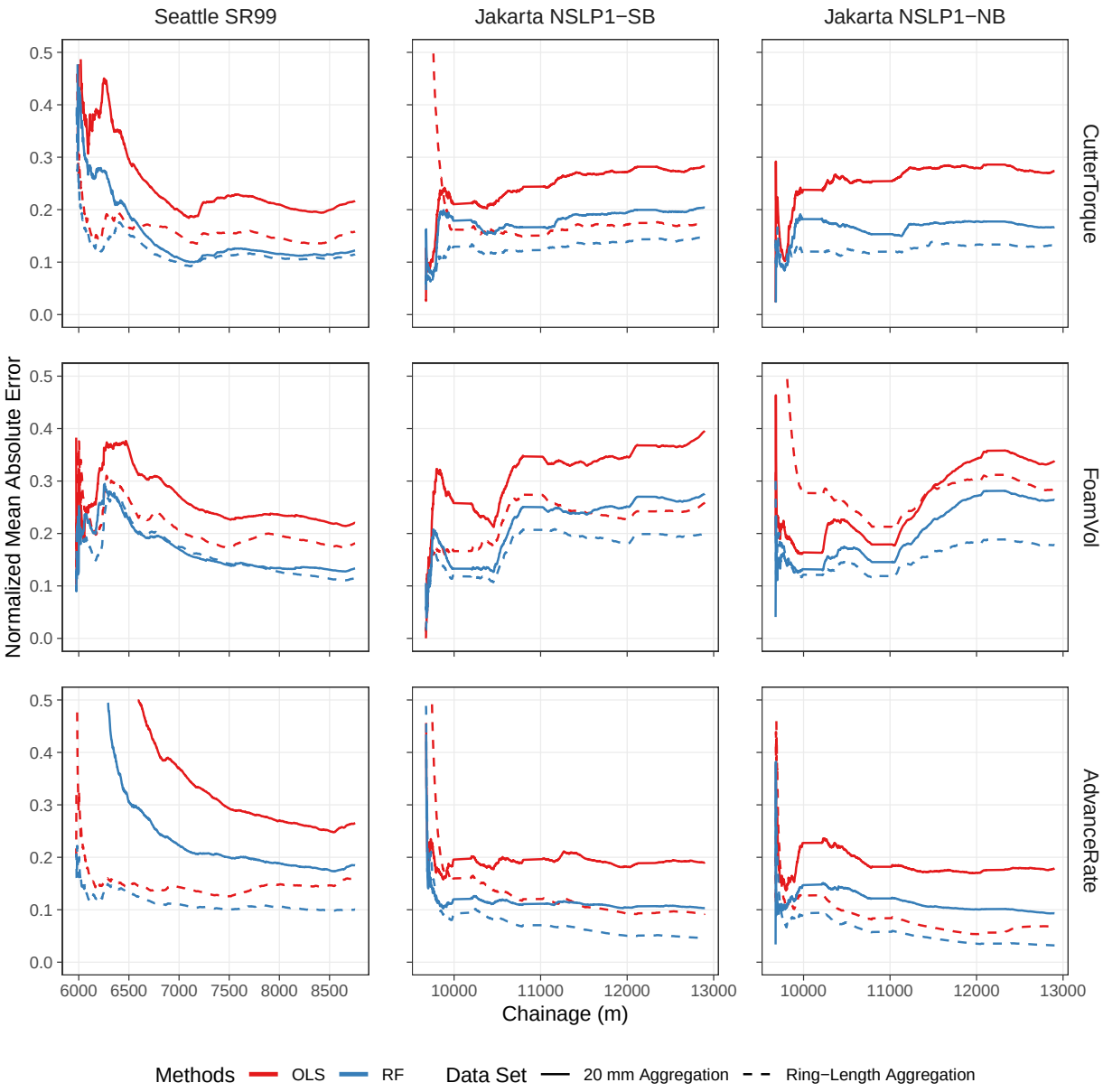

Figure 5.9: Comparison of prediction performance between prediction models trained using fine and coarse data aggregation levels.

variance of the coarse aggregate data (in both training and testing data). However, despite the difference, both data aggregation levels produced similar prediction behaviors (error patterns) along the chainage. This indicates that data aggregation does not substantially affect the behaviors of prediction models. However, it affects the prediction performance, where models with a coarser data aggregation level will gain higher prediction performance due to the lower variance. This suggests that different preparation methods (data aggregation or smoothing methods) may be used to obtain the general behaviors of prediction models, but the prediction performance comparison is only valid at the same data aggregation level.

Similarly, OLS models with coarse aggregate data produced better prediction performance than those with fine aggregate data. However, an exception should be made for cutter torque predictions of the NSLP1-NB data set. In this case, the OLS model with coarse data aggregation produced extremely high errors. This is likely caused by response outliers or high leverages in the prediction results, which are statistical phenomena of extremely high residual errors due to the presence of multicollinearity among the predictor variables (James et al., 2013). These problems can potentially occur in OLS-based regression methods. This shows that RF models produced better performance and more stable predictions than the OLS models, which suggests another advantage of using machine learning to develop EPB TBM prediction models.

5.4.4 Data Reduction and Information Loss

In this investigation, the prediction models were trained using five data reduction scenarios, i.e., the coarse aggregate data (approx. 1% of the fine aggregate data) and random sampling of the fine aggregate data with the percentage of 1%, 5%, 10%, and 30%. The prediction model was also trained using 100% of the fine aggregate data to provide a benchmark performance. Note that different from the previous investigation, all the trained models here were tested using the fine aggregate data to measure the effect of data reduction.

Figure 5.10 presents the effects of data reduction on the performance of the SR99 cutter torque prediction models. The left and right panels show the OLS and RF prediction results, respectively. Both figures show that prediction models trained using the coarse aggregate data produced the lowest prediction performance (the highest errors), even compared to models trained using the 1% random sample of the fine aggregate data, which has a similar number of observation data points. This demonstrates the limitation of prediction models trained using coarse aggregate data. As discussed previously, data aggregation reduces the variance of the features. Thus, the model could not capture the variance of fine aggregate data in the testing set. This suggests that a prediction model should be trained using a data set with the same aggregation level as the future (testing or prediction) data. The model may suffer substantial reductions in prediction performance if trained using the aggregated (or smoothed) data set.

Furthermore, Figure 5.10 also presents the effects of information loss on prediction performance. The figure shows that the performance of the RF models increases with the increase in sampling percentage. The RF model trained using only 30% of the fine aggregate data

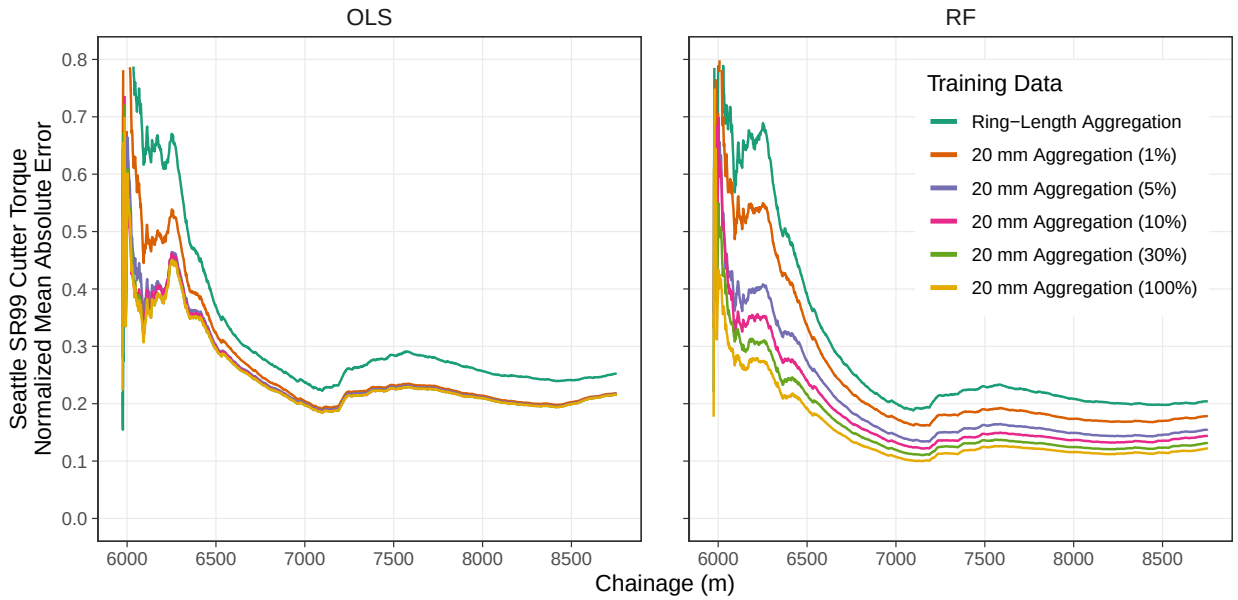

Figure 5.10: Effects of data reduction on the prediction performance of the models.

shows only a slight performance reduction (about 0.01 difference) compared to the complete data set (100%). This shows a trade-off between the prediction performance and the computation cost in RF prediction models. In contrast, the effects of information loss due to data reduction cannot be observed in the OLS models. The OLS models produced almost identical prediction performance in all randomized sampling data reduction scenarios. As a linear parametric method, OLS fits the predictors' relationships to be always linear. Therefore, in unbiased conditions, the same linear relationship will be constructed no matter how many data points are used to train the model (to fit the prediction line). This suggests that a small percentage of randomized sampling can be adequate to build an OLS prediction model. Nevertheless, this also means that more available data will not provide additional values for the OLS model.

5.4.5 Effects on Model Computability

Tunneling operations involve incremental learning along the chainages. In this case, the developed data-driven models should be able to learn continuously and incrementally along the tunnel chainages. This can be done by dynamically re-training the model using previously generated data. However, it should be noted that the amount of operation data generated during tunneling continuously increases and can be massive (e.g., can easily achieve TBs), depending on the tunnel length and the construction period. Therefore, computation

scalability should be an important consideration in developing the model.

Figure 5.11 presents the computation profile of RF model training in different data sizes. In this investigation, all the computations were conducted using a single thread. The left panel shows the required computation time along the chainages. It can be observed that the required computation time continuously increases as the amount of training data increases. This highlights the importance of computation scalability for data-driven modeling in tunneling. The right panel shows the computation speed-up, which compares the computation performance of a model relative to the reference model (i.e., 100% training data). The figure shows that reducing the training data to 30% could speed up the computation to about 4-5 times without sacrificing the prediction performance significantly (Figure 5.10). Training with only 5% and 1% data could increase the speed-up to about 40 and 300 times, respectively, with more trade-offs in the prediction performance.

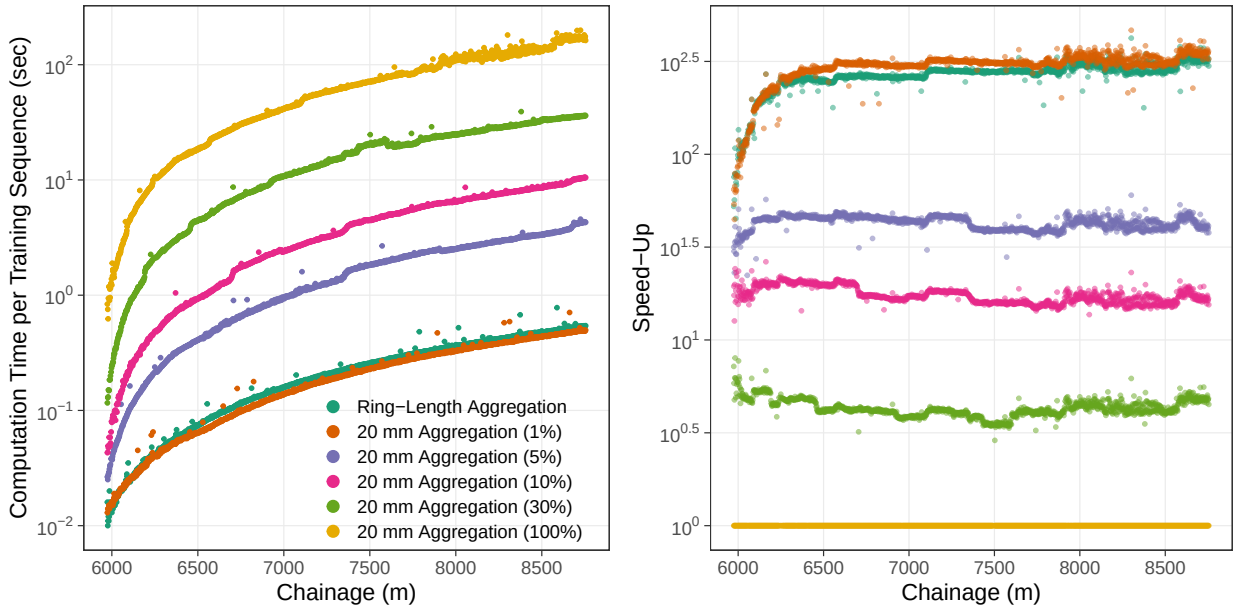

Figure 5.11: Effects of data size on the computation time (left panel) and speed-up (right panel) of the RF models.

Figure 5.11 also shows that the model trained using the coarse aggregate data had similar computation performance to the model trained using 1% of the fine aggregate data, as both had a similar number of observation data points. However, the coarse aggregate data model produced substantially lower prediction performance (see Figure 5.10). This suggests that randomized sampling should be the preferred option for reducing training data size, which confirms the study by Kammerer et al. (2020).

5.4.6 Effects on Model Stability

In this study, the stability of prediction models is represented as the sensitivity of the feature importance ranks in different data aggregation levels and tunneling cases. This can be visualized and measured by plotting feature importance ranks produced by prediction models trained using the 20 mm aggregate data on the x-axis and the ring aggregate data on the y-axis. Visualization of this plot is shown in Figure 5.12. The plot will align perfectly in a diagonal line if both data sets produce the same ranks. In contrast, the plots will be scattered if the ranks are different.

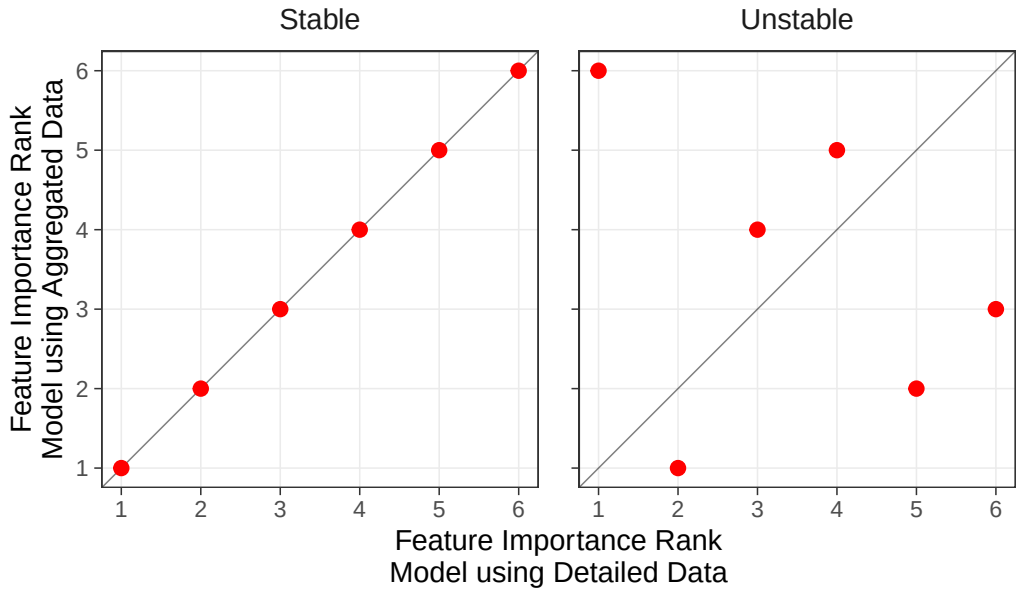

Figure 5.12: Visualization of feature importance comparison plot and the indication to stable and unstable ranks.

Stable Feature Importance Ranks. Figure 5.13 presents prediction models with relatively stable feature importance ranks, where diagonal patterns can be observed. The top and the middle panels present the feature importance rank of the thrust force and cutter torque models, respectively. These panels show that both data aggregation levels produced models with comparable feature importance ranks. This applies to both prediction methods (OLS and RF) and all tunneling cases. The ranks were also consistent. The result shows that cutter torques, thrust forces, chamber pressures, and foam volumes were strongly associated.

This is sensible since cutter torques should be associated with the geologic characteristics (Maher, 2015; Sousa, 2010; Sousa & Einstein, 2012) and proportional to the thrust forces at the same geologic characteristics (Ates et al., 2014; Jakobsen et al., 2013). Furthermore, the

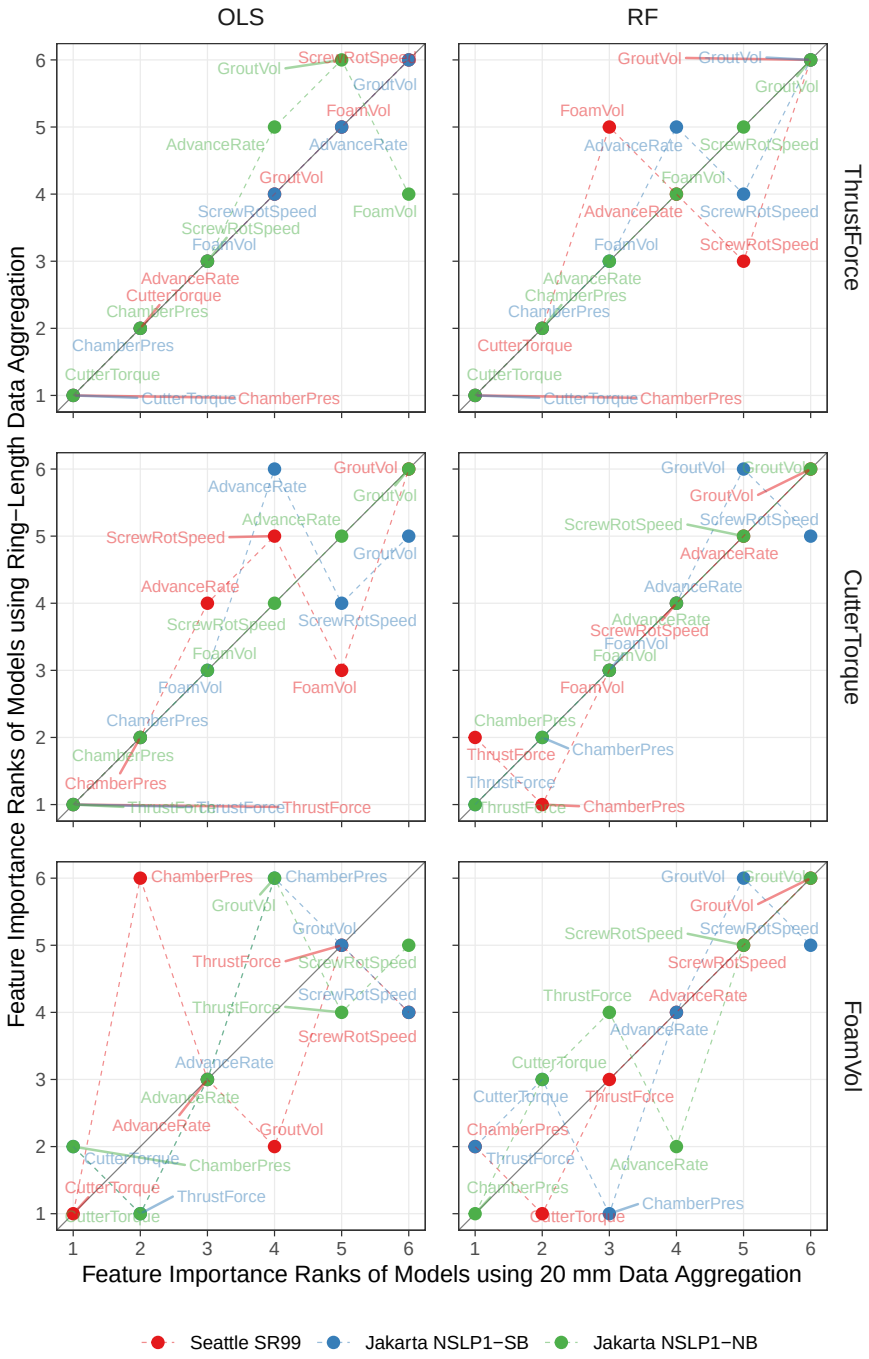

Figure 5.13: Examples of prediction models with relatively stable feature importance ranks. Diagonal patterns of the data points can be observed.

chamber pressures and foam conditioners should also be associated with geologic characteristics. Chamber pressures are continuously regulated during tunneling to balance the earth pressure, which is a function of the confining pressure at a particular depth (Guglielmetti, 2008; JSCE, 2016; Maidl et al., 2013). The deeper the tunneling, the larger the confining pressure, and the more likely the geology has higher strength characteristics. Foam conditioners are typically injected into the excavation face during tunneling to condition the excavated geologic materials (e.g., lower strength, higher plasticity, and better flow behavior) (Peila, 2014; Peila et al., 2019; Thewes & Budach, 2010) and therefore should be related to the geologic characteristics.

The bottom panels present the feature importance rank of the foam volume models. The RF models show that chamber pressures, cutter torques, thrust forces, and advance rates are important predictors of foam volumes. This is sensible since, as discussed previously, foam volume should be associated with the geologic characteristics, and the geologic characteristics should also be associated with the chamber pressure, cutter torque, and thrust force. Several past studies have indicated that TBM advance rates are associated with foam conditioners (Apoji et al., 2022a; Maher, 2015; Roby & Willis, 2014; X. Ye et al., 2017). However, the OLS models produced more rank disagreement. Some results are also less convincing. For example, in the SR99 case, the OLS model trained using the coarse aggregate data considered grout volumes as a high-rank feature and chamber pressures as a low-rank feature. This is different from all other models. This indicates that OLS coefficients can be more sensitive (less stable) to data aggregation effects.

Unstable Feature Importance Ranks. Figure 5.14 presents two prediction models with unstable feature importance ranks. It can be observed that the plots show scattered data points without any apparent patterns. This means models trained with the two different data aggregation levels produced substantially different feature importance ranks. The feature importance ranks were also inconsistent, even for the NSLP1 twin tunnels, which were excavated in the same geologic conditions. The feature importance rank disagreement might be due to none of the considered predictors dominating the response feature or the presence of missing predictors (i.e., other variables outside the considered predictors that strongly governed the response feature).

The advance rate models can be evidence (Figure 5.14, top panels). TBM advance rates are typically regulated by adjusting the other control parameters, such as the thrust force and cutter rotation speed. The target advance rate is determined based on various decision factors, such as the required chamber pressure, geologic conditions, target tunneling performance, TBM behaviors, and even personal driving preference (Sutcliffe 1996). These factors might not be captured in the considered predictors. The grout volume models can be another example of evidence (Figure 5.14, bottom panels). Backfill grouting is a task to close the gap between the newly installed tunnel ring segments and the excavation opening. The grout is typically injected from the shield tail until a certain injection pressure or volume, depending on factors such as the estimated overburden pressure or excavated materials. Thus, grout volumes may not be associated with other features considered in this study.

Measuring Model Stability. The results show that prediction models trained using

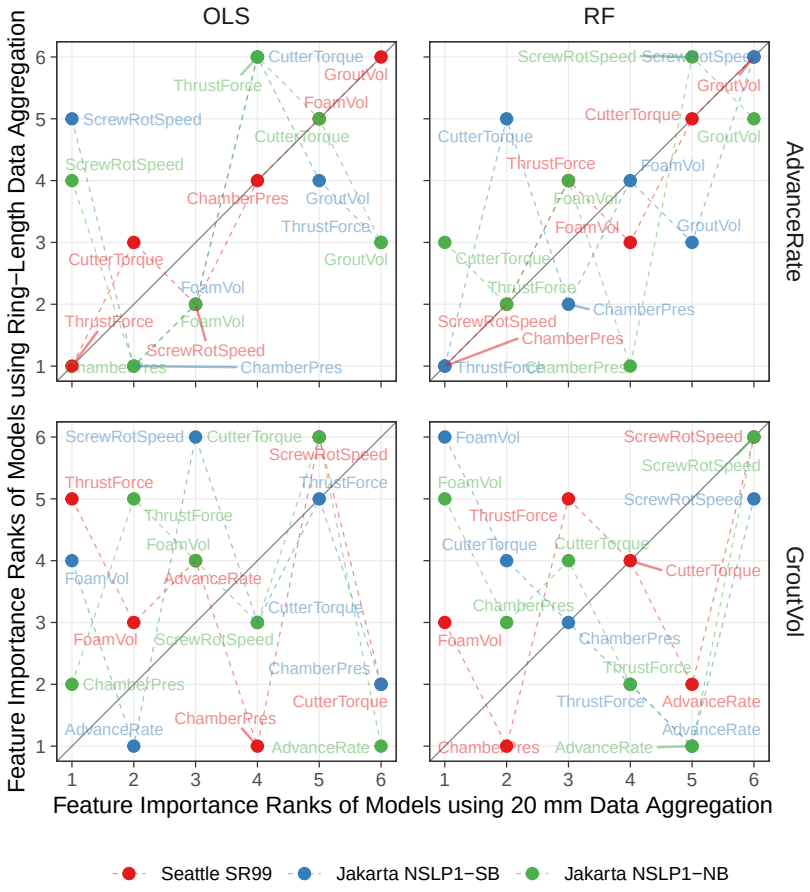

Figure 5.14: Examples of prediction models with unstable feature importance ranks. The data points were scattered without any identifiable patterns

different data aggregation levels may or may not produce similar feature importance ranks. Comparable feature importance ranks could be achieved by models with adequate predictor features, such as the thrust force and cutter torque models. In contrast, substantial disagreements in the ranks were produced by models with insufficient predictor features, such as the advance rate and grout volume models. This indicates that a prediction model with any data aggregation levels may produce a valid feature importance rank in the condition of sufficient predictor features. This also suggests that the stability of feature importance ranks can be used to indicate and measure the sensitivity of a prediction model in different data aggregation levels.

Quantitative measurement of the models' stability was done by computing the coefficient of determination, R^2, of the feature importance comparison plots. If the models produce exactly the same feature importance ranks, the plot will align perfectly in a diagonal line

with an R^2 of 1.0. Therefore, a high R^2 (i.e., closer to 1.0) indicates a stable model. In contrast, a low R^2 (i.e., closer to 0) represents disagreements in the ranks, which indicates an unstable model. Figure 5.15 presents the R^2 of every prediction model for all considered features. The figure shows that most of the considered features produce fairly good stability (i.e., $R^2 > 0.5$), except the advance rate and grout volume models, which have been discussed previously. The figure also shows that, overall, RF models produced better stability than the OLS models. This result is interesting since nonlinear nonparametric machine learning models have been widely seen to be less stable than linear parametric methods such as OLS. This implies that machine learning is a valuable tool when working with complex and nonlinear data sets such as TBM data.

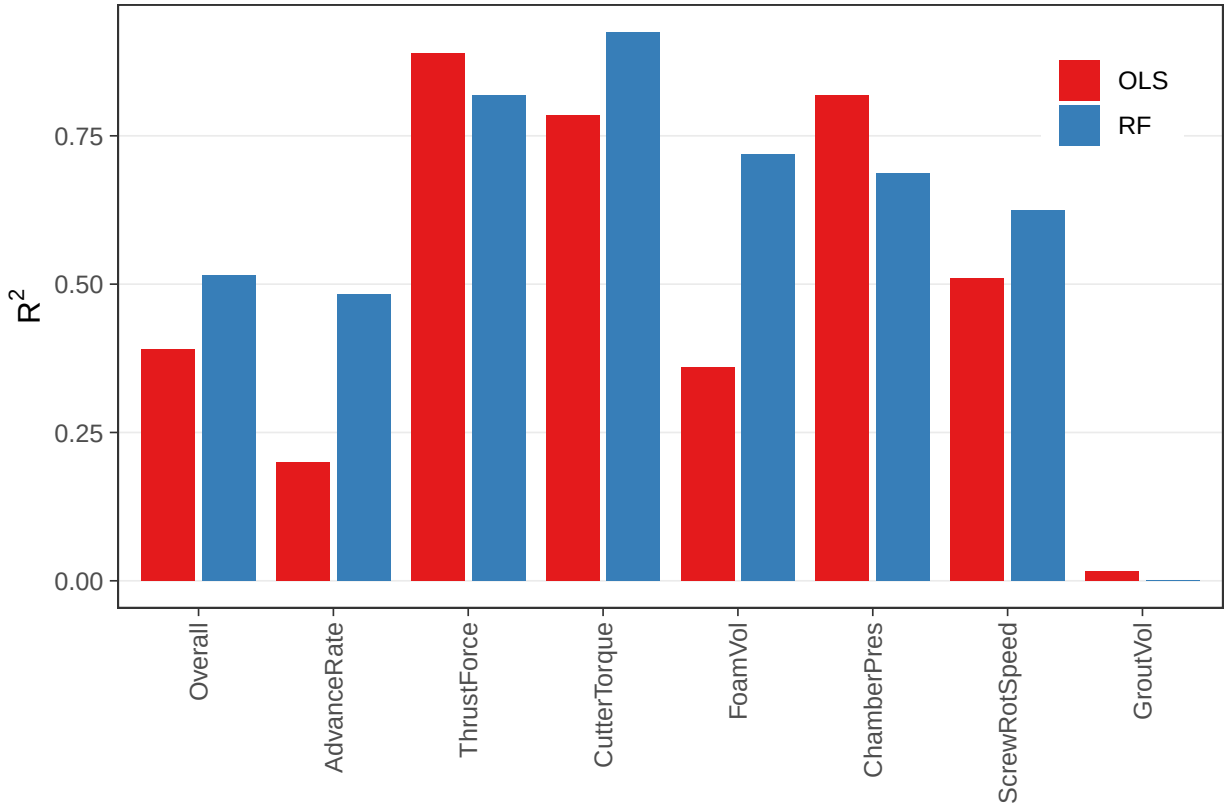

Figure 5.15: Coefficient of determination (R^2) produced from feature importance comparison plots of all prediction models.

5.5 Conclusions

This study has investigated the effects of data aggregation on the statistical characteristics and prediction models of TBM data. The investigation was conducted using three EPB TBM data sets from two tunneling cases, i.e., the SR99 highway tunnel in Seattle, Washington, USA, and the NSLP1 MRT railway twin tunnels in Jakarta, Indonesia. The observation data points for each data set were prepared in two data aggregation levels, i.e., fine aggregate data (20 mm chainage length) and coarse aggregate data (ring-length data points). The effects of data aggregation were analyzed in the framework of the models' predictability, computability, and stability . The main findings of this study are summarized in the following.

- **Effects on statistical parameters**. Data aggregation does not substantially affect the central tendency of TBM data. However, it changes the data variance and distribution. Data with a coarser aggregation level tend to have lower variance. The effects on data distribution can be more sensitive and less predictable.

- **Effects on model predictability**. In the conditions of sufficient observations and predictors, prediction models developed using different data aggregation levels may produce comparable trends of prediction performance. However, models with a coarser aggregation level enjoy higher prediction performance due to the lower variance. This applies to both parametric linear regression (OLS) and nonlinear nonparametric machine learning (RF) models. Therefore, different preparation methods (data aggregation or smoothing methods) may be used to obtain the general behaviors of prediction models, but the prediction performance comparison is only valid at the same data aggregation level.

- **Data reduction and computability**. The training data set of a prediction model should be at the same aggregation level as the testing data set (or future prediction data). Training the model using data using coarser aggregation may cause it fails to capture the variance in the predictions. In reducing training data size, random sampling should be preferred to aggregation. With the same quantity as aggregated data, random sampling may produce a higher prediction performance.

- **Effects on model stability**. The result suggests that prediction models with any data aggregation levels may produce a valid feature importance rank in the condition of sufficient predictor features. Thus, the stability of the rank can be used to indicate and measure the sensitivity of the models in different data aggregation levels. Stable feature importance ranks can be achieved if the predictor features are sufficient to represent the response feature. In contrast, unstable feature importance ranks may imply insufficient (or missing) predictors.

In addition, this study has demonstrated the importance of utilizing nonlinear nonparametric machine learning methods in working with complex data sets such as TBM data,

which is the product of causal-effect interactions of operator control decisions, TBM behaviors, and geologic conditions. This study has shown that RF models consistently outperformed OLS models in both predictions and stability.

These findings have also provided some bases for more research questions, as the following.

- **Finding the optimal aggregation**. Further study can be performed using more data aggregation levels and varying the types of observation data points (i.e., spatial vs. temporal). This may lead to an optimal data aggregation level, which will be an important milestone in developing a unified data preparation method for TBM data.
- **Finding the threshold of sufficient data**. This study shows that the statistical parameters and prediction behaviors jittered in the early chainages but converged and stabilized later. This indicates that there may be a threshold level where the training data can be considered sufficient to produce stable and reliable predictions. Finding this threshold will provide important information on developing data-driven models in tunneling.
- **Evaluating the reliability of feature importance ranks**. This study demonstrates that feature importance analysis is the foundation of measuring the stability of prediction models. However, there are various methods to measure the importance score, and each method can result in different importance scores and ranks (Genuer et al., 2010). Thus, selecting the best method for determining the feature importance rank is still debatable. The permutation-based feature importance was argued to be less susceptible to bias than the standard impurity-based importance (Altmann et al., 2010; Strobl et al., 2007). Nevertheless, this method may not be free from problems, as it may can be misleading, particularly in the presence of strong dependency among the predictor features (Hooker et al., 2021). Further investigation should be performed using different feature importance analysis methods to obtain more robust conclusions.

Chapter 6

Effects of Feature Selection on Prediction Models

6.1 Introduction

6.1.1 Background

Tunnel boring machines (TBM) continuously generate numerous operation data every few seconds during tunneling. Depending on the TBM design, the produced operation data can reach hundreds or even thousands of variable features. This massive data contains valuable information about the tunneling process, which is essential for TBM control decision-making. Various data-driven models have been proposed to utilize the data to support tunneling operations. Most proposed models used machine learning algorithms to predict a response feature based on the high-dimensional predictor features. Many studies have demonstrated encouraging results, suggesting the possibility of further potential to utilize TBM data.

Despite this potential, the data also has an inherent problem. Many of the features are highly correlated. In data-driven models, strongly correlated features cause redundancy, increase the computation cost, and further complicate the model's interpretability. Furthermore, they may produce multicollinearity-related problems in linear prediction models such as the ordinary least squares (OLS). This problem leads to some questions: What are the most appropriate features to be included when developing a tunneling data-driven model? How should the feature selection be performed? What are the effects of the selection? Selecting features from TBM data can be problematic due to the massive size and the level of complexity. Note that the data is a product of causal effect interactions of the ground conditions, TBM behaviors, and human operator decisions. To date, less effort has been given to understanding these features and their effects on the models.

6.1.2 Related Works

In early development, TBM data-driven models were developed using only a few pre-selected TBM features. These features were typically selected based on experience and simple empirical relationships. For example, Suwansawat and Einstein (2006) selected five earth pressure balance (EPB) TBM features to develop artificial neural network (ANN) models for estimating tunneling-induced ground movements, i.e., face pressure, penetration rate, pitching angle, tail grouting pressure, and percent of tail grout filling.

In contrast, Boubou et al. (2010) selected 10 EPB TBM features to develop similar models, i.e., advance rate, cutter torque, thrust force, chamber pressure, tail grouting pressure and volume, horizontal and vertical guidance parameters, the time required for excavation and installation one tunneling lining, and total work for one ring excavation. These examples show that the selected TBM features could vary, even for similar models (ANN) and with the same prediction response (tunneling-induced ground movements).

In the next development, the selection was conducted more systematically using feature selection algorithms (FSA). This allowed more TBM features to be evaluated. Maher (2015) conducted a comprehensive analysis using various FSAs, such as filter, wrapper, and ensemble-based methods, to identify strong predictors for EPB TBM advance rate from hundreds of sensor measurements. These algorithms could identify strongly and weakly related features. However, they often returned inconsistent results, where different algorithms produced different ranks. Furthermore, these algorithms could not capture the mixed causal effect relationships of the features. For example, increased belt conveyor speed should be the effect of the increased advance rate. Therefore, belt conveyor speed should not be a strong predictor of advance rate. These problems show that the standard FSAs are insufficient for the EPB TBM data sets.

To deal with these problems, Maher developed a novel ensemble FSA called the ensemble network normalization algorithm (JENNA). This algorithm could produce better feature selection result stability than a conventional ensemble FSA (Saeys et al., 2008). Furthermore, some selected features from the northbound and southbound tunnel data sets were similar, which indicates the algorithm's capability to generalize from different tunnel data sets. Nevertheless, many of the selected features consisted of individual sensor records. This can mislead the interpretation. For example, foam is typically injected into the excavation face to condition the ground through multiple pipes. If the algorithm returns only a few numbers out of all the foam injection pipe sensors as the selected features, this should not be interpreted that only those particular selected foam injection sensors are the strong predictors. This implies the generalizability problem in FSA-based EPB TBM feature selection.

Recently, more variations of methods have been proposed for TBM feature selection. Many of them were developed using machine learning. Meschke et al. (2019) used embedding and clustering algorithms to select input features for penetration rate and settlement prediction models. They used the t-SNE algorithm (Maaten & Hinton, 2008) to project the original features into an embedding space. Then, the clustered the embedded feature points using the HDBSCAN algorithm (Campello et al., 2013). Feature points with clusters were

validated using the Spearman correlation. Only one feature point was selected to represent every cluster with a relative correlation value less than a specified threshold. Feature points with no cluster were considered unique features. Mokhtari and Mooney (2020) used a feature filtering technique called RReliefF (Robnik-Šikonja & Kononenko, 2003) to select input features for an advance rate prediction model. They used a pair-wise correlation of high-ranking features to identify feature redundancies. The redundant features were removed, remaining a single representative feature.

6.1.3 Objectives

The discussed studies highlight the variation of features used in TBM data-driven models. Until now, a unified approach to selecting TBM features has not been available (Marcher et al., 2020; Sheil, Suryasentana, Mooney, et al., 2020). This limits the interpretability and generalizability of the models. Recent studies have indicated that feature selection guided by domain knowledge may improve the stability of the results, such as in transport economics (Groves & Gini, 2015), healthcare (Radovanovic et al., 2015), and biological data (Raghu et al., 2017). However, no study has investigated the implementation of knowledge-based feature selection for TBM data. The effects of the feature selection on the prediction models have not also been identified.

This study aims to (i) propose a generalized knowledge-based taxonomy of EPB TBM features and (ii) investigate the effects of the knowledge-based feature selection on prediction models. The knowledge-based taxonomy was developed by synthesizing various tunneling information such as published literature and guidelines, standard codes, as well as the experience of engineers and operators. The effects of the feature selection on the models were investigated in the framework of the models' predictability, computability, and stability (generalizability) (B. Yu & Kumbier, 2020).

Machine learning models have been widely seen as "black box" models. The models can yield high prediction accuracy but are difficult to be interpreted. Understanding the input features should be a necessary attempt to open the models. The proposed taxonomy is expected to serve as a guide for EPB TBM feature selection and facilitate a unified method of TBM data preparation.

6.2 Data

6.2.1 Tunneling Cases

This study was performed using the same EPB TBM data sets used in Chapter 5, i.e., (i) the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA, and (ii) the North-South Line Phase-1 (NSLP1) Mass Rapid Transit (MRT) railway twin tunnels in Jakarta, Indonesia. More discussion of these tunneling cases can be found in Chapter 5.

6.2.2 Data Description and Preparation

As discussed in Chapter 5, the three data sets consisted of enormous tunneling operation measurement records generated by numerous sensors installed on the TBM. To limit the scope, this study only considered numerical features from continuous measurements. All categorical features (e.g., features of status levels and switch on/off) were removed. Furthermore, this study focused on the primary EPB TBM processes during the tunneling excavation phase, i.e., excavating, advancing, steering, ground conditioning, earth pressure balancing, muck discharging, and tail grouting processes. Features related to the non-excavation phase, such as segment lining erection, and features related to mechanical and electrical components, such as gear and oil conditions, were not considered.

The original observation data points of the three data sets are in 5 seconds. However, to minimize the data noise and computation cost, this study used ring-length aggregate spatial series as the observation data points. This means an observation data point is the average measured values within a ring length for gauge metrics (e.g., features of pressures, forces, torques, and speeds) or the final measured value of a ring for cumulative increment metrics (e.g., features of volumes). Note that the data aggregation should not affect the validity of the results. As concluded in Chapter 5, in the condition of sufficient observation data points and predictor features, this aggregation should not substantially affect the model's prediction behaviors (error trends) and generalizations (feature importance rank).

Subsequent to the initial preparation, the data sets were cleaned according to the following steps, (i) volume normalization, (ii) removal of non-excavation observation data points, (iii) removal of erroneous records, and (iv) removal of constant features. More discussion of these steps can be found in Chapter 5. The number of features and observation data points returned from all of these data preparation are summarized in Table 6.1.

Figure 6.1 presents the correlation networks of the three prepared data sets. Each node in the networks represents an EPB TBM feature. The clustered nodes indicate strongly correlated features. The blue and red lines represent positive and negative correlations, respectively. The stronger the color, the higher the correlation value. The plots were built using a correlation threshold of 0.5, meaning only correlations higher than this value are shown. This figure visualizes high multicollinearity among the EPB TBM features.

6.3 Methods

This study aims to propose a generalized knowledge-based taxonomy of EPB TBM features and investigate the effects of feature selection on prediction models. This was achieved by conducting three main tasks, as listed below.

1. **Developing a taxonomy for selecting features**. This task was performed to develop a generalized taxonomy of EPB TBMs. This was achieved by combining knowledge-based feature hierarchy and correlation-based filtering.

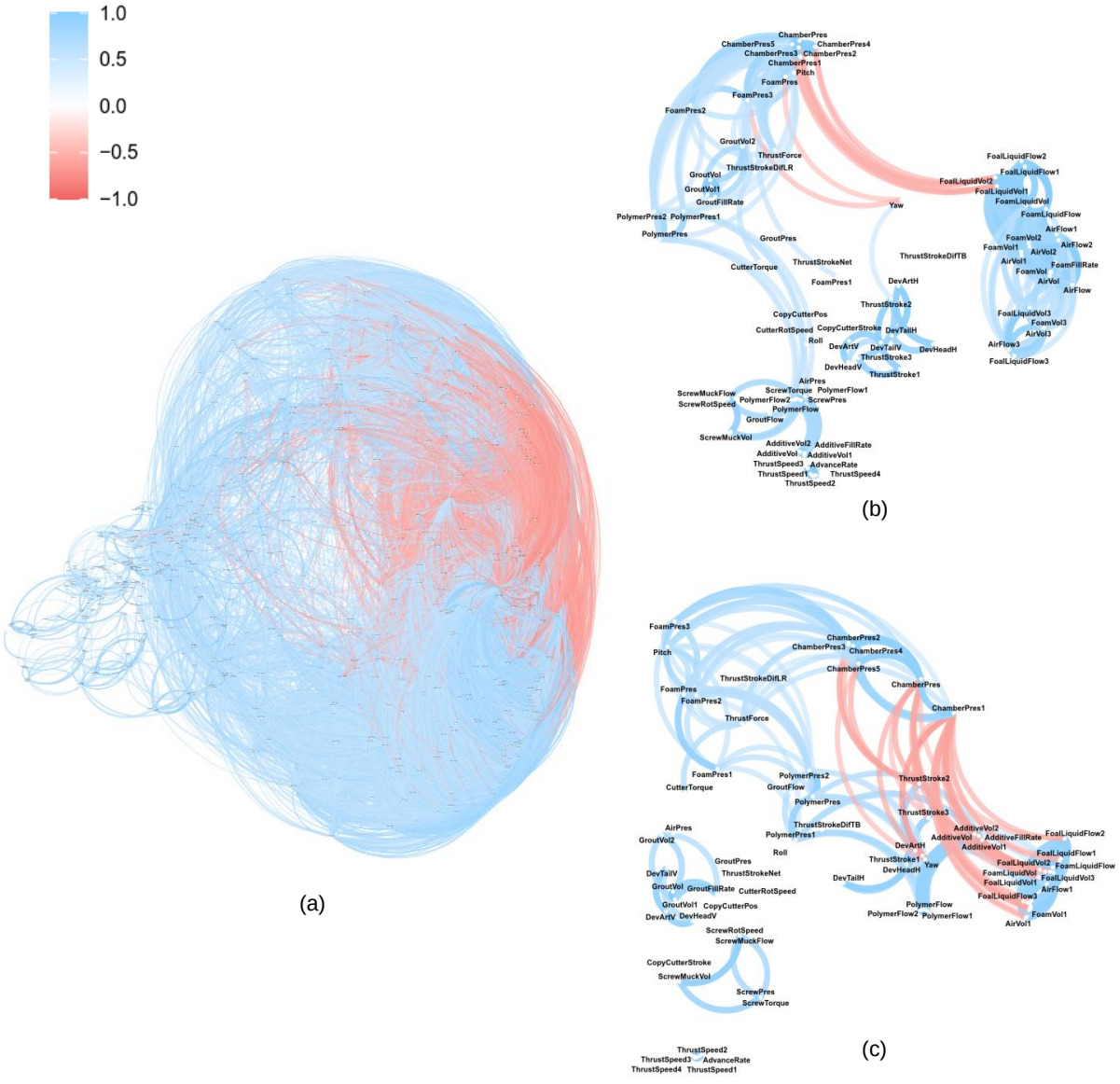

Figure 6.1: Correlation networks of EPB TBM data: (a) SR99, (b) NSLP1-SB, (c) NSLP1-NB

Table 6.1: Dimension of the data sets before and after data preparation.

Data Dimension	Data Preparation	SR99	NSLP1-SB	NSLP1-NB
Features	Original	5975	467	467
	After removing non-continuous variable features	2306	247	247
	After removing features not directly related to main EPB TBM functions	601	79	79
Observation	Original	1416	1771	1777
	After removing non-excavation phases, missing data, and erroneous records	1253	1684	1695

2. **Evaluating multicollinearity**. This task was performed to demonstrate the presence of multicollinearity in EPB TBM data and evaluate the effects on the prediction models. This was achieved by measuring the variance inflation factor (VIF) of the data and developing least squares-based regression models.

3. **Investigating the effects of feature selection on the prediction models**. This task was performed to the effects of the selected feature set on prediction models. This was achieved by developing models using two different feature sets (i.e., with and without the feature selection) and investigating the models' predictability, computability, and generalizability.

More descriptions of the methods can be found in the following sections.

6.3.1 Hierarchical Feature Selection

The EPB TBM feature taxonomy was developed by combining knowledge-based feature hierarchy and correlation-based filtering. The knowledge-based feature hierarchy was constructed by synthesizing various tunneling information such as published literature and guidelines, standard codes, as well as the experience of engineers and operators. This resulted in four levels of EPB TBM feature hierarchy, i.e., (i) the process level, (ii) the component level, (iii) the unit measurement level, and (iv) the individual sensor level.

1. **Process level**. At the process level, all the EPB TBM features were categorized into seven partial processes of EPB tunneling, i.e., excavation, advancing, steering, ground conditioning, earth pressure balancing, muck discharging, and tail grouting processes.

This category was determined based on the main EPB TBM processes (Chapter 2). Note that all processes unrelated to the tunneling excavation phase, such as segment lining erection, were not included.

2. **Component (and sub-component) level**. At the component level, features in each process group were categorized into several groups of related TBM system components. For example, the excavation process is associated with features related to the cutter head and copy cutter systems. The advancing process is associated with features related to TBM advance rate and thrust systems. Groups of sub-components were included in some components with more complexities.

3. **Unit measurement level**. Features in each group of components (or sub-components) that measure the same unit were grouped at the unit measurement level. For example, the thrust system was divided into thrust speed (mm/min), force (kN), pressure (bar or kPa), and stroke (mm).

4. **Individual sensor level**. The individual sensor level consisted of features produced by the same sensor. For example, a TBM may have several pressure sensors installed on different thrust jack groups, foam injection pipes, and locations in the excavation chamber.

The feature taxonomy was developed by aggregating every feature at the individual sensor level into a feature at the unit measurement level. Subsequently, features at every process level were selected by considering (i) Spearman correlation, (ii) feature redundancy, and (iii) applicability on different EPB TBM data sets.

6.3.2 Multicollinearity Analysis

Multicollinearity is a condition when the predictors of a regression model are correlated. The presence of multicollinearity in the model can be measured using the variance inflation factor (VIF). Consider y as a response feature with $x_1, x_2, \ldots, x_p$ as the predictors. The VIF of the predictors can be evaluated by first regressing every predictor i as a response to the other predictors, i.e.,

$$\begin{aligned} x_1 &= \alpha_0 + \alpha_1 x_2 + \alpha_2 x_3 + \cdots + \alpha_{p-1} x_p + e \\ x_2 &= \alpha_0 + \alpha_1 x_1 + \alpha_2 x_3 + \cdots + \alpha_{p-1} x_p + e \\ &\cdots \\ x_p &= \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_{p-1} x_{p-1} + e \end{aligned} \tag{6.1}$$

where α is the regression coefficients, and e is the residual errors. Then, the VIF of every predictor i can be calculated as

$$VIF_i = \frac{1}{1 - R_i^2} \tag{6.2}$$

where R_i^2 is the coefficient of determination of every estimated predictor i. The multicollinearity analysis was performed by measuring the VIF of data with and without the feature selection. The effects of multicollinearity on prediction models were also investigated by developing regression models using the two feature sets.

6.3.3 Model Setup

The effects of feature selection on prediction models were investigated by developing models using two different feature sets, i.e., with and without the hierarchical feature selection. Two TBM features were selected as the response feature, i.e., cutter torque and thrust force. Each response feature was predicted using the other features as predictors. The schematic diagram of the prediction model is presented in Figure 6.2. Similar to the prediction model in Chapter 5, this model does not consider correlations between observation data points (Zeger & Liang, 1992).

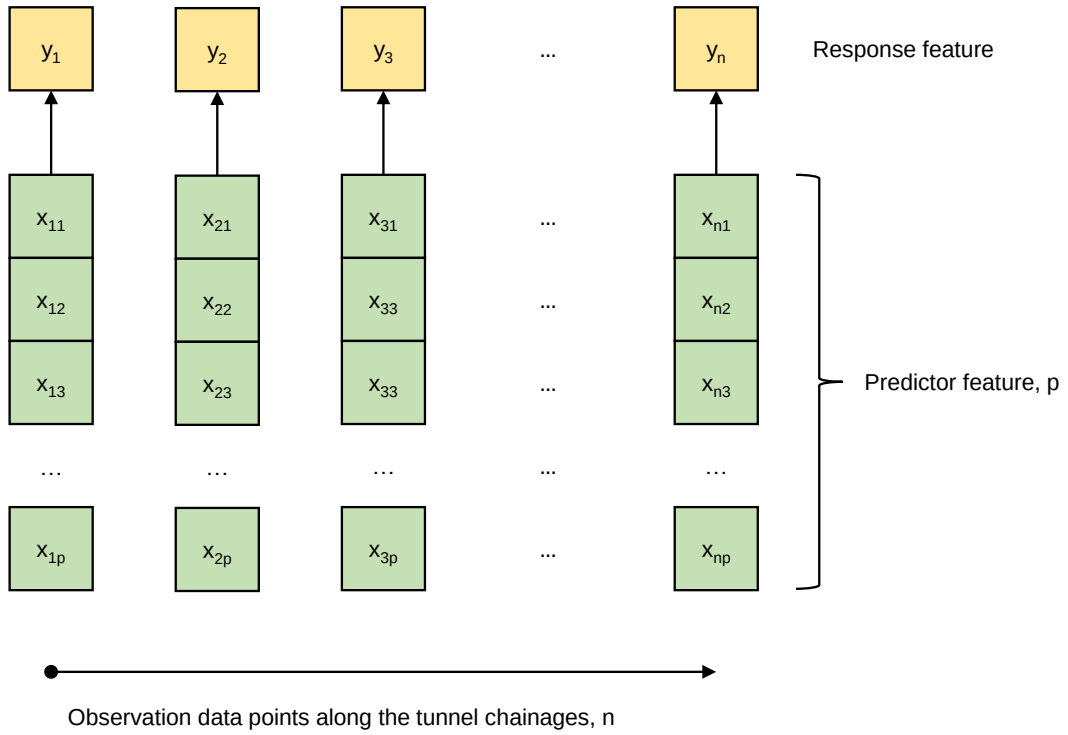

Figure 6.2: Schematic diagram of the prediction model.

The data was split in a static random scheme with a training and testing ratio of 70:30 (Figure 6.3). Note that this static random scheme does not represent the actual tunneling process, which should be dynamic and sequential. However, this scheme is considered adequate since this investigation aims to compare the effects of different feature sets, not to simulate the tunneling process.

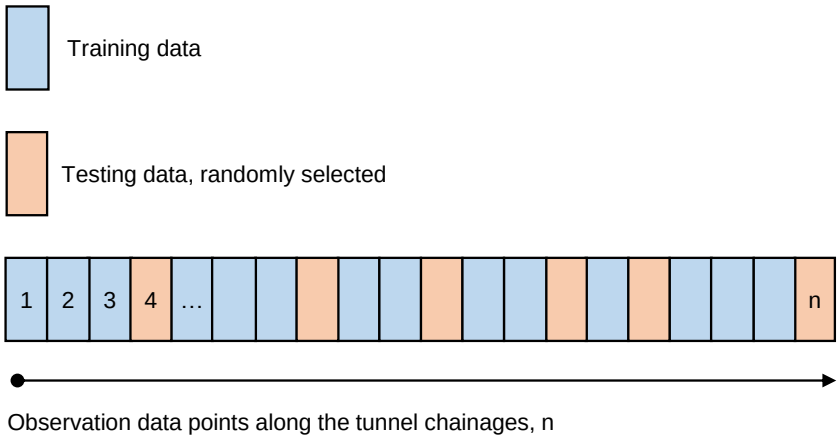

Figure 6.3: Schematic diagram of static random data splitting for model training and testing.

6.3.4 Prediction Methods

Four prediction methods were used in this study, i.e., Ordinary Least Squares (OLS), Ridge regression, the Least Absolute Shrinkage and Selection Operator (the Lasso), and Random Forests (RF).

- **OLS**, as discussed in Chapter 5, is a widely used linear regression method that uses regression coefficients to define a response variable, y_i, as a linear function of the predictor variables, $x_1, x_2, \ldots, x_p$. This method was selected to represent parametric linear prediction methods.
- **Ridge regression** is a type of regularized linear prediction method. This method has the same RSS as OLS, but the regression coefficients are estimated by minimizing the RSS with a constraint. This can reduce the variance and make the prediction results more robust to overfitting, thus leading to better generalization. This method is commonly used to deal with data sets with multicollinearity problems. In this study, ridge regression was performed using the algorithm implementation by Friedman et al. (2010).

- **The Lasso** is another type of regularized linear prediction method. This method estimates the regression coefficients by minimizing the RSS with an absolute constraint. Unlike Ridge regression which keeps all predictors in the prediction results, the Lasso may shrink some of the predictors to zero. This shrinkage capability leads to better interpretability of the prediction results. The Lasso was selected as a linear embedded feature selection method. In this study, the Lasso was performed using the algorithm implementation by Friedman et al. (2010).

- **RF** is an ensemble-supervised learning algorithm that aggregates a large number of decision trees to produce predictions (Breiman, 2001). RF was selected to represent nonparametric nonlinear machine learning and embedded feature selection methods. This study used the RF fast implementation in C++ and R (Wright & Ziegler, 2017).

More discussion of these prediction methods and their mathematical bases can be found in Chapter 4.

6.3.5 Model Hyperparameters

Hyperparameter analysis was performed on all the models to select the best hyperparameter configuration. The analysis was done by conducting repeated cross-validation (i.e., five times repeat, 10-fold CV) in various ranges of RF hyperparameters, i.e.,

(i) the number of trees (ntrees).

(ii) the possible number of features that are randomly selected to split at each node (mtry),

(iii) the minimum node size, and

The ntrees define how the forest is to be generated (i.e., the number of constructed independent decision trees). Both mtry and the minimum node size define the complexities of the constructed trees. The estimated response variance was used as the split rule in this regression problem (Wright & Ziegler, 2017). This defines how a single decision tree is to be constructed. The model performance was evaluated using the out-of-bag samples (OOB).

The hyperparameter analysis results are presented in Figure 6.4 and Figure 6.5. These figures show that decent model performance could be obtained in adequate number of ntrees (e.g., more than 100), mtry of about a third of the predictors, p, and small number of minimum node size. This applies to the three different tunneling cases. Therefore, this study used ntrees of 500, mtry of p/3, and minimum node size of 5. This configuration agrees with the recommendation from previous studies such as Probst et al. (2019).

6.3.6 Model Evaluation and Feature Importance

The prediction models were evaluated using three performance metrics, i.e., the coefficient of determination (R^2), mean absolute error (MAE), and normalized mean absolute error

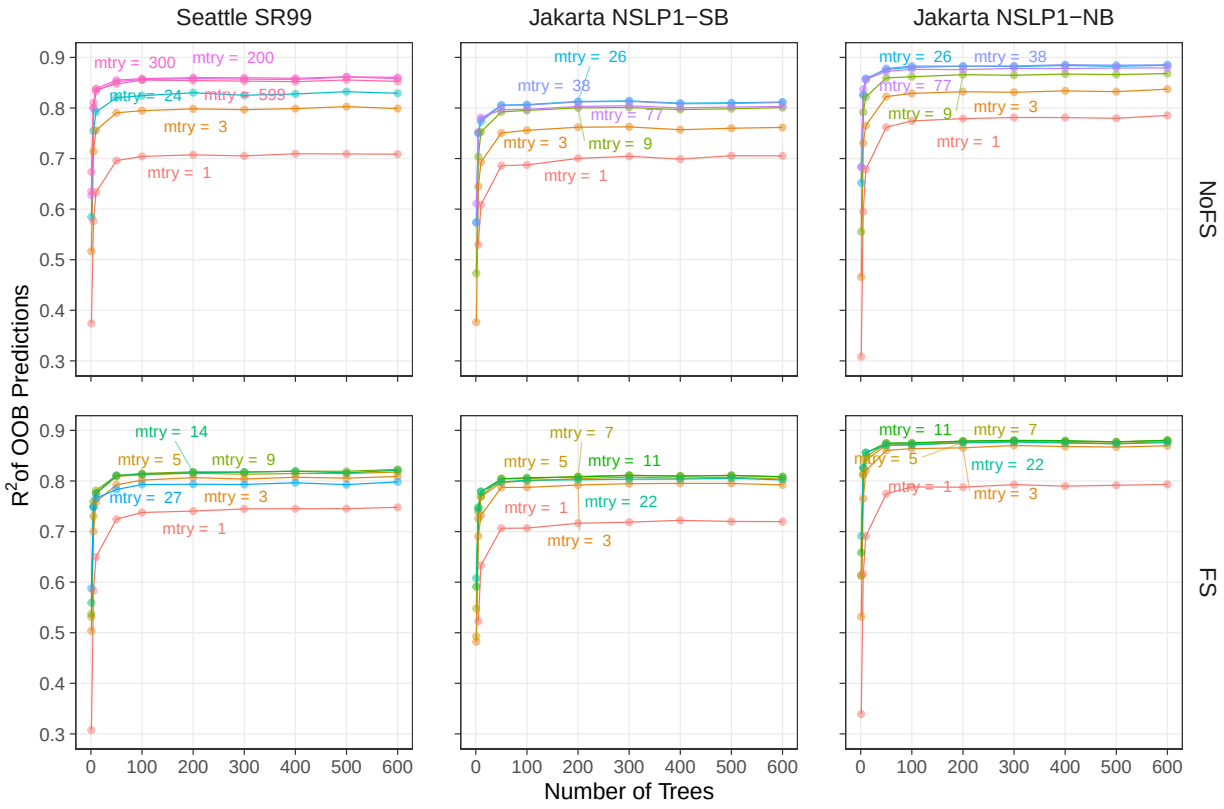

Figure 6.4: Effects of ntrees and mtry hyperparameters on models with different feature sets and tunnelling cases.

(NMAE). The R^2 was used as the goodness of fit and is defined as

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}. \tag{6.3}$$

The MAE and NMAE was used to accommodate comparisons of MAE across the features with different units, as discussed in Chapter 5.

The generalizability of the prediction models was evaluated with the same method as Chapter 5. Briefly, a model can be considered generalizable if it has stable feature importance ranks in various model settings and input, i.e., different prediction methods, data aggregation levels, and tunneling cases. The feature importance rank of OLS models was obtained from the absolute value of the regression coefficients. The feature importance rank of RF models was obtained from the permutation importance scores. More discussion on the permutation feature importance can be found in Chapter 5.

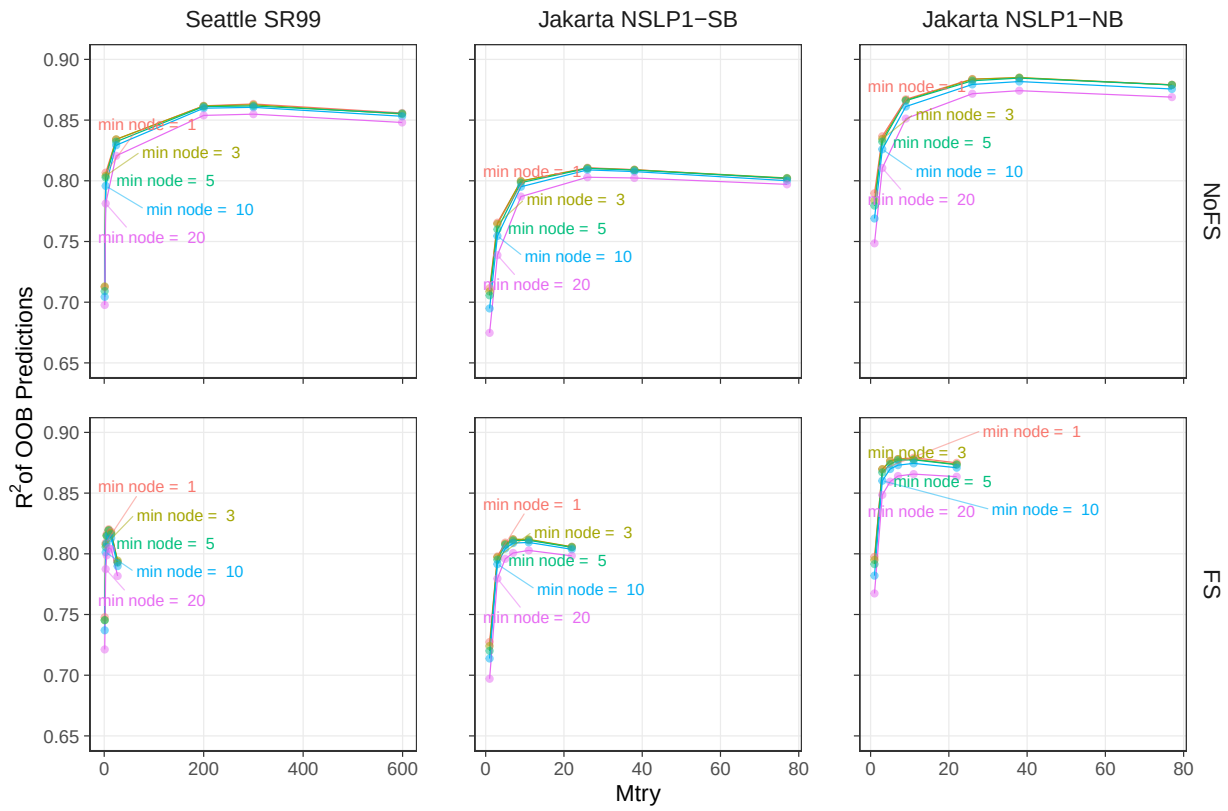

Figure 6.5: Effects of mtry and minimum node size hyperparameters on models with different feature sets and tunnelling cases.

6.4 Results and Discussion

6.4.1 Generalized Feature Taxonomy

As described in Section 6.3.1, the taxonomy was developed based on four levels of hierarchy, i.e., (i) the process level, (ii) the component level, (iii) the unit measurement level, and (iv) the individual sensor level. Figure 6.6 presents the proposed taxonomy of EPB TBM features. It can be seen that the process level consists of seven groups of EPB tunneling processes, i.e., excavating, advancing, steering, ground conditioning, earth pressure balancing, muck discharging, and tail grouting. Each group branches into several related EPB TBM system components. For example, the excavation process consists of features related to the cutter head and copy cutter systems. The ground conditioning process consists of features related to foam, polymer, slurry, and additives. Note that some components with more complexities may have sub-component levels to allow a better representation of the system. For example,

foam can be subdivided into features related to foam, air, foaming liquid, water, foaming agent, and polymer agent.

Figure 6.6: Generalized taxonomy of EPB TBM features. The generalized features are highlighted in yellow.

Each component (or sub-component) branches further into features with the same measurement units. For example, the cutter head system consists of torque (kNm) and rotation speed measurements (rpm). The foam conditioning system consists of volume (m3), flow (m3/min), and injection pressure (kPa) measurements. The unit measurement level

is formed by combining every associated individual sensor. The individual sensors are not shown in the figure due to the excessive number. This aggregation is important since the combined value of measurements contains more generalizable information than the individual sensor measurement. For example, an EPB TBM can have several foam pipes, with a flow sensor installed in each pipe. The total foam flow should provide more information on the overall tunneling process than foam flow in individual pipes. (Note that the information at the individual sensor level can also be useful in some detail tunneling operation tasks.)

The generalized EPB TBM features are highlighted in yellow. These features were selected by considering (i) Spearman correlation filtering in every group of processes, with a threshold value of 0.8, (ii) feature redundancy, where only one feature was selected from comparable measurements, e.g., flow and volume, and (iii) applicability on the three tunneling data sets, which enables feature generalization. Figure 6.7 presents the correlation networks of the selected features. It can be observed that the networks are more decipherable (less crowded) than networks without feature selection (Figure 6.1). The nodes are more spread, and the links have lighter colors. This indicates lower correlations among the retained features.

6.4.2 Multicollinearity Problems

The presence of multicollinearity in the EPB TBM data sets was measured using the VIF. Figure 6.8 presents the VIF of several EPB TBM features with and without the hierarchical feature selection (HFS). Note that the x-axis is presented in a log scale due to the extreme values. Typically, a VIF of 5 can be considered high multicollinearity (Sheather, 2009). This figure shows that without the feature selection, most features have extreme VIFs. Some features even reach hundreds of VIF or above, which indicates serious multicollinearity issues. This high VIF appears in all three data sets. The VIF drops after applying the feature selection with no VIF higher than 10. This result quantitatively validates the presence of severe multicollinearity in EPB TBM features.

Multicollinearity may cause problems when developing linear regression-based prediction models such as OLS. Figure 6.9 presents the prediction results of (a) cutter torque and (b) thrust force using all other features as predictors. The models were developed using the SR99 data set without the hierarchical feature selection. The black and red colors represent the actual values and the results of OLS predictions, respectively. Some extreme discrepancies (prediction errors) can be observed, i.e., between chainages 6250 and 6300 in the cutter torque model and between chainages 7200 and 7600 in the thrust force model. Note that the y-axis is presented in a log scale due to the extreme errors. These errors are in different locations, indicating that they are not caused by outliers in the predictors. Furthermore, the outliers of the data sets have been removed, as described in section 6.2.2.

These extreme errors are most likely caused by statistical problems known as response outliers and high leverage (James et al., 2013). These problems can occur when the predictors of an OLS model are multicollinear. This can happen even if the predictors are within normal ranges without any outliers. The multicollinearity problems can be eradicated by

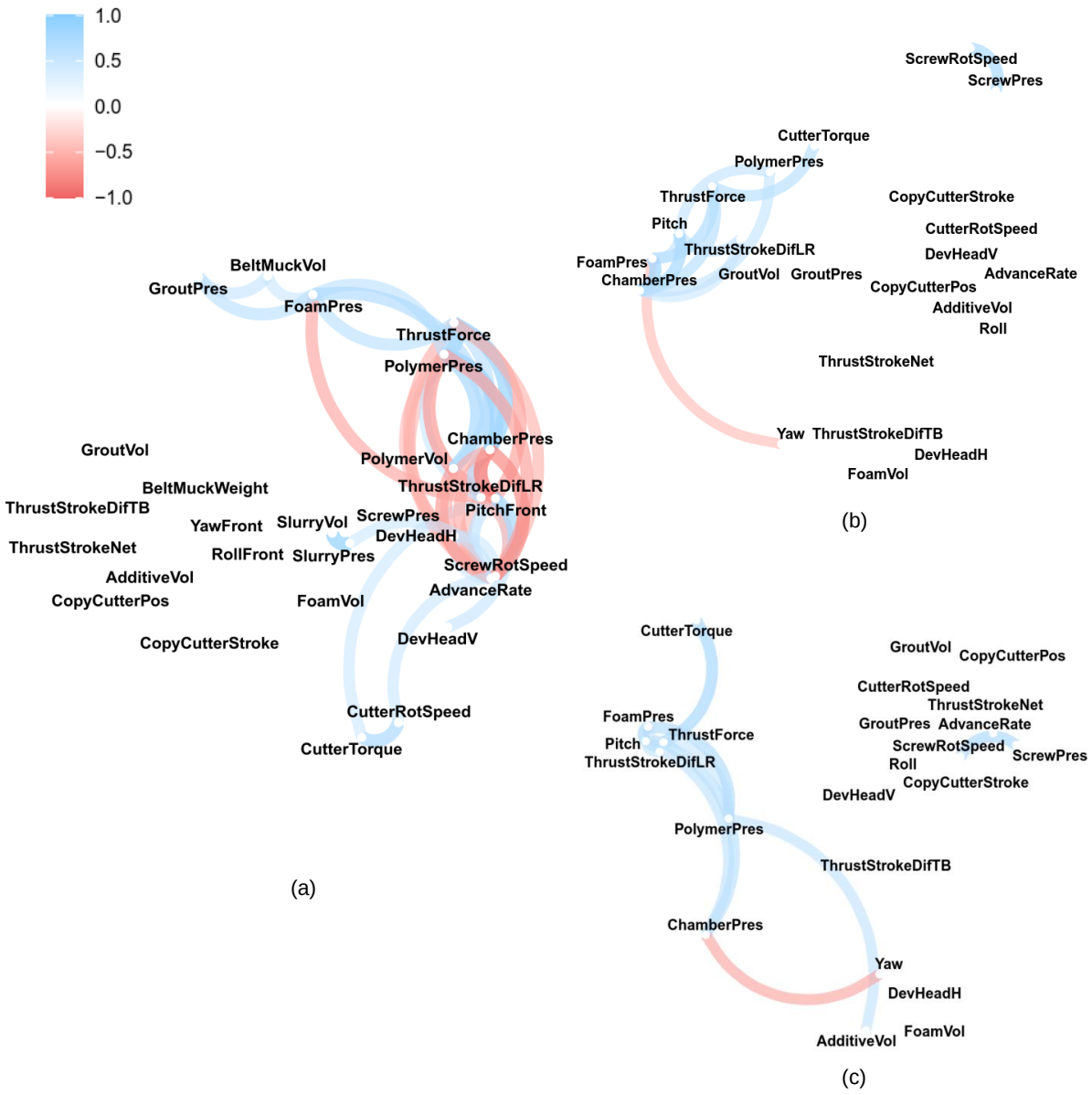

Figure 6.7: Correlation networks of EPB TBM data with the hierarchical feature selection: (a) SR99, (b) NSLP1-SB, (c) NSLP1-NB.

Figure 6.8: Variance inflation factors (VIF) of the EPB TBM data with and without the hierarchical feature selection.

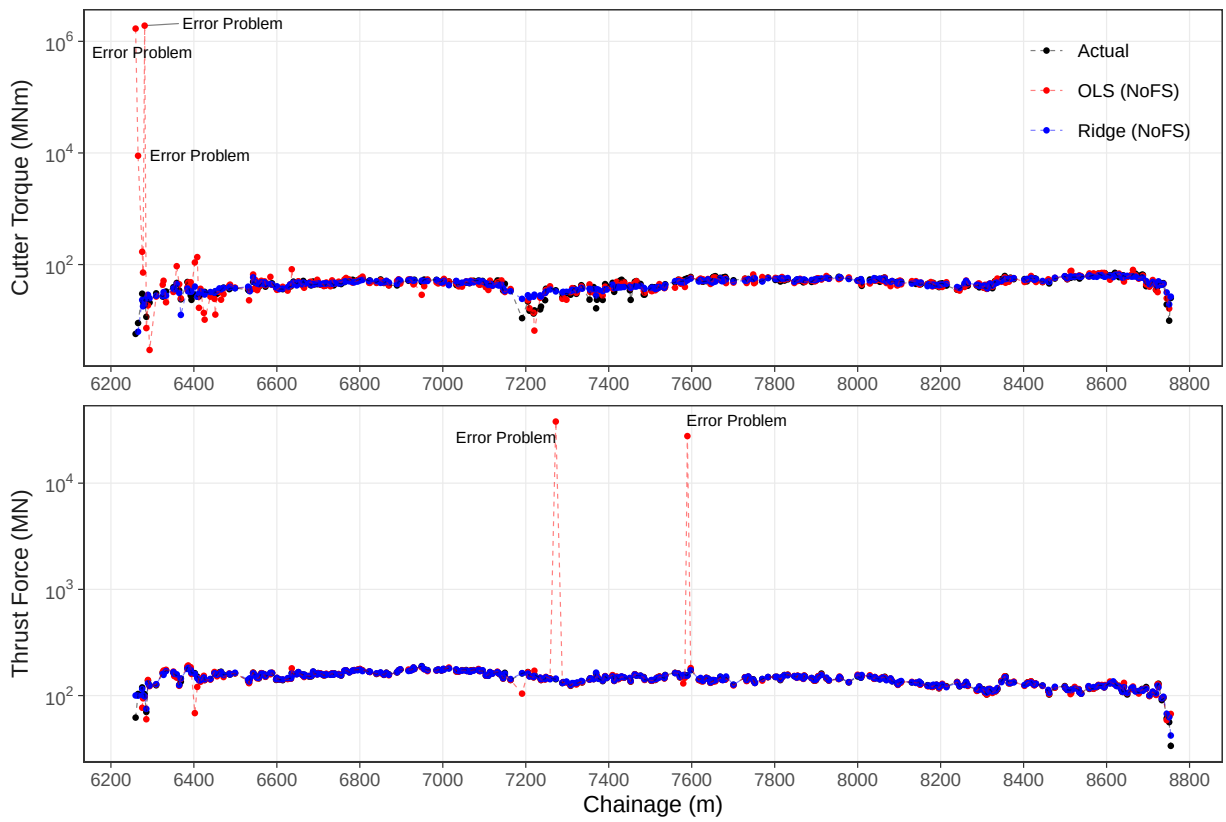

Figure 6.9: Extreme errors of OLS predictions due to severe multicollinearity in the data.

regularization, such as using Ridge regression. The blue color represents the ridge regression results. It can be observed that the ridge regression model did not produce extreme errors, even though this model was developed using the same predictor as the OLS model. This result provides evidence of multicollinearity problems in TBM prediction models.

The prediction performance of cutter torque and thrust force models in two feature sets, i.e., with and without the hierarchical selection, was investigated using four different prediction methods, i.e., OLS, Ridge, the Lasso, and RF. Table 6.2 presents the overall performance, which is quantitatively represented as MAE, NMAE, and R2. The result shows that OLS models without the feature selection produced unreasonable overall prediction performance. The extremely inflated errors were due to the presence of multicollinearity. This can be confirmed since these errors did not occur in OLS models with the selected feature set. Furthermore, the problems also did not occur in regularized methods (i.e., the Ridge and Lasso) and nonlinear nonparametric machine learning method (i.e., RF)

This table also shows that the hierarchical feature selection did not substantially affect

Table 6.2: Overall prediction performance of different models and feature sets.

Response Feature	**Predictor Set**	**Prediction Method**	**MAE**	**NMAE** (%)	R^2
Cutter torque	Before HFS	OLS	454952.78	980.85	0.04
		Ridge	3696.72	7.97	0.82
		Lasso	4007.50	8.64	0.80
		RF	3019.19	6.51	0.89
	After HFS	OLS	4366.99	9.41	0.76
		RF	3013.62	6.50	0.89
Thrust force	Before HFS	OLS	1016535.74	711.34	0.03
		Ridge	1981.47	1.39	0.98
		Lasso	880.56	0.62	0.97
		RF	3414.66	2.39	0.96
	After HFS	OLS	5153.47	3.61	0.91
		RF	4249.89	2.97	0.94

*HFS: Hierarchical Feature Selection

the RF prediction performance. The cutter torque prediction models with the two different feature sets produced the same relative error of 6.5% and R^2 of 0.89. The thrust force model with the selected feature set produced a relative error of 2.97% and R^2 of 0.94, which is only slightly lower than the counterpart model with a relative error of 2.39% and R^2 of 0.96. This result indicates that prediction models developed using the selected feature set can be more robust to severe errors, especially due to multicollinearity problems, highlighting the importance of feature selection for TBM data sets. This result also suggests that the proposed features do not substantially affect the prediction performance, especially if the model is developed using a nonparametric nonlinear machine learning prediction method such as RF.

6.4.3 Effects on Model Predictability

Figure 6.10 compares the distribution of cutter torque (top panels) and thrust force (bottom panels) prediction values produced by models developed using different feature sets. The left, middle, and right panels present the results from the SR99, NSLP1-S, and NSLP1-N data sets. The black lines represent the measured cutter torque and thrust force values. The red and blue lines represent the OLS and RF predictions, respectively. The solid and dashed lines represent the model developed with and without the hierarchical feature selection,

respectively. The figure shows that both of the feature sets produced comparable prediction distributions. This applies to both OLS and RF models. The RF models produced almost identical distributions. In comparison, the OLS models produced slight discrepancies in some parts of the distributions. Interestingly, the OLS model developed using the selected feature set produced closer distribution to the measured values. This suggests that the hierarchical feature selection may not substantially affect the distribution of prediction values from RF models. However, the selection can be beneficial for OLS prediction models.

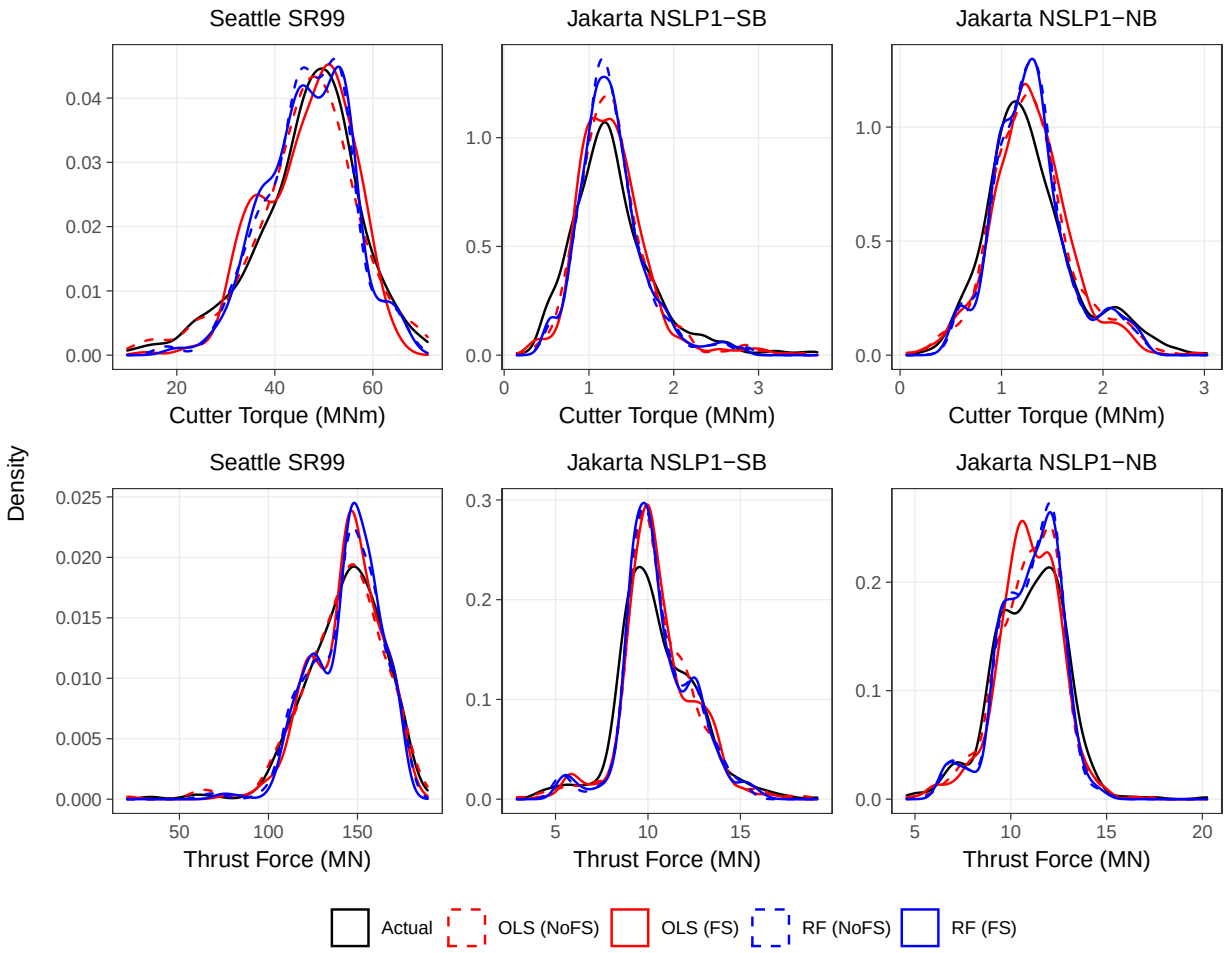

Figure 6.10: Comparison of prediction value distributions produced by models with different prediction methods, feature sets, and tunneling cases.

In more detail, Figure 6.11 compares every prediction point produced by both feature sets. The panel arrangement is the same as Figure 6.10. The x-axes represent predictions

from models developed using without the feature selection. The y-axes represent predictions from models with the selected feature set. The red and blue points represent the OLS and RF predictions, respectively. The figure shows that prediction points in every panel formed a diagonal pattern. This indicates that both of the feature sets produced comparable prediction points. The RF models consistently outperformed the OLS models, which can be observed by narrower prediction bands of the diagonal pattern.

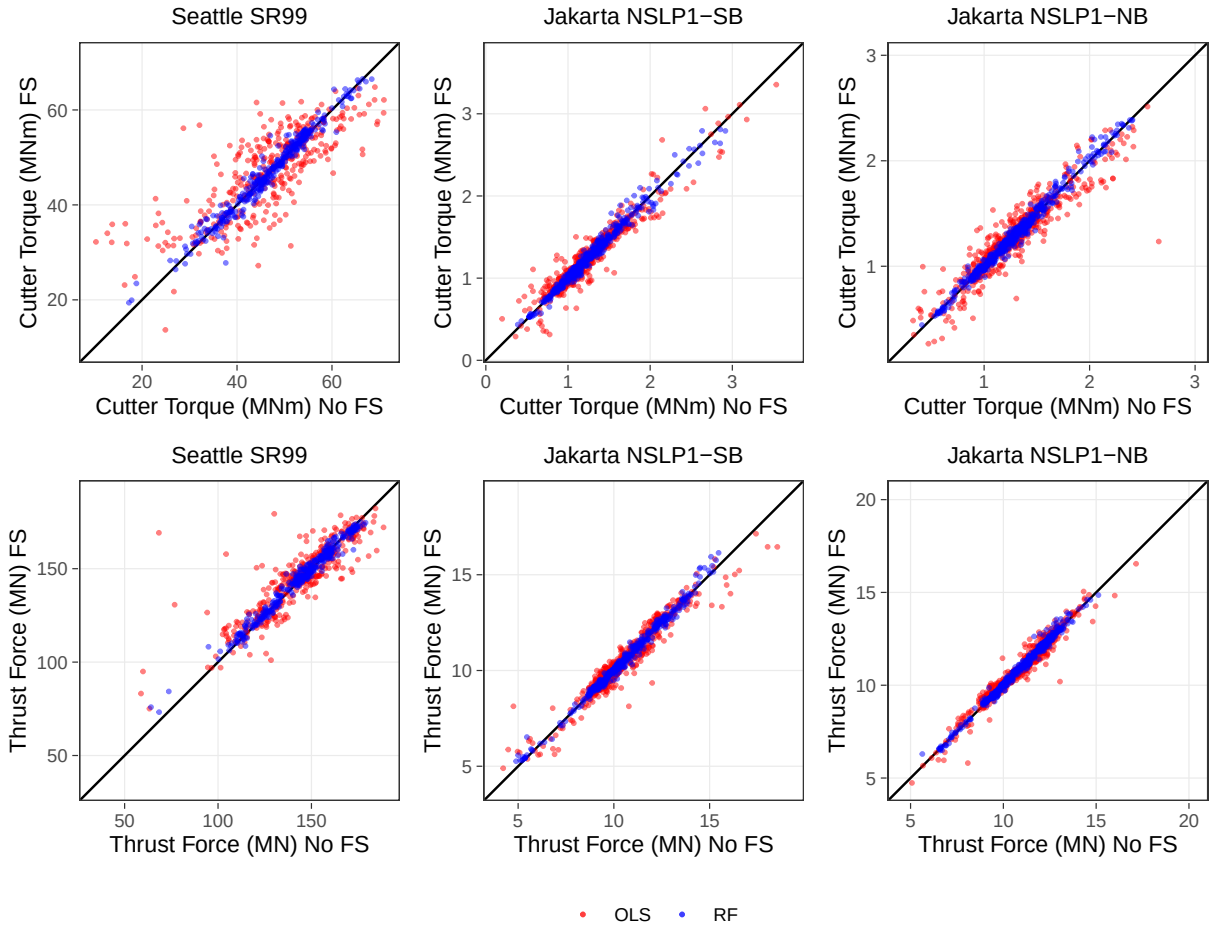

Figure 6.11: Comparison of prediction values produced by models with different prediction methods, feature sets, and tunneling cases.

The RF models reliably produced a high R2 of at least 0.98. This applies to both response features and in all the considered tunneling cases. The OLS models produced slightly lower R2 of at least 0.73 and 0.93 for the SR99 and NSLP1 data sets, respectively. It should

be noted that OLS models developed without the feature selection produced some extreme errors due to the multicollinearity problems, as discussed in Section 6.4.2. Thus, the R2 was computed by excluding those prediction points. This comparison further suggests that the hierarchical feature selection may not substantially affect the prediction values. This also highlights the advantage of machine learning when developing data-driven models using complex and nonlinear data sets such as TBM operation data.

6.4.4 Effects on Model Computability

During tunneling, prediction models should be continuously retrained to adapt to new TBM data generated every few seconds. Thus, computation performance is a critical consideration when developing the models. Figure 6.12 presents the computation time of RF models with different numbers of features. All models were trained using the same computing parameters to allow a comparison of the computation performance, i.e., a single process with 500 trees. Expectedly, the figure shows that higher numbers of features produced more unpredictable computation performance. Reducing the features from 500 to 100 can cut the computation time by more than half. A further reduction to 25 features can cut the computation time by another half. This demonstrates the benefit of using fewer features in a prediction model. This also suggests the advantage of using the hierarchical feature selection for TBM data-driven models, where comparable prediction performance can be obtained with significantly less computation cost.

6.4.5 Effects on Model Stability

Figure 6.13 presents the top five predictor features of cutter torque obtained from models with different prediction methods, feature sets, and tunneling cases. The first row of the figure presents feature ranks produced by RF models developed using the selected feature set. Interestingly, the figure shows similar features in all the tunneling cases. Thrust force, chamber pressure, and shield attitudes (i.e., pitch and yaw angles) were consistently identified as strong predictors of cutter torque. This result is sensible. Cutter torque should be proportional to the thrust force at the same geologic material strength (Ates et al., 2014; Jakobsen et al., 2013). Cutter torque can be associated with the chamber pressure since both should correlate to the strength of the excavated geologic materials (Sramoon et al., 2002; Sugimoto & Sramoon, 2002). Deeper tunneling depth will likely convey higher earth pressure and stiffer geologic materials. Shield attitudes may also be related to cutter torque since it may affect the contact pressure between the cutter face and the excavation face and between the shield friction and the ground (Sramoon et al., 2002; Sugimoto & Sramoon, 2002; Sugimoto et al., 2007).

The figure also shows fairly consistent ranks in all the tunneling cases. Both NSLP1 models produced identical feature ranks, even though these tunnels were excavated using different EPB TBMs and by different operators. Nevertheless, some discrepancies can be observed. The SR99 model identified cutter rotation speed as the top predictor. This is different from

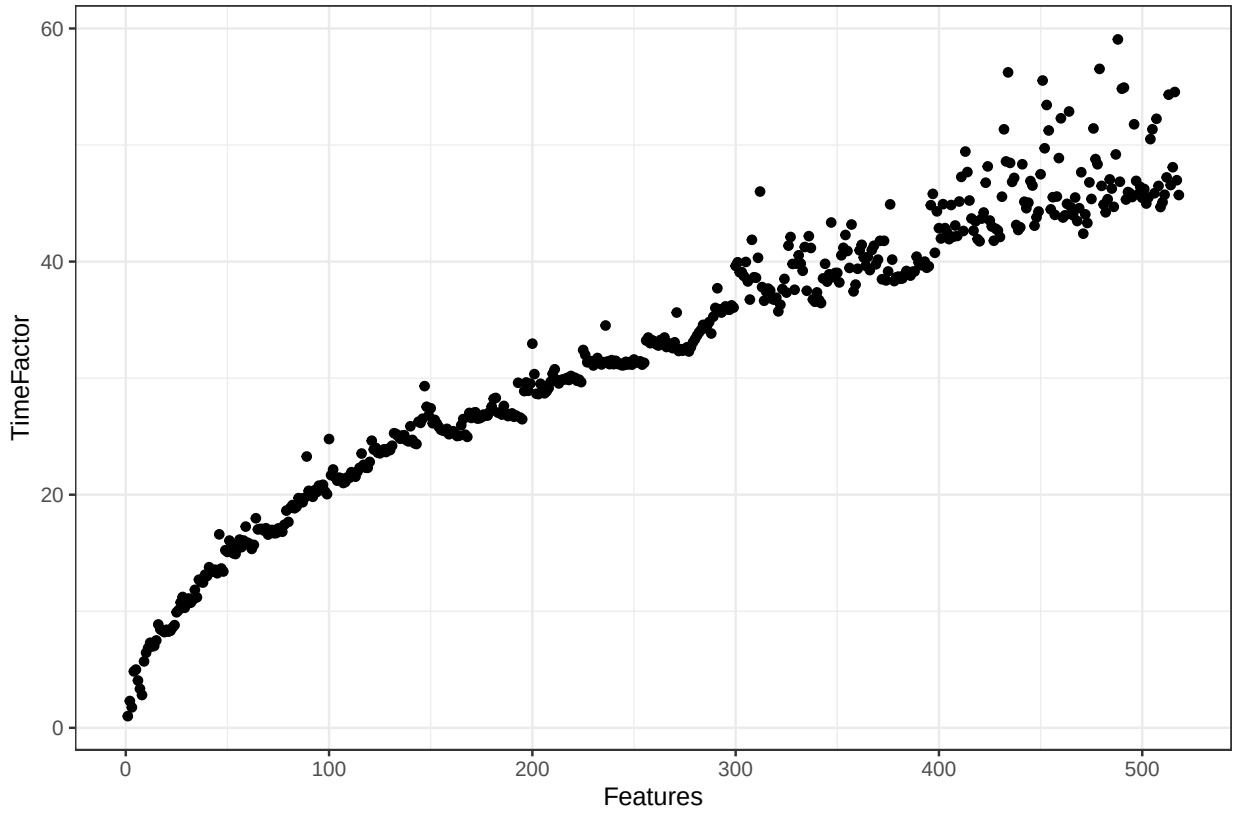

Figure 6.12: Computation performance of an RF model with different numbers of predictor features

the NSLP1 models, where cutter rotation speed was not considered an important predictor. However, it should be noted that the cutter rotation speed was maintained at about 0.9 rpm in both the NSLP1 tunneling cases. In this condition, this feature could not provide any indication of cutter torque and, therefore, cannot be the predictor. In addition, both NSLP1 models identified foam injection pressure as an important predictor. This is sensible since foam is typically injected into the ground to condition the strength characteristics of the excavated geologic materials, which will eventually affect the cutter torque.

The second row of the figure presents the feature ranks of OLS models developed using the selected feature set. These models consistently identified thrust force and chamber pressure as the two strongest predictor features of cutter torque. The models also identified features related to shield attitudes and ground conditioners as strong predictors. However, more inconsistency can be observed in the ranks. For example, the NSLP1-SB model identified the shield attitudes as yaw angle and thrust stroke difference (top-bottom). In contrast, the

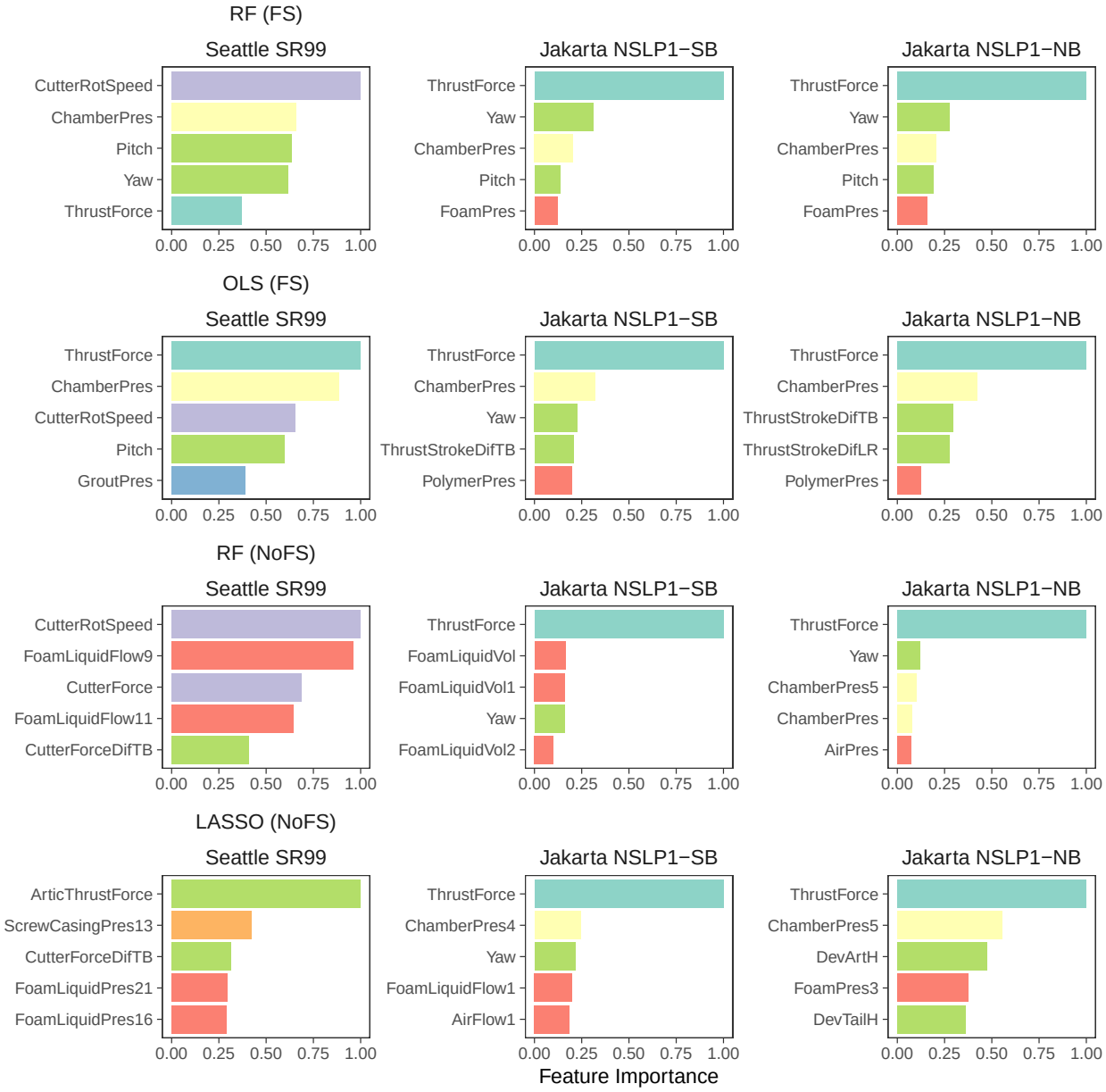

Figure 6.13: Comparison of top predictors identified in cutter torque models with different prediction methods, feature sets, and tunneling cases.

NSLP1-NB model identified the attitudes as the thrust stroke differences (top-bottom and left-right). Furthermore, the SR99 model identified the attitudes as pitch angle only. In addition to the consistency issue, a less sensible feature can also be observed in the result. The SR99 model identified grout pressure as a strong predictor feature of cutter torque. Grouting pressure is applied to the shield's tail and may not directly correlate to the cutter torque at the excavation face.

The consistency in the RF feature ranks indicates the generalizability of the models. This suggests the capability of nonparametric nonlinear machine learning models to obtain some generalization in data sets with complex relationships, such as TBM operation data. This result oppose a widely accepted premise that OLS is more interpretable and generalizable than machine learning (e.g., B. Efron (2020)). Less consistency of the OLS results indicates problems encountered by the parametric linear models to fit complex and nonlinear relationships within the TBM data. This suggests the limitation of parametric linear models for generalizing the complex data sets.

The third and fourth rows of the figure present the feature ranks obtained from the RF (the third row) and the Lasso (the fourth row) models. The models were developed using the feature set without the selection to investigate the performance of embedded machine learning-based feature selection. The Lasso and RF were selected to represent linear and nonlinear embedded feature selection methods, respectively. These figures show that the produced feature importance ranks are highly inconsistent. No patterns can be observed between the two prediction methods and the three tunneling cases.

In many cases, most of the ranks are filled by individual sensor measurements. For example, in the SR99 data set, the Lasso model identified two foam liquid pressure sensors as cutter torque predictors. The RF model also identified two foam liquid flow sensors as predictors. Similarly, both the NSLP1 models identified many individual sensors of foam liquid volume and chamber pressure in the ranks. These are the same problem encountered by Maher (2015). This indicates that the prediction models were strongly biased toward those individual sensor measurements. In this condition, no matter how good a model produces prediction performance, no generalization and no new understanding can be drawn. This suggests the importance of guided feature selection in developing data-driven models using TBM data.

Figure 6.14 presents feature ranks from thrust force prediction models. Overall, the produced ranks lead to similar conclusions to the cutter torque models (Figure 6.13). The figure shows that RF models with the selected feature set identified almost identical predictors in all the tunneling cases, i.e., cutter torque, chamber pressure, shield attitudes (i.e., pitch and yaw angles), and ground conditioner. The difference is only in the type of ground conditioner, i.e., the SR99 model identified polymer pressure, while the NSLP1 models identified foam pressure. This might be due to the different geologic conditions of the tunneling cases. The NSLP1 site in Jakarta is dominated by clay and silt. In comparison, the SR99 site in Seattle is dominated by glacial deposits, which vary from clay and silty soils to sandy and gravelly soils. More polymer is required to condition sandy soils (A. Merritt et al., 2023). This might be the reason why the polymer is a strong predictor in the SR99 model.

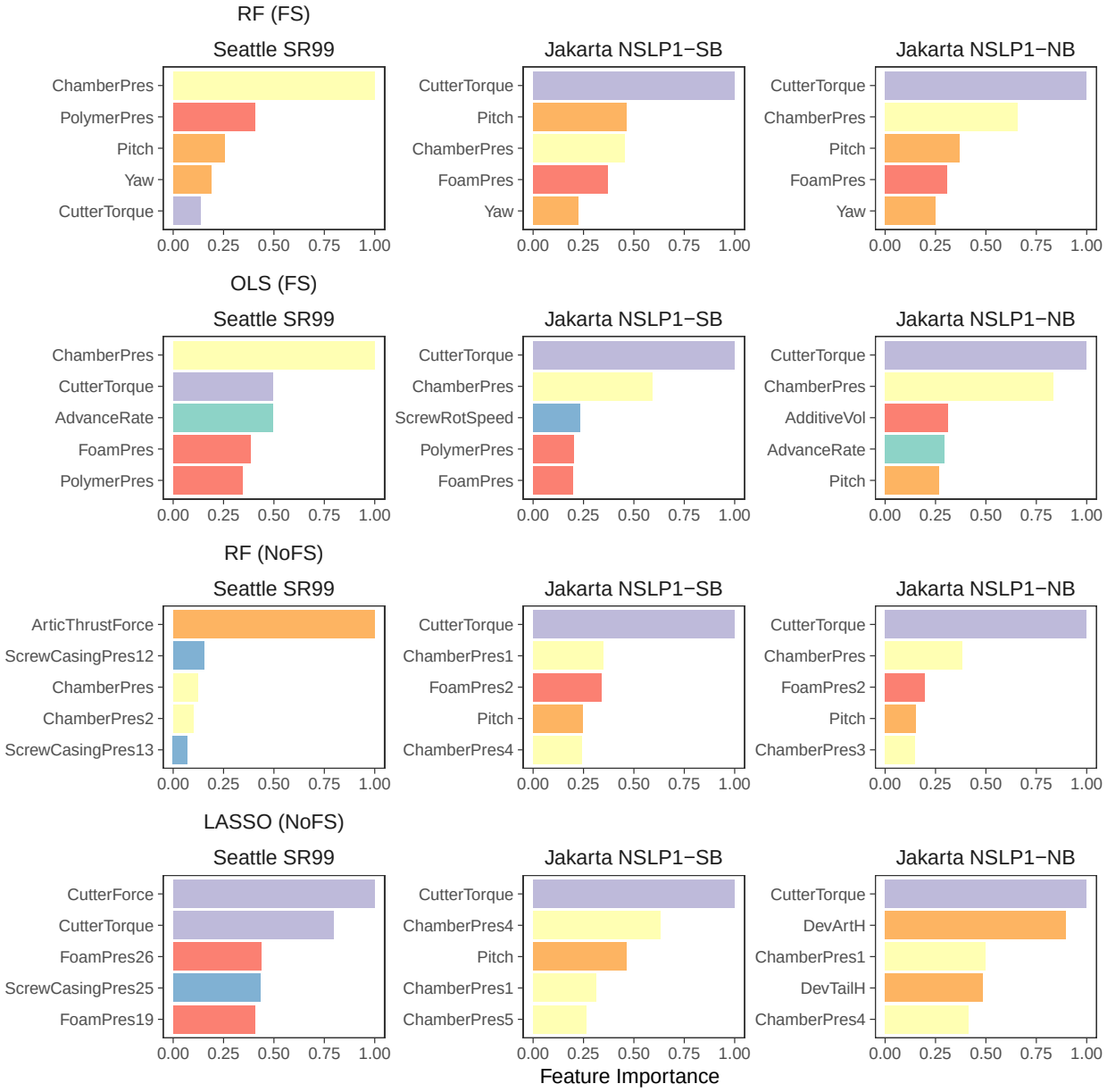

Figure 6.14: Comparison of top predictors identified in thrust force models with different prediction methods, feature sets, and tunneling cases.

The OLS model with the selected feature set also produced similar features, but with more inconsistencies. For example, all cases consistently identified cutter torque, chamber pressure, and features related to ground conditioners as important predictors to thrust force. However, there are variations in the presence of advance rate (in the SR99 and NSLP1-NB), screw rotation speed (in the NSLP1-SB only), and shield pitch angle (in the NSLP1-NB only). Similar to the cutter torque, thrust force models without the feature selection identified predictors that are biased toward individual sensor measurements. The ranks are highly inconsistent and difficult to generalize. This result provides more evidence of the importance of knowledge-guided feature selection in developing stable and generalizable prediction models.

6.5 Conclusions

This study has quantitatively validated the presence of multicollinearity in TBM data. This presence may result in multicollinearity-related problems, such as response outliers and high leverage. This highlights the importance of TBM feature selection when developing data-driven models. This study has also proposed a general taxonomy of EPB TBM features. The taxonomy was developed based on four levels of hierarchy, i.e., (i) the process level, (ii) the component level, (iii) the unit measurement level, and (iv) the individual sensor level. The hierarchy allows the domain knowledge guides the feature selection process. Effects of the proposed hierarchical feature selection on prediction models have been investigated in the framework of the model's predictability, computability, and stability (generalizability). The main findings of this investigation are summarized in the following.

- **Effects on model predictability and computability**. Prediction models developed using the selected feature set should be more robust to severe multicollinearity problems. Furthermore, the models can produce comparable prediction performance to models developed using the original features with substantially less computation cost. Computation performance is critical when developing data-driven tunneling models since the models should be continuously retrained to adapt to newly generated TBM data.

- **Effects on model generalizability**. These results suggest that prediction models developed using the proposed hierarchical feature selection can produce consistent features and importance ranks in different tunneling cases, indicating better model generalizability. Furthermore, machine learning appears to offer better model generalization than parametric linear methods when using data sets with complex and nonlinear relationships such as TBM data. Developing generalizable models is essential for scientific discovery in tunneling, where most data come from observations, not controlled experiments.

Further studies may include expanding the TBM variable features and improving the reliability of the feature importance analysis. The original TBM data sets include numerous

discrete variable features that indicate more operator control decisions. However, the scope of this study was limited to numerical features from continuous measurements only. More studies can be performed to combine the categorical and numerical features and investigate their effects on the prediction models.

This study has demonstrated encouraging results on the potential of using feature importance analysis to assess model generalizability. This reveals that different tunneling cases with significantly different geologic conditions, EPB TBMs, and operators can produce consistent features and ranks, which indicate the generalizability of the models. However, as discussed in Chapter 5, selecting the best method for determining the feature importance rank is still debatable (Altmann et al., 2010; Hooker et al., 2021; Strobl et al., 2007). Further investigation using different feature importance analysis methods should be performed to obtain more solid conclusions.

Chapter 7

Interpreting Geologic Conditions during Tunneling

Some contents of this chapter have been published in Apoji, D., Fujita, Y., and Soga, K. (2022). *Soil Classification and Feature Importance of EPBM Data Using Random Forests*. Geo-Congress. American Society of Civil Engineers. 520–528.
https://doi.org/10.1061/9780784484029.052. (With permission from ASCE)

7.1 Introduction

7.1.1 Background

The geologic conditions along a tunnel alignment significantly govern the tunneling performance and risk profile. From the dynamical systems perspective, the ground condition is the primary external stimulation (disturbance) of a tunneling process. Changing ground conditions will change the system's dynamics and eventually alter the tunneling outputs, i.e., tunnel boring machine (TBM) behaviors and performance. In this condition, the operators must adjust the control parameters to maintain the target tunneling performance. Therefore, TBM operators need to be continuously aware of the changing ground conditions during tunneling.

Detecting the changing ground conditions during tunneling has always been one of the most challenging tasks in tunneling operations. The tunnel geologic maps can provide a guide. This map is prepared based on available geotechnical data obtained in the design stage. However, it should be noted that the map is interpreted from a limited number of boreholes at discrete locations. The sampled area (i.e., boreholes) may only cover less than one percent of the total tunneling area. Furthermore, geologic conditions always have inherent variability and uncertainty, especially since tunnel alignments can span several kilometers. Therefore, the map cannot accurately capture the variability of ground conditions at a detailed level.

Also, this variability cannot be observed directly during TBM driving. Unlike driving a car, TBM operators cannot see outside. Thus, they cannot see the ground conditions at the excavation face. They can see the muck (excavated soil materials) discharged to the belt conveyor. However, the muck cannot fully represent the ground conditions at the excavation face. The muck has spent some time in the excavation chamber. Thus, there is a delay time between when the soil is excavated and discharged to the conveyor belt. Furthermore, the properties of the excavated soils have been altered by injections of ground conditioners and the mixing process.

Practically, TBM operators can only infer the ground conditions based on the operation data (Garcia et al., 2021). TBM operation data contains various information on the tunneling operation, such as the status and performance of the cutter, thrust, excavation chamber, ground conditioning system, tail grouting system, and muck discharged system. This data is typically generated every few seconds by numerous sensors in the TBM. Interpreting numerous amounts of data in streams can be challenging for humans. Furthermore, manual interpretation can be unsystematic and subjective and may lead to human bias and errors. Different operators may have different ways of interpreting the ground conditions. This problem leads to the development of tunneling geologic information prediction systems, which are intended to provide systematic and accurate ground interpretation during tunneling.

7.1.2 Related Works

Statistical methods. In early development, Yamamoto et al. (2003) proposed a geostatistical method to infer the ahead geologic conditions during tunneling. They used geostatistical techniques to transform bore logs and TBM operation data into geologic information with three-dimensional (3D) temporal and spatial distributions. In principle, the method used correlation coefficients to convert the drill energy coefficient into excavating energy and rock mass strength. The spatial distribution of these parameters was modeled using the mean and covariance functions. Finally, the distribution of these parameters ahead of the TBM face was estimated using the Kriging method. This is an early implementation of data-driven methods to develop tunneling geologic information prediction systems.

Sousa and Einstein (2012) proposed a probabilistic method to develop a geologic prediction system during tunneling. The proposed model used Bayesian Networks (BN), a probabilistic graph model, to transform TBM operation data into the ahead geologic conditions. They examined 31 different BN models by considering various configurations with various earth pressure balance (EPB) TBM features as the variables, i.e., the penetration rate, cutter torque, cutter force, thrust force, and muck weight (on the conveyor belt). The training data set consisted of 720 rings chosen randomly from a section in the Porto Metro tunnel. The data observation data points were in ring aggregation. Each ring is labeled by its geologic category, i.e., soil, rock, or mixed ground conditions. Sensitivity analyses were performed to determine the most suitable model. They found that the best model was built using the penetration rate, cutter torque, and cutter force.

Machine learning methods. The use of AI and machine learning techniques to tackle this problem has emerged in the past few years. Q. Zhang et al. (2019) proposed a method to transform TBM operation data to rock mass types using machine learning. They examined three machine learning algorithms, i.e., Support Vector Machine (SVM), Random Forests (RF), and k-Nearest Neighbors (kNN). The model was developed using four TBM features as the input, i.e., the advance rate, cutter torque, cutter rotation speed, and thrust force. The data was obtained from a tunnel project in Jilin, China. The original data contained 12 million observation data points and was compressed to 5014 leaf node entries using the balanced iterative reducing and clustering using hierarchies algorithm (BIRCH). Each observation data point was labeled with the rock mass type, which was categorized into five types of potential rock mass using K-means++, an unsupervised learning algorithm. The data were sequentially split into training and testing data sets with the proportion of (i) 20:80, (ii) 50:50, and (iii) 80:20, respectively. Cross-validation was performed in 10-folds to tune the hyperparameters of every machine learning model and evaluate the performance. They concluded that the prediction models could produce reasonably good results even with limited training data. The best prediction performance was produced by the SVM model.

Zhao et al. (2019) evaluated the performance of various machine learning algorithms to classify geologic types. They developed geologic prediction models using eight machine learning algorithms, i.e., kNN, Bayesian Linear Regression (BLR), SVM, Decision Tree (DT), RF, CatBoost, Extreme Gradient Boosting (XGBoost), and Artificial Neural Network (ANN). The original TBM operation data consisted of 72 TBM features, such as cutter torque, thrust force, advance rate, and fuel tank temperature. Unlike the previous studies, Zhao et al. did not directly use TBM features as the input. Instead, they augmented the features using Principal Component Analysis (PCA) and retained 95% of the variance information. The original data contained 4.6 million observation data points. However, only 88 rings (from 1364 rings) had the geologic information. The data were labeled with six geologic units, which were categorized according to the similarity of their properties. The data were split into training and testing data sets with a proportion of 70:30, respectively. Unfortunately, no information if the data were split randomly or sequentially.

They concluded that, in most cases, dimension reduction could improve the performance of the models. However, it should be noted that dimension reduction (i.e., using PCA) may reduce the model interpretability as the original features are transformed from variables that represent physical measurements to scores that mix the original features. Furthermore, they also concluded that the ANN model with a specific architecture could outperform the other models. However, this result should be cautiously taken. It appears that the other machine learning models were not tuned, as no information was provided on hyperparameter tuning and cross-validation in the modeling process. It should be noted that the performance of most machine learning models can be improved by using properly tuned hyperparameters (Du et al., 2021; Schratz et al., 2019; L. Yang & Shami, 2020).

Erharter et al. (2019) compared the performance of neural network-based models in transforming TBM data to rock mass types. They examined two types of ANN algorithms, i.e., Multilayer Perceptron (MLP) and Long-Short-Term Memory (LSTM). The models were

developed using TBM features as the input, i.e., advance rate, penetration rate, cutter rotation speed, cutter torque, thrust pressure, thrust force, crown-support-cylinder pressures (left and right), and path of crown-roof-support-cylinder. Some calculated variables were also added to the feature set, i.e., the rotary share of the specific energy and the torque ratio. The data was obtained from The Brenner Base Tunnel, a railway tunnel between Austria and Italy. The observation data point interval of the original data was 10 seconds. These observation data points were aggregated into 4 cm spatial points using linear interpolation. The data was labeled with four geologic classes representing the rock mass behaviors. The data was split into sections to obtain the training and testing data sets. The tuning of the ANN architectures and hyperparameters was done by trial and error. They found that both models showed satisfying accuracies, with the LSTM model outperforming the MLP more consistently.

Real-time interpretation. Erharter et al. (2020) extended the development into a real-time (or online) rock classification system during tunneling. They used an ensemble LSTM model to classify TBM data into rock behavior types. The data set was the same as the previous study (Erharter et al., 2019). The ensemble method was used to increase the prediction stability due to the inherent stochastic behavior of ANNs. The model was trained and tested in a continuous loop to develop the real-time rock classification system. The training batch was performed every 400 m TBM advance when new data streams were available. The concluded that the ensemble method effectively improved the overall accuracy and prediction stability. It was also mentioned that changing the learning rate may have substantial effects on the overall accuracy of the model.

H. Yu (2021) investigated methods to characterize the encountered ground condition based on EPB TBM data during tunneling by employing two machine learning frameworks, i.e., (i) supervised learning using multinomial logistic regression and (ii) semi-supervised learning using a similarity matrix and label propagation method. The models were developed using eight EPB TBM features as the input, i.e., advance rate, cutter rotation speed, cutter torque, chamber pressure, chamber's vertical gradient, screw conveyor rotation speed, screw conveyor torque, and belt conveyor muck weight. The data was obtained from the Northgate Link Extension tunneling project in Seattle, Washington. The observation data points were aggregated in ring size. For each ring, only the stable state of the excavation advancement phase was considered. The data was labeled with the fractional representation of four encountered soil types according to 83 available boreholes. The models were trained sequentially by adding more boreholes to simulate the tunneling process. They concluded that both models performed similarly, given sufficient training data. However, the semi-supervised learning model performed substantially better in limited training data.

7.1.3 Objectives

The review demonstrates the possibility of systematically and quantitatively interpreting the geologic conditions based on TBM data. Nevertheless, it also reveals some gaps, as listed below.

- Most of the models performed categorical classification. This may not be ideal as the actual ground conditions may not have a clear boundary between geologic types. Furthermore, the categorical classification system does not allow for identifying the gradual transition changing of soils which is prevalent in nature. In addition, most of the previous studies were done for rock classification. Limited studies have been performed to develop soil interpretation models.

- Most of the studies used limited TBM features. While it is true that increasing numbers of (unrelated) features may increase the risk of overfitting, selecting only limited features, especially in a complex dynamical system such as tunneling, may not be able to capture all of the actual processes (i.e., underfit the process). For example, most of the studies included no features related to the ground conditioning system. It is evident that the injected ground conditioners affect other variables such as advance rate, cutter torque, and screw conveyor torque. Omitting these features means neglecting these critical interactions.

- Most of the previous studies focused on prediction performance. Fewer efforts have been made to the model interpretability of the model and to generalize the important features that contain the relevant geologic information. This is important to provide a better generalization to the conclusion, especially since every previous study used different tunneling data sets due to the unavailability of openly accessible tunneling data.

This study aims to (i) propose a supervised artificial intelligence (AI) system that is capable of interpreting the encountered geologic transition based on TBM operation data in real-time during tunneling, and (ii) investigate TBM features that contain most of the geologic information. The geologic interpretation model was developed a supervised learning algorithm and used 36 TBM features from different tunneling partial processes as the input predictors. The importance of each predictor in the model was investigated using three feature importance analysis methods.

7.2 Data

7.2.1 Tunneling Case

This study used a data set of the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA. More discussion of this tunneling case can be found in Chapter 5. This data set was considered suitable for the geologic interpretation system simulation since the tunneling was conducted in various geologic deposits with substantial ground transitions.

7.2.2 Geologic Conditions

Geologic conditions along the SR99 tunnel were obtained from the geotechnical baseline project reports (WSDOT, 2010a, 2010b, 2010c). Based on the reports, the geology of the site consisted of younger Holocene deposits and older Vashon and Pre-Vashon (Pleistocene) deposits. These geologic deposits may come from different geologic origins and processes (e.g., glacial and non-glacial units). However, for the tunneling purpose, the deposits were classified into engineering soil units (ESU) based on their physical characteristics. At the ground surface and shallow depth, the geologic conditions were dominated by fill and younger soil deposits, i.e.,

- Engineered and Non-engineered fill (ENF),
- Recent Granular Deposits (RGD), and
- Recent Clay and Silt (RCS).

The tunnel was primarily constructed below the fill deposits, where the geologic conditions were dominated by overconsolidated glacial and non-glacial pre-Vashon geologic units, i.e.,

- Till Deposits (TD),
- Cohesionless Sand and Gravel (CSG),
- Cohesionless Silt and Fine Sand (CSF),
- Cohesive Clay and Silt (CCS), and
- Till-Like Diamict (TLD).

In terms of their engineering behaviors, CCS is cohesive, while CSG and CSF are cohesionless. TD may have a more complex behavior since it is a cohesive mixture of gravel, sand, silt, and clay. TLD is generally cohesionless but may have layers and lenses of tills (2010c). The interpretation of the geologic conditions along the tunnel can be seen on the geologic map as presented in Figure 7.1.

Supervised learning algorithms require labels as the ground truth. This study considered boreholes to be the ground truth and the specified ESUs of each borehole within the tunnel diameter as the labels. The boreholes were envisaged to represent geologic conditions for a 50-ft (about 15.2 meters) radius from the drilled location. Multiple ESU might be encountered at the tunneling face. The labeling was done depending on the analysis approaches, i.e., classification or regression analysis.

- **Classification**. In the classification approach, the labels were represented as five borehole categories, characterizing the mixtures of ESU at each borehole location. Five labels were created based on the borehole categories (Table 7.1).

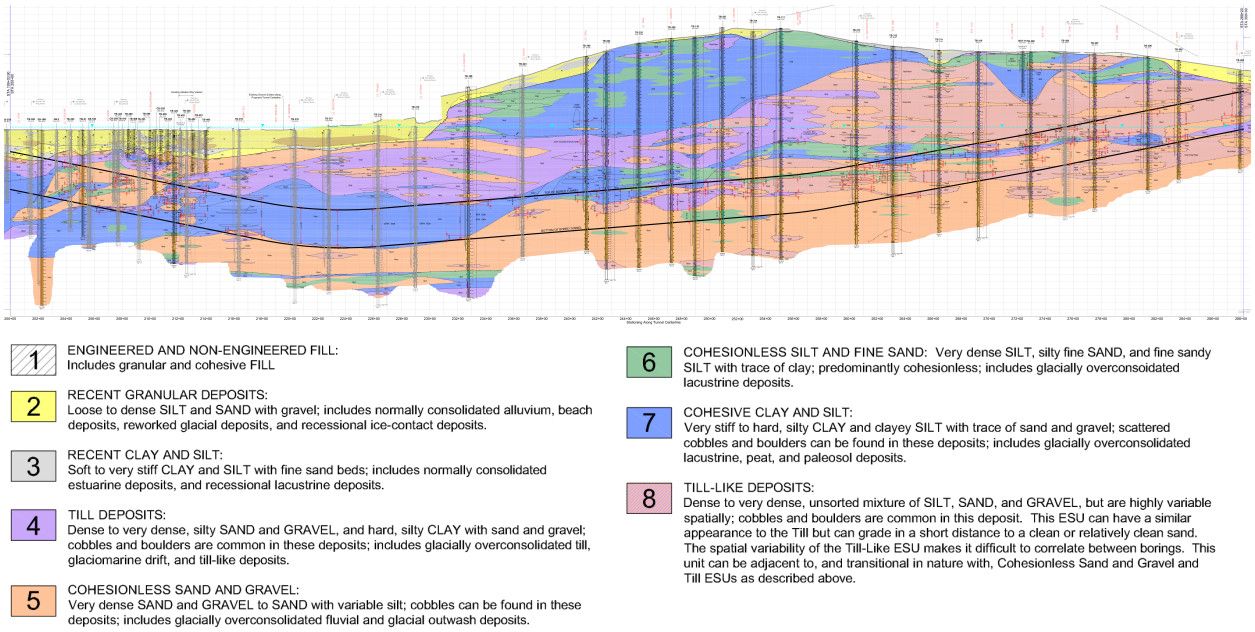

Figure 7.1: Geologic map of the SR99 tunnel (WSDOT, 2010b, 2010c).

- **Regression**. In the regression approach, the labels were represented by the percentage of every ESU within the tunnel diameter (Table 7.2). This means each borehole contained the percentage of multiple ESUs. The percentage of each ESU varied from 0 (the ESU did not exist in the borehole) to 100% (the only ESU within the tunnel diameter of the borehole).

7.2.3 Data Description and Preparation

Similar to Chapter 5, this study used ring-length aggregate spatial series as the observation data points. Also, the study was focused on the continuous data of the primary EPB TBM system, i.e., main features in the tunneling processes. To avoid bias toward spatial locations, features related to TBM steering and attitudes were not included (i.e., pitch and yaw angles of the shields, earth pressures in the chamber). To further condense the number of features and enable more interpretability of the analysis results, records from same sensor types were summed together. This produced a total of 36 features to be used in the analysis (Table 7.3).

The data was prepared according to the following steps, (i) removal of non-excavation observation data points, (ii) aggregating features from the same sensor types and volume

Table 7.1: Borehole labels for the classification model.

BH	Chainage	Borehole Label
TB326	20150	CCS_TLTD
EB18B	20250	CCS_TLTD
TB325	20425	CSGCSF_CCS
TW21	20525	CSGCSF_CCS
CB108	20575	CSGCSF_CCS
TB324	20775	CSGCSF_CCS
TB323	20925	CSGCSF_CCS
TB322	21175	CCS_CSG
TB320	21275	CCS_CSG
TB319	21625	CCS_CSG
TB318	22025	CCS_CSG
TB317	22275	CCS_CSG
TB316	22625	CCS
TB315	22900	CCS
TB108	23275	CCS_CSG
TB223	23675	CSGCSF_CCS
TB109	24125	CSGCSF_CCS
TB224	24275	TLTD_CSGCSF_CCS
TB225	24500	CSGCSF_CCS
TB226	24725	TLTD_CSGCSF_CCS
TB110	24900	TLTD_CSGCSF_CCS
TB227	25100	CSGCSF_TLTD
TB111	25500	CSGCSF_TLTD
TB312	26310	TLTD_CSGCSF_CCS
TB311	26650	TLTD_CSGCSF_CCS
TB310	26925	TLTD_CSGCSF_CCS
TB309	27300	TLTD_CSGCSF_CCS
TB308	27550	TLTD_CSGCSF_CCS
TB307	27750	CSGCSF_TLTD
TB305	28350	CSGCSF_TLTD
TB304	28790	CSGCSF_TLTD

Table 7.2: Soil labels for the regression model.

BH	Chainage	TD	CSG	CSF	CCS	TLD
TB325	20425	7.7	47.9	26.7	17.7	0
TW21	20525	0	89.4	0	10.6	0
CB108	20575	0	63.7	15.4	20.9	0
TB324	20775	4.6	61.3	16	12.8	5.3
TB323	20925	15.6	68.8	4.6	11	0
TB322	21175	15	22.3	0	55.9	6.8
TB320	21275	-8.9	23.3	5.1	40.2	40.2
TB319	21625	7.1	6.9	0	85.9	0
TB318	22025	0	4.2	0	95.8	0
TB317	22275	0	0	0	96	4
TB316	22625	0	0	0	100	0
TB315	22900	0	0	0	100	0
TB108	23275	0	20	0	80	0
TB223	23675	6.4	54.7	0	27.9	11
TB109	24125	0	78.1	10.6	11.3	0
TB224	24275	0	38.1	0	35.9	25.9
TB225	24500	0	69.7	8.8	15.9	5.5
TB226	24725	44	50.6	0	5.4	0
TB110	24900	36.7	16.7	26.6	20	0
TB227	25100	24.8	51.4	23.8	0	0
TB314	25310	0	64.1	10.4	0	25.6
TB111	25500	10	48.5	6.2	0	35.4
TB313	26050	0	66.2	8.8	0	25
TB312	26310	0	47.3	19.4	14.6	18.6
TB311	26650	0	44.2	12.1	12.7	31
TB310	26925	13.8	43.8	13.6	14.8	13.8
TB309	27300	14.4	57.3	3.6	5.4	19.3
TB308	27550	0	61.3	8.7	11	19
TB307	27750	3.6	53	0	0	43.3
TB306	28125	0	7.7	2.5	31.2	58.6
TB305	28350	2.5	61.4	0	0	36.1
TB304	28790	16.3	18.9	0	0	64.8

Table 7.3: List of EPB TBM features used in the supervised geologic interpretation model.

No.	Feature	Unit	Sub-Component	Group
1	Cutter torque	kNm		Excavation
2	Cutter head force	kN		
3	Cutter rotation speed	rpm		
4	Thrust force	kN		Advancing
5	Thrust stroke	mm		
6	Advance rate	mm/min		
7-8	Shield roll	deg	Front, rear	Attitudes
9	Foam agent volume	m^3/m		Ground conditioning
10	Foam agent flow	m^3/min		
11	Foam liquid volume	m^3/m		
12	Foam liquid flow	m^3/min		
13	Foam liquid pressure	kPa		
14	Foam volume	m^3/m		
15	Foam flow	m^3/min		
16	Foam pressure	kPa		
17	Polymer agent volume	m^3/m		
18	Polymer agent flow	m^3/min		
19	Polymer volume	m^3/m		
20	Polymer flow	m^3/min		
21	Polymer pressure	kPa		
22	Chamber bulk density	kPa		Earth pressure balancing
23-24	Screw rotation speed	rpm	#1, #2	Muck discharging
25-28	Screw pressure	kPa	#1A, #1B, #2A, #2B	
29	Screw torque	kNm		
30	Screw muck volume	m^3/m		
31-32	Belt muck volume	m^3/m	Front, rear	
33-34	Belt muck weight	Ton	Front, rear	
35	Grout volume	m^3/m		Backfill grouting
36	Grout pressure	kPa		

normalization, (iii) removal of features that might contain biased information toward TBM spatial location, (iv) removal of erroneous records, and (v) labeling observation data points located adjacent to available boreholes. More discussion of some of these steps can be found in Chapter 5. Some examples of the data with the soil labels (classification) are shown in Figure 7.2.

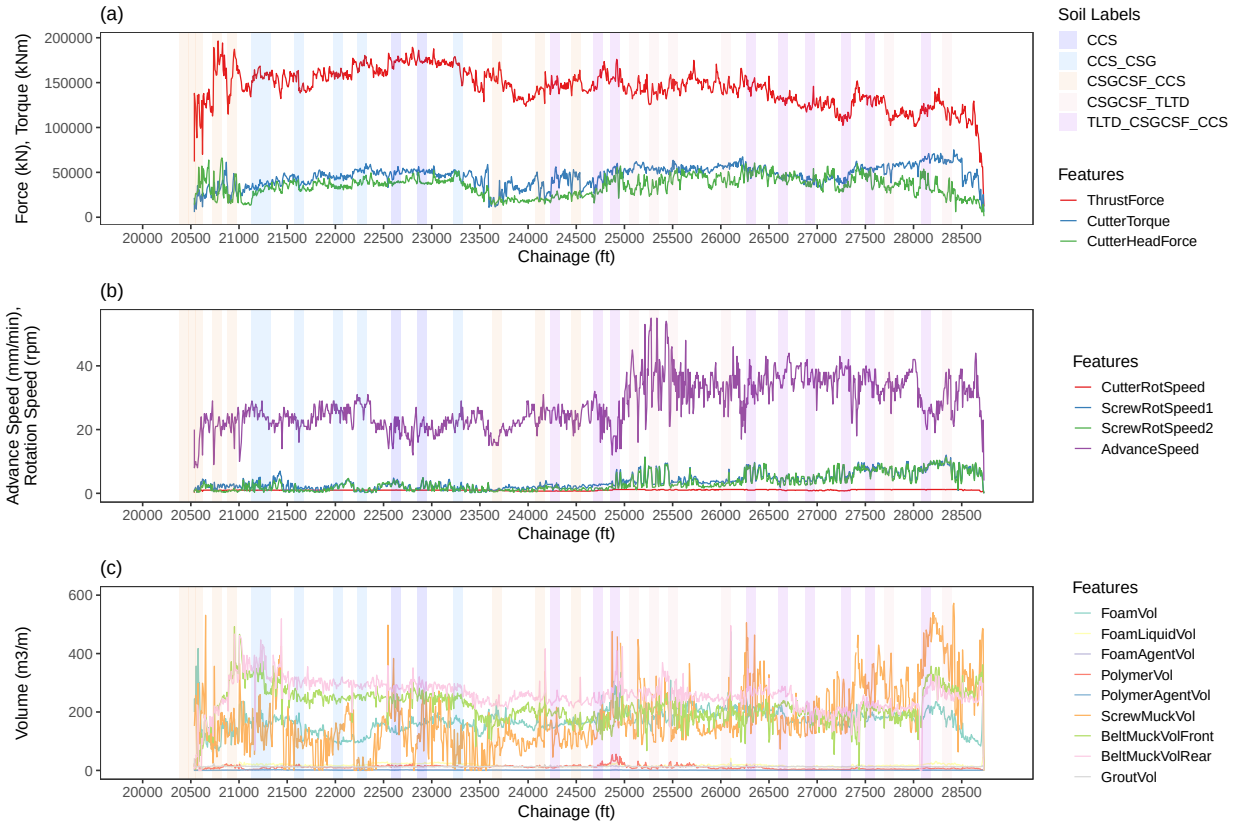

Figure 7.2: Examples of EPB TBM operation data along the tunnel alignment: features related to (a) forces, (b) speeds, (c) volumes. Background colors represent the soil (classification) labels.

7.3 Methods

The proposed geologic interpretation system was developed under the hypothesis that TBM operation data contain information related to the encountered ground conditions. TBM can be analogous to cone penetration tests (CPT), a geotechnical investigation method that has been widely used to characterize soil stratification at a site. Both TBM and CPT

penetrate the ground and produce a set of continuous data. However, CPT only returns a few variable features, such as the tip and friction resistances. Furthermore, the data is produced in a controlled procedure, e.g., a fairly uniform penetration mechanism and rate. Therefore, the data can be interpreted using a simplified method. This is not the case in TBM, where the data are high-dimensional, massive, and produced from complex interactions of ground conditions, TBM behaviors, and operator decisions. Machine learning was expected to be capable of exploiting patterns within the data and transforming them into geologic information.

7.3.1 Model Setup

The geologic interpretation system was developed using two prediction approaches, i.e., classification and regression. In the classification-based prediction approach, the model was used to transform TBM operation data into the probabilities of the borehole categories (Figure 7.3). In contrast, in the regression-based prediction approach, the model was used to transform TBM operation data into the percentage of every specified ESUs. The multiple outputs were obtained by developing multiple prediction models simultaneously, i.e., a single model for each ESU. Subsequently, all of the single model's prediction results were proportioned linearly so that the total percentage of all ESUs is 100% (Figure 7.4).

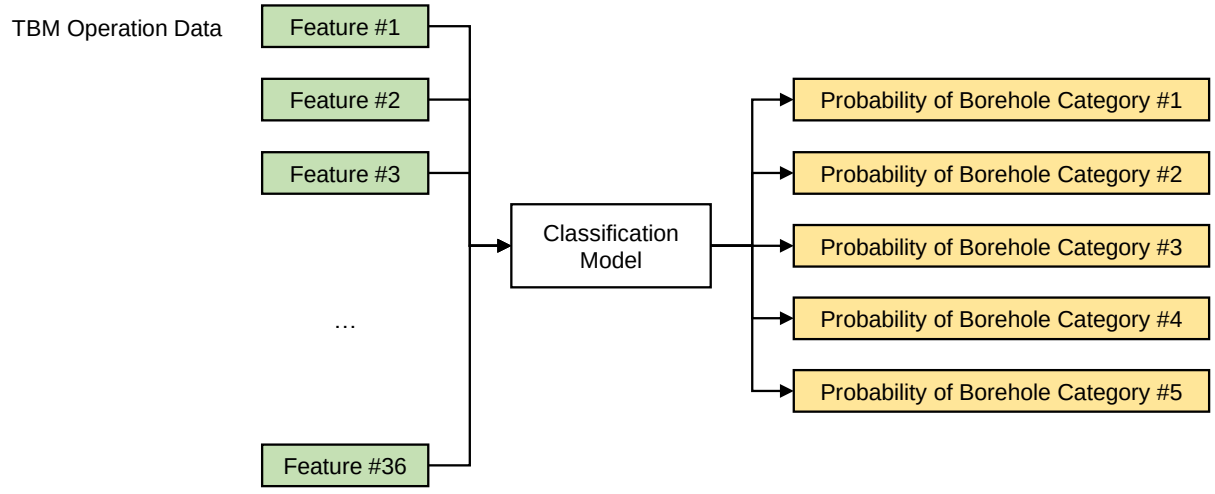

Figure 7.3: Schematic diagram of the classification-based prediction approach.

Three models were developed from these approaches depending on the training-testing scheme and the model output.

(i) **Static Random Classification Model**. The static random classification model was developed as a benchmark model. In this model, the labeled data was randomly split

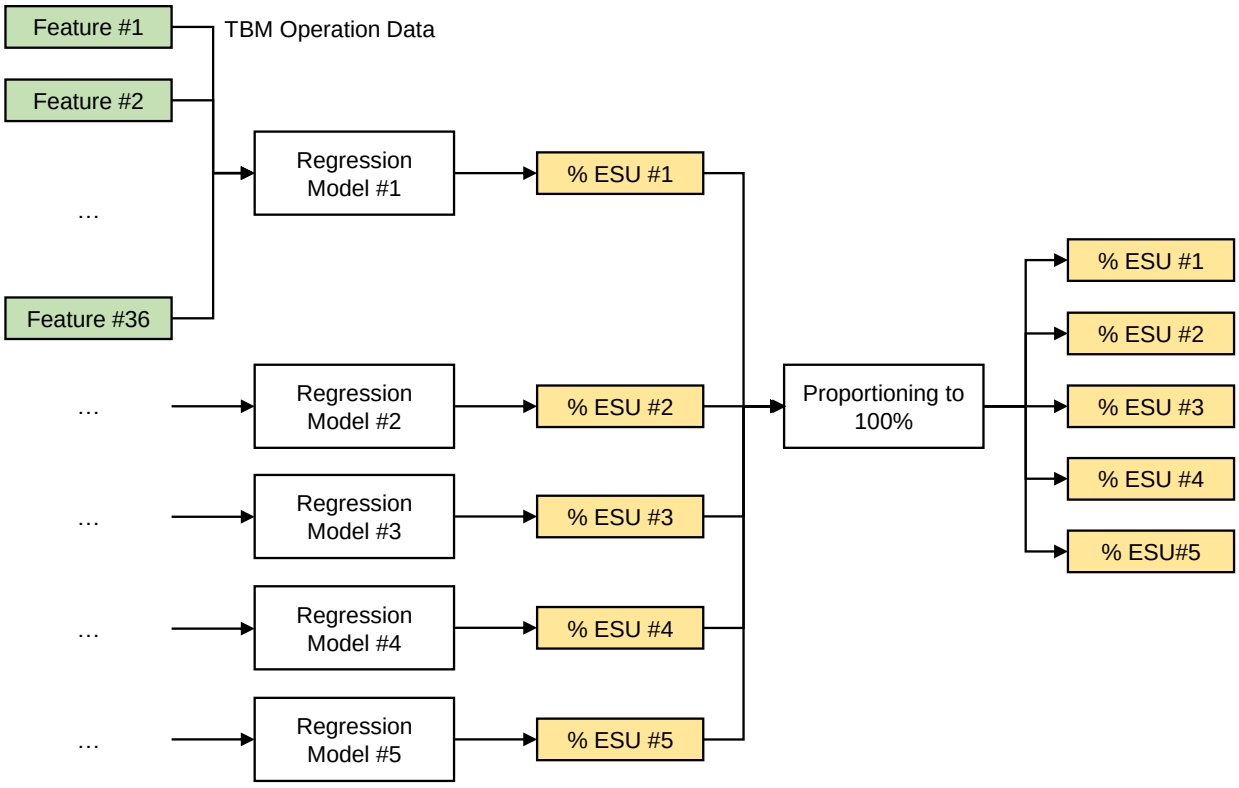

Figure 7.4: Schematic diagram of the regression-based prediction approach.

into training and testing data sets with a ratio of 80:20. This scheme was developed to follow standard practice in machine learning analysis. The randomly selected data ensures minimum bias in the training and testing data to produce the best classification results. However, this scheme would not be applicable in real tunneling cases as the TBM data is generated sequentially. The training data was up-sampled due to the imbalance class distribution (Sun et al., 2009). Besides the labeled data, the model was also used to classify the unlabeled data. A schematic diagram of this model is shown in Figure 7.5.

(ii) **Dynamic Sequential Classification Model**. The dynamic sequential classification model was developed to follow the actual tunneling process. In this model, the classifications were performed sequentially at each observation data point (testing data) based on past observations (training data). In other words, the model was retrained using observations from previous chainages to classify the encountered soils at the current chainage. The training data was also up-sampled at every retraining phase. This

scheme might contain more bias in the training and testing data but can be implemented in real tunneling cases. A schematic diagram of this model is shown in Figure 7.6.

(iii) **Dynamic Sequential Regression Model**. The dynamic sequential regression model was developed as another scheme that can be implemented in real tunneling cases. The training and prediction (testing) schemes are the same as the dynamic sequential classification model, i.e., continuous retraining using the previous observations to predict the encountered soils at the current chainage. Different from the classification models, no up-sampling was performed on the regression model.

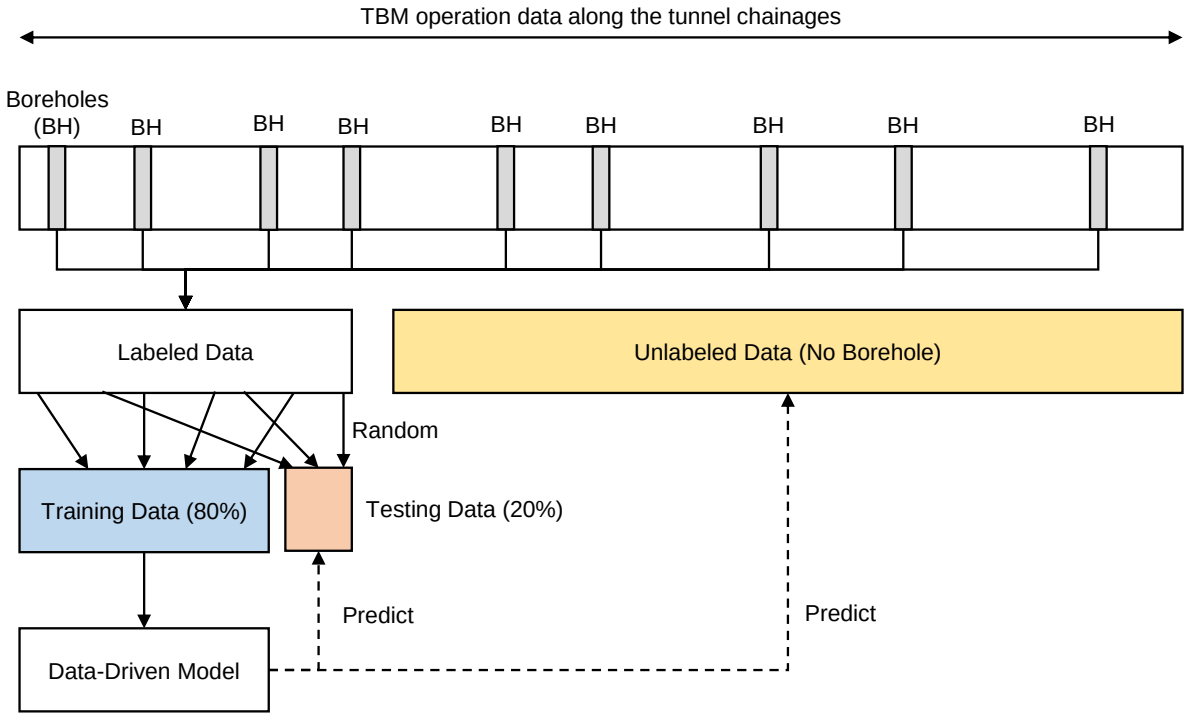

Figure 7.5: Schematic diagram of the static random model.

7.3.2 Prediction Methods

The classification and regression models were developed using Random Forests (RF), an ensemble-supervised learning algorithm that produces the predictions by aggregating a large

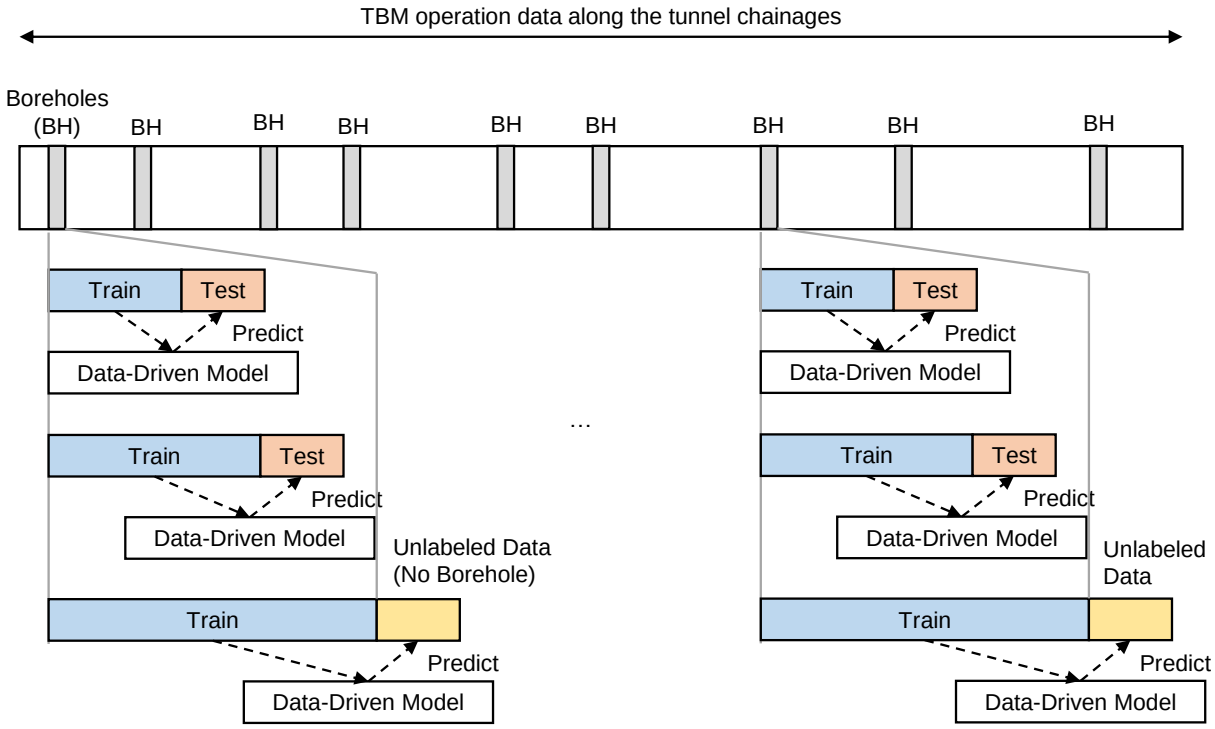

Figure 7.6: Schematic diagram of the dynamic sequential model.

number of decision trees (Breiman, 2001). More discussion and mathematical background of this algorithm can be found in Chapter 4. This study used the RF fast implementation in C++ and R (Wright & Ziegler, 2017).

7.3.3 Model Hyperparameter

Hyperparameter analysis was performed on the static random classification model to investigate the effects of the hyperparameters on the model performance. The analysis was also performed to select the best hyperparameter configuration. Similar to Chapter 6, five times repeated 10-fold cross validation (CV) was conducted in various ranges of RF hyperparameters, i.e., (i) ntrees, (ii) mtry, (iii) minimum node size, and (iv) split rule. More discussion on the ntrees, mtry, and minimum node size can be found in Chapter 6.

The split rule defines how a single decision tree is to be constructed. In this study, two split rule methods were evaluated, i.e., the Gini (Breiman et al., 1984), and the extremely randomized trees (Extratrees) (Geurts et al., 2006) split rules. The Gini split rule determines

the best partitioning features based on the best cut point of each randomly selected feature (mtry). Typically, the cut point can be defined by the Gini index. In contrast, the Extratrees split rule determines the best partitioning features based on a random cut point of each randomly selected feature (mtry). Due to this randomization, the Extratrees algorithm requires less computation cost. Note that none of the split rules has been proven to be better in classification performance. The performance depends on the dataset and its properties (Probst et al., 2019). The model performance was evaluated using the out-of-bag samples (OOB).

7.3.4 Feature Importance

Feature importance analysis was performed on the static random classification model to explore the role of each feature in the classification. There are various methods to measure the importance score. Each method can result in different importance scores and ranks (Genuer et al., 2010; Strobl et al., 2007). Therefore, this study employed three measurement methods, i.e., impurity, permutation, and conditional permutation feature importance methods.

- **Impurity**. The impurity-based importance is the basic RF feature importance measurement method. This method measures the mean decrease in impurity (the Gini index) of all splits in all constructed decision trees (Breiman, 2001).
- **Permutation**. The permutation-based importance measures the difference between the accuracy of the original data and the accuracy of the data with that particular feature being permuted (Gregorutti et al., 2017; Nicodemus et al., 2010).
- **Conditional permutation**. The conditional permutation-based importance measures the importance similar to the permutation-based importance but performs the permutation within a conditioning grid provided by the fitted model (Strobl et al., 2008).

By employing several measurement methods, it was expected that the general tendency of the feature importance rank could be captured, and a stronger conclusion could be drawn from the results.

7.4 Results and Discussion

7.4.1 Effects of Hyperparameters

Figure 7.7 presents the effects of the RF hyperparameters on the static random classification model. Figure 7.7(a) and Figure 7.7(b) show the effects of different ntrees and mtry on the model performance (i.e., OOB classification accuracy). These figures show the importance of ntrees on the model's classification accuracy. Both figures show the increase of model performance on the increase of ntrees (i.e., increasing classification accuracy to about 0.98). The

performance tends to be stable after ntrees reach a threshold value. This result demonstrates that specifying adequate ntrees is vital for better classification performance. However, increasing it after the threshold value may only increase the computation cost without any additional benefit to the model performance. In this study, ntrees of 500 were selected for the geologic interpretation models. This value was selected to ensure both decent prediction performance and reasonable computation cost.

The models in Figure 7.7(a) and Figure 7.7(b) were constructed using the Gini split rule (Breiman, 2001) and the Extratrees split rule (Geurts et al., 2006), respectively. Figure 7.7(a) shows that the Gini split rule produced the lowest accuracy when mtry was equal to the number of features (mtry = p = 36). The accuracy increased at lower numbers of mtry. In contrast, Figure 7.7(b) shows that Extratrees split rule produced the lowest accuracy at mtry of 1. It appears that Extratrees split rule produced more stable accuracy in the variation of mtry, except the single one. Therefore, Extratrees split rule was selected for the models.

Figure 7.7(c) and Figure 7.7(d) present the effects of different mtry and minimum node size on the model performance. Figure 7.7(c) shows that Gini split rule achieved the peak accuracy at about mtry of 6, which is equal to $\sqrt{p}$. This result is in agreement with previous studies (Bernard et al., 2009; Probst et al., 2019). The accuracy dropped at larger numbers of mtry. In contrast, Figure 7.7(d) shows that Extratrees split rule produced relatively more stable accuracy at mtry of $\sqrt{p}$ to p. Both split rules produced their lowest accuracy at the minimum node size of 20. Besides this value, the variation in the accuracy was less significant. Therefore, mtry of $\sqrt{p}$ and a minimum node size of 10 was selected for the models.

7.4.2 Classification-Based Interpretation

Figure 7.8 presents the results of the static random classification model, juxtaposed with the geologic map and the borehole classes. The boreholes were colored based on the dominant ESU of each label. The grey color denotes the unlabeled data, where no boreholes were available. The classification results were determined based on the highest probability produced at each data point, as shown in the bottom panel of the figure. The black bar denotes classification errors. The static random classification model produced decent classification accuracy of 0.933 (93.3%). Note that this accuracy applies to the labeled data only. The classification results of the unlabeled data could only be qualitatively compared to the geologic map (which is also an interpretation by geologists).

The figure shows that the classification probability may reveal the geologic transition along the tunnel horizon. For example, both the geologic map and the classification probability show the transition from predominantly CCS to predominantly CSG between boreholes TB108 and TB223. In further detail, the model could also capture localized geologic conditions. For example, the model captured an increasing probability of mixed soils just before borehole TB312, where the geologic map indicates the beginning of the CCS layer on top of the tunnel horizon, creating mixed geologic conditions.

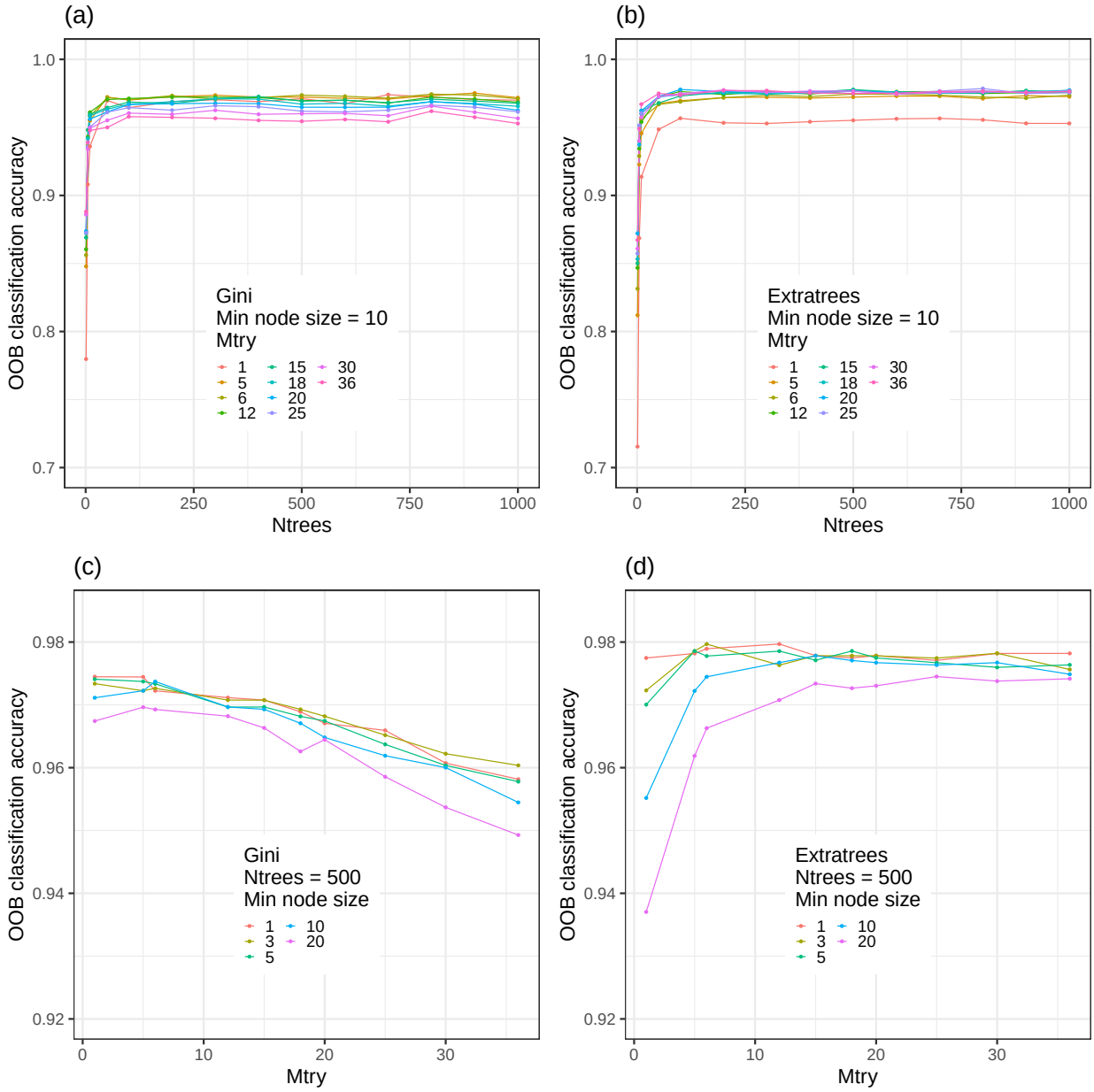

Figure 7.7: Effects of ntrees and mtry hyperparameter (top panels) and effects of mtry and minimum node size hyperparameters (bottom panels) on models constructed with different split rules, i.e., the Gini (left panels) and Extratrees (right panels).

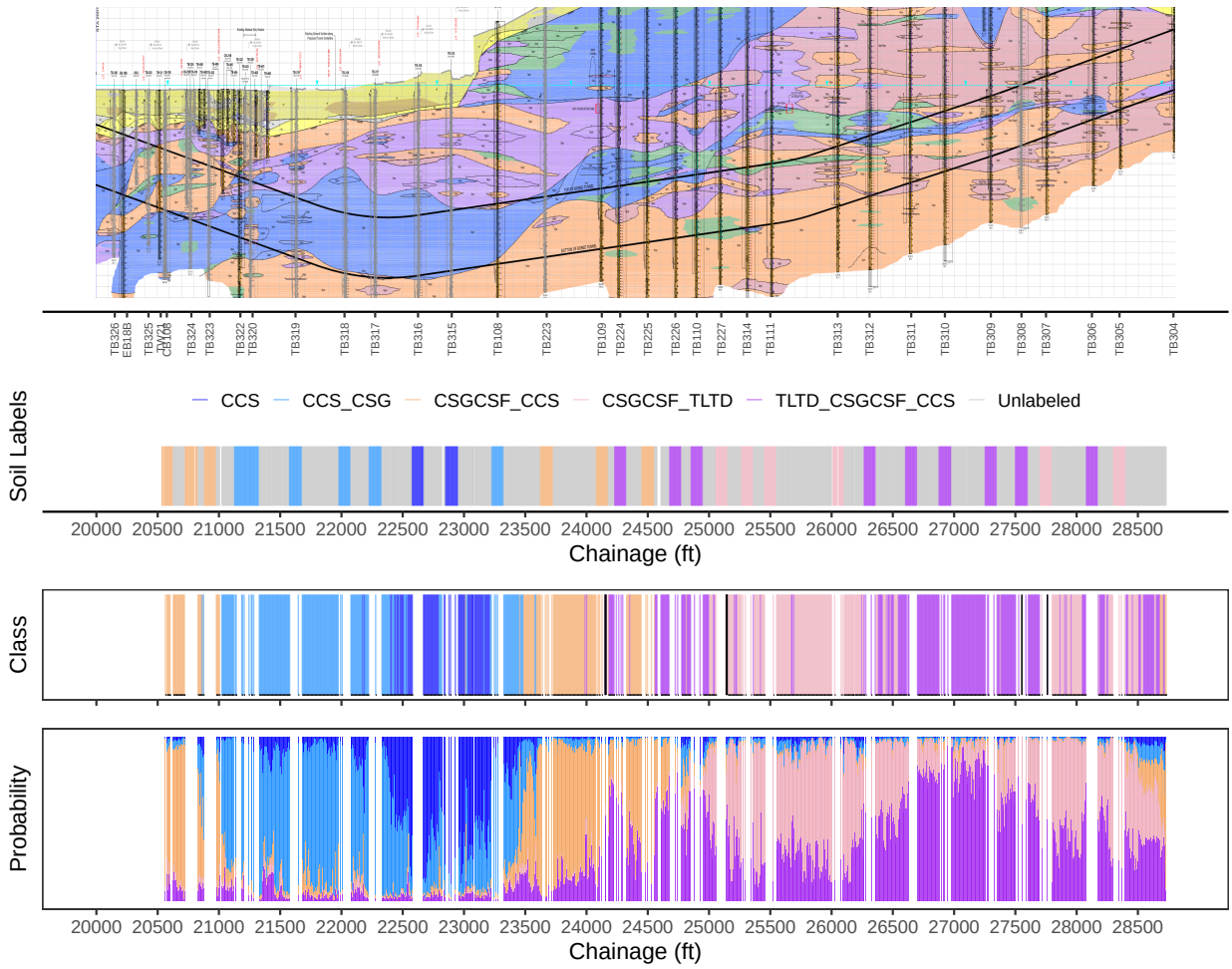

Figure 7.8: Geologic interpretation along the tunnel chainages produced by the static random classification model. Black bars denote classification errors.

Figure 7.9 presents the results of the RF dynamic sequential classification model. It can be observed that more errors were produced by the dynamic model, with the accuracy decreased to 0.913 (91.3%). This lower accuracy was expected since the training data for the dynamic model was not randomized, hence containing more bias. Furthermore, the model was only trained with very limited data and labels in the early parts of the tunnel alignment. This caused weak performance in classification analysis. However, the performance seems to have improved (i.e., similar to the static model) after the model received adequate training data and labels.

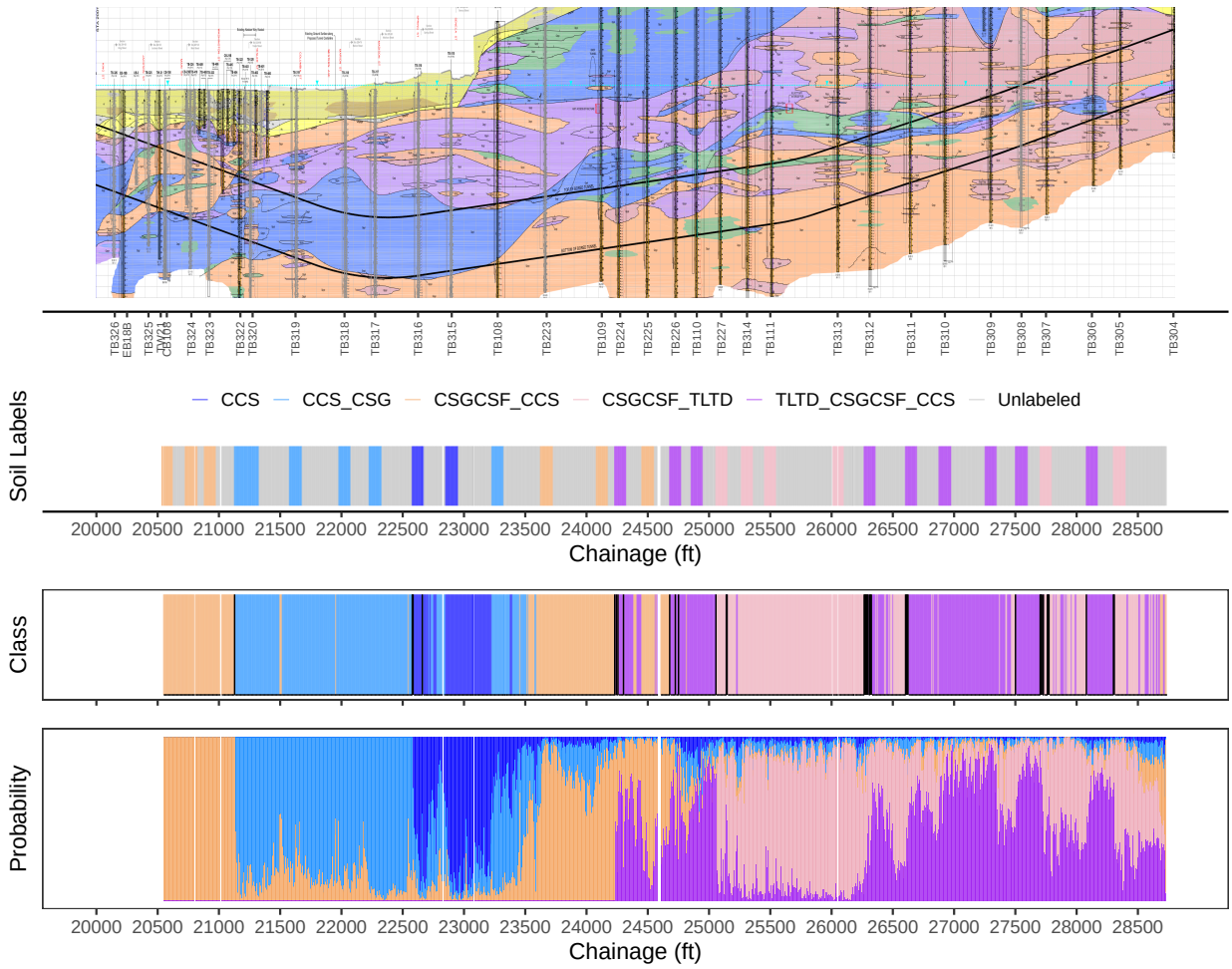

Figure 7.9: Geologic interpretation along the tunnel chainages produced by the dynamic sequential classification model. Black bars denote classification errors.

7.4.3 Regression-Based Interpretation

Figure 7.10 presents the results of the dynamic sequential regression model. Note that the regression model produced the estimated percentage of every ESU in each borehole. The estimated ESUs produced by the model were colored in accordance with the ESU colors in the geologic map. This allows a direct comparison between the model and the geologic map. The figure shows that the model produced comparable stratification patterns to the geologic map. For example, the model could capture the increase and decrease of CCS at the first half of the tunnel alignment, and the domination of CSG at the end half of the tunnel alignment. The figure also shows that the model could indicate a transition between CCS and CSG in the early tunneling phase, i.e., between chainage 21000 and 22000, where the training data was still very limited.

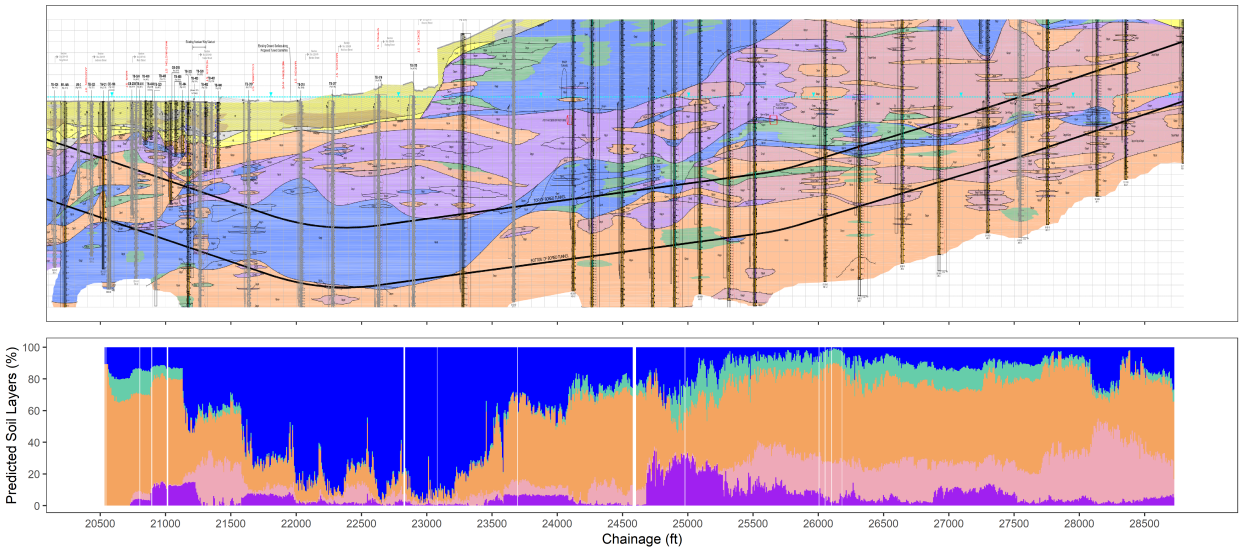

Figure 7.10: Geologic interpretation along the tunnel chainages produced by the dynamic sequential regression model.

The strong capability of the regression model can be observed in more detail in Figure 7.11. This figure compares several geologic features between chainage 22500 to 25500 ft. The figure shows that the regression model can detect the transition of predominantly CCS to predominantly CSG at about chainage 23500 ft (a). Note that this transition (i.e., the changing percentages of the ESUs) was produced merely based on patterns within the TBM operation data. No borehole was available at about this chainage. The model could also detect a small pocket of TLD about chainage 24200 (b) , large pockets of TD and CSF at

about chainage 24800 ft (c) and 25000 ft (d), repectively. These results are comparable to the geologic map.

Nevertheless, it is interesting to see a slight difference between the map and the model. The model indicated an increase of CCS at chainage 24000 ft (e). This geologic feature is not indicated in the geologic map. Unfortunately, there is no way to confirm the correct interpretation. However, note that the geologic map was interpreted by the geologists based on available boreholes. Since there is no borehole available at this location, the interpretation might be made merely by interpolating the adjacent boreholes with some consideration from the judgment of the geology processes. In contrast, the regression model produced the interpretation based on continuous TBM data. Therefore this interpretation may have more quantifiable justification.

These results demonstrate that the proposed supervised AI system can be employed to infer the encountered geologic conditions along tunnel alignments quantitatively and systematically based on continuous TBM operation data. The geologic transitions can be detected using either the classification or regression approaches. However, the regression approach can estimate the percentage of every soil type in a borehole, enabling a direct geologic interpretation and comparison to the geologic map. Furthermore, the regression approach allows transition detection even from early tunneling phases, where the classification approach performs poorly due to limited available training data.

7.4.4 Feature Importance

Figure 7.12 presents the feature importance analysis results, where features with higher ranks imply larger "weights" in the static random classification model. The relative importance was standardized to a maximum score of 1.0. Note that the relative comparison of the feature ranks is more important than the importance scores. The features are shown in color according to the EPB TBM feature groups, i.e., features related to (i) cutter, (ii) thrust, (iii) foam conditioner, (iv) polymer conditioner, (v) shield attitude, (vi) chamber, (vii) screws, (viii) belt conveyor, and (ix) tail grouting.

Less important features. It is intuitive to assume that the measured muck characteristics in the screws and the belt conveyors can be strong indicators for the encountered ground conditions. However, all the employed feature importance analysis methods show that features related to the screws and the belt conveyors are consistently at moderate to lower importance ranks. The reason might be the time delay experienced by the encountered soils at the TBM face to reach sensors at the screw and belt conveyors. The excavated soils reach the screws after spending some time in the bulkhead chamber and arrive at the belt conveyor after spending more time in the screws. Thus, these sensors measure muck characteristics from the encountered soils at a few rings behind.

Furthermore, the measured muck characteristics (i.e., weights and volumes) might not accurately represent the encountered soil characteristics. Due to the injection of ground conditioners, the physical properties of the muck can be significantly different from the physical properties of the undisturbed grounds. Subsequently, the muck is mixed and pressurized

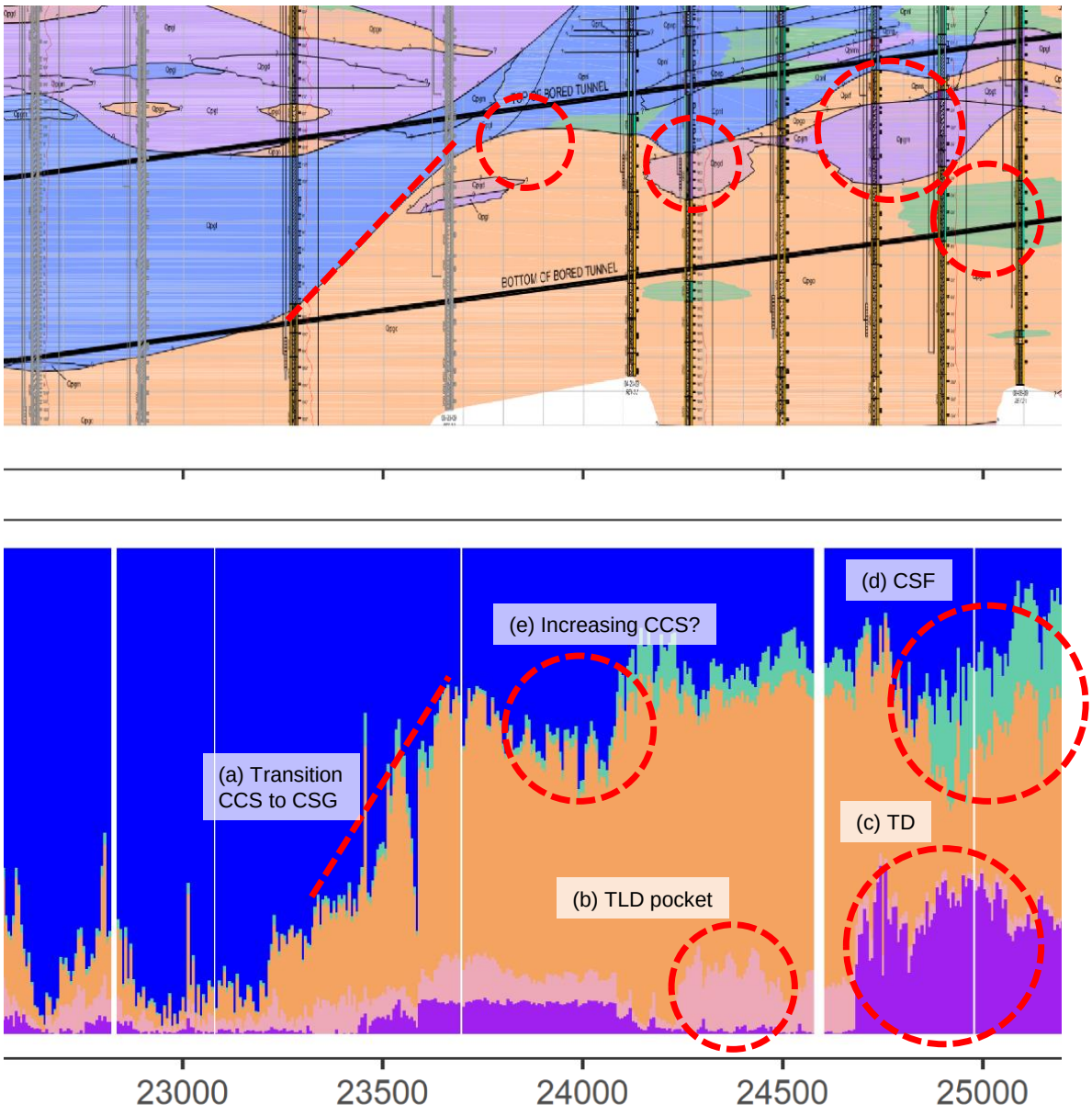

Figure 7.11: Comparison of geologic conditions interpreted by human (the geologic map) and by the AI system.

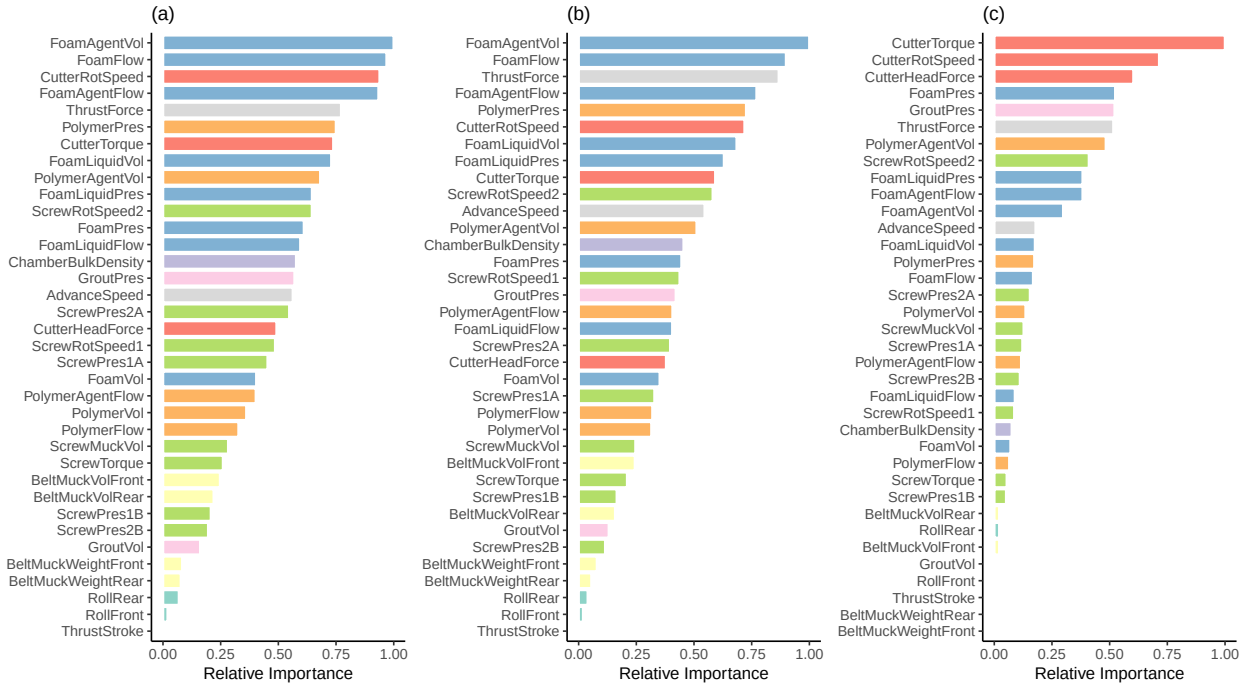

Figure 7.12: Feature importance ranks produced by (a) impurity, (b) permutation, and (c) conditional permutation methods.

inside the chamber, which can further alter its characteristics. The flow and volume of the muck inside the screws and the belt conveyor are regulated according to the screw rotation speed, which is adjusted by the operators to achieve a specified chamber pressure (Borghi, 2006; A. S. Merritt, 2004). These factors significantly affect the muck characteristics in the screw and belt conveyors.

Important features. All feature importance analysis methods consistently produced features related to cutter (i.e., cutter torque and cutter rotation speed), thrust (i.e., thrust force and advance rate), and ground conditioning systems (i.e., foam and polymer) at high ranks. This is expected as the cutter and thrust systems directly interact with the encountered grounds (Sousa & Einstein, 2012). In contrast, another feature from the thrust group, i.e., thrust stroke, was consistently at the lowest importance rank. Thrust stroke is typically pushed to its maximum stroke length or the specified tunnel lining segment length. Therefore, the stroke length values should not depend on the encountered soils. This can be a sign that the feature importance analysis returned sensible results.

All features related to ground conditioners, except the foam volume, are consistently at high ranks. The foaming agent (or surfactant) and the foam liquid (or solution) volumes were recorded by measuring their flows using flowmeters. In contrast, the foam volume

was obtained by combining the foam liquid and injected air volumes. The air volume was approximated based on the air injection pressure. Thus, the measurements of the foaming agent and liquid volumes should be more accurate than the foam volume. This inaccuracy might be the reason why foam volume is at a different rank compared to the other features in the ground conditioning system.

The association between ground conditioners and the encountered soils seems obvious. Ground conditioners are typically injected into the encountered soils at the cutter head to achieve favorable soil flow characteristics (e.g., high compressibility, low strength, low permeability). The favorable soil characteristics can be achieved by applying ground conditioners in appropriate mixtures, depending on the encountered soil type. Laboratory tests were usually conducted on retrieved soil samples before the tunneling began to get some guidelines on the ground conditioning mixture (Peila, 2014; Peila et al., 2019).

However, applying ground conditioners during tunneling is not straightforward. First, the encountered grounds at the TBM face are unknown. Second, the effective characteristics of the mixture are not only determined based on the intrinsic properties of the ground conditioners but also depend on the specific soil properties, e.g., porosity (Thewes et al., 2012; Thewes & Budach, 2010), and specific pressure levels (Mori, 2016) at every location along the tunnel alignment. This information is impossible to obtain in the pre-tunneling stages. Therefore, in reality, the TBM operators must continuously adjust the conditioners to achieve favorable soil flow characteristics during tunneling (JSCE, 2016).

7.4.5 The Importance of Feature Interactions

Figure 7.12 reveals that some TBM features may have larger weights in the classification model. However, this does not mean that each of those features has an exclusive relationship that can be used directly for the ground interpretation. This premise is demonstrated in Figure 7.13. This figure presents the accuracy of the RF models that made the classification using a single predictor feature only. The figure shows that a single feature could only produce a classification accuracy of about 0.5 or below. This applies to all features, including the high-rank features (i.e., thrust, cutter, and ground conditioning systems). An accuracy of 0.5 or lower means that the classifier model did not perform better than a random guess. The low accuracy indicates that information in a single feature may not be adequate for soil classification. This suggests that it is the interactions among the TBM features that contain the "true" information about the geologic conditions.

7.5 Conclusions

This study has proposed a supervised AI system to interpret the geologic conditions based on TBM operation data. The proposed system can infer the encountered geologic transitions in real-time during tunneling. Thus, it can be employed as a tool to systematically infer the

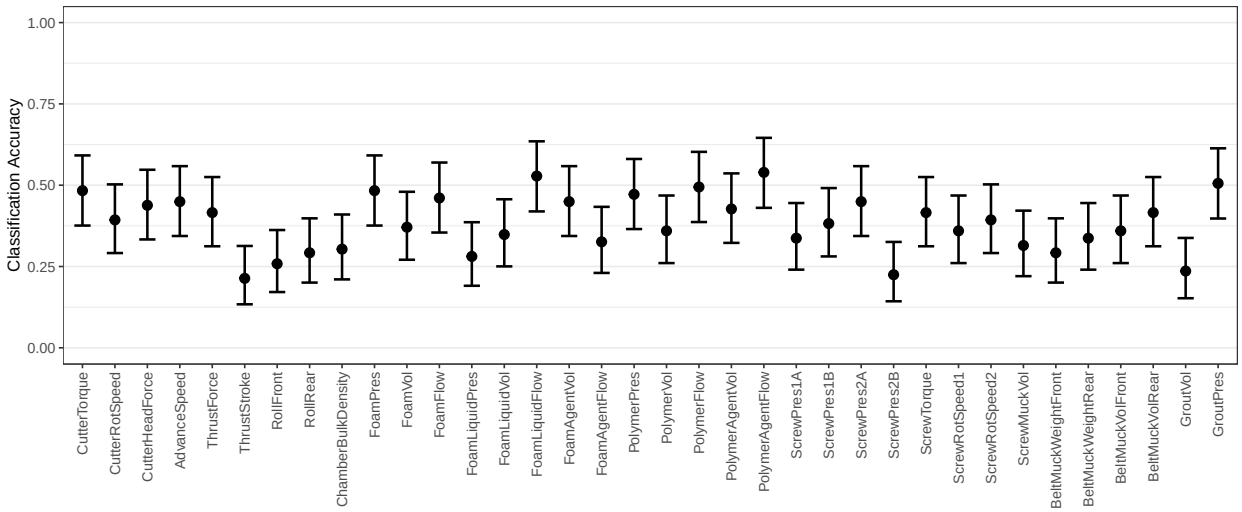

Figure 7.13: Soil classification accuracy when RF models were trained using a single predictor feature only.

encountered geologic conditions along tunnel alignments. The main findings of this study are summarized in the following.

- **Prediction model**. This study demonstrates that the geologic interpretation model can be developed using RF, which can deliver stable and decent prediction performance with simple hyperparameter tuning. This is important considering the sequential process in tunneling, where the model needs to be dynamically retrained whenever new training data become available (i.e., when the TBM arrives at a borehole location).

- **Detecting geologic transition**. The geologic transition detection can be performed using classification or regression approaches, i.e., by representing geologic conditions as probabilities of several geologic labels or by representing the thickness of every geologic deposit at borehole locations as numerical percentages. The regression approach is computationally more expensive than the classification approach since it requires multiple regression models to be built (one for each geologic deposit type). However, it can estimate the percentage of every soil type in a borehole, enabling a direct geologic interpretation. This approach also enables better prediction performance from early tunneling phases, where the classification approach performs poorly due to limited available training data.

- **Geologic information in TBM features**. This study quantitatively shows that TBM data contain information about the geologic conditions. The feature importance

analysis has shown that features related to the cutter, thrust, and ground conditioning systems contain relatively larger "weights" in the geologic interpretation model. However, the information in a single feature may not be adequate for interpretation, even if it is from a highly important feature. This indicates that information about the geologic conditions is contained in the interactions of the features (i.e., nonlinear normalization of the features). Capturing these interactions is essential to infer the geologic conditions encountered during tunneling. This highlights the importance of nonparametric and nonlinear machine learning algorithms to exploit complex interactions of TBM data.

Based on these findings, more work can be done to further develop the proposed geologic interpretation system. This may include the following.

- **Machine learning algorithms and extrapolation problems**. It is important to investigate the effects of different machine learning algorithms on the model, especially how they behave in extrapolation problems. The TBM may always encounter new geologic conditions along the alignment, forcing the model to extrapolate. It is widely known that machine learning, especially the decision tree-based algorithms such as RF, may perform poorly in extrapolation problems (Ebert et al., 2014; Webb et al., 2020).

- **Change of optimal hyperparameters along the tunnel alignment**. Investigating how the optimal hyperparameters change along the tunnel alignment can be important since tunneling involves sequential learning, where the model should be continuously retrained. The new data may affect the best hyperparameter configuration. Highly fluctuating hyperparameter configurations mean the model should be continuously tuned (e.g., by cross-validation), which may cause a computation time bottleneck during tunneling.

- **Model interpretation**. Understanding the model is essential to build some generalizations from the model. This is also critical to ensure model reliability. However, this task can be challenging since the appropriate way to interpret the model, e.g., by measuring the feature importance, is still an active research area (as discussed in Chapters 5 and 6). Some efforts may include exploring different model interpretations and feature importance analysis methods, as well as implementing the model in different tunneling cases.

Chapter 8

Detecting Geologic Anomalies during Tunneling

Some contents of this chapter have been accepted in Apoji, D. and Soga, K. (2023). *Soil Clustering and Anomaly Detection Based on EPBM Data using Principal Component Analysis and Local Outlier Factor*. Geo-Risk. American Society of Civil Engineers.

8.1 Introduction

8.1.1 Background

Detecting the changing ground conditions during tunneling has always been one of the most critical tasks in tunnel boring machine (TBM) operation. The TBM operators need to be continuously aware of the changing ground conditions. The geologic map can provide a guide. However, it only contains limited information since it is interpreted from limited boreholes at discrete locations. Therefore, during tunneling, the operators must continuously infer the encountered geologic conditions based on TBM operation data (Garcia et al., 2021). The TBM operation data is typically generated every few seconds by numerous sensor measurements. Interpreting this kind of data stream can be challenging and lead to unsystematic interpretation, human bias, and errors.

Many studies have been produced to develop a more systematic method to infer the encountered geologic conditions based on TBM operation data. Many used data-driven approaches such as using probabilistic models (Sousa & Einstein, 2012) and machine learning (Apoji et al., 2022b; Erharter et al., 2020; Q. Zhang et al., 2019; Zhao et al., 2019). Most of the developed data-driven methods used a supervised learning method to convert the TBM operation data into the predicted ground classifications. Supervised learning is a powerful prediction tool. However, it strongly depends on available training data and the ground truth labels. In tunneling, the ground truth (i.e., borehole data) is very sparse and limited. It typically only captures less than one percent of the geologic conditions along

the tunnel alignment. Furthermore, supervised ground prediction methods will always have problems when the TBM encounters new ground conditions (i.e., soil units) that have not been experienced before. Unfortunately, this will happen all the time in tunneling due to the variability of geology.

8.1.2 Related Works

Unsupervised learning is a potential tool to tackle this problem. Unlike supervised learning, unsupervised learning algorithms do not require training data and labels. The algorithms find structures and patterns in the given data set as is. Some previous studies have tried to implement unsupervised anomaly detection methods in tunneling. Sheil, Suryasentana, and Cheng (2020) explored various anomaly detection methods for micro-tunneling applications. The study indicates that non-parametric anomaly detection methods may perform better than density-based methods. However, the conclusion was derived only from univariate data set as input (i.e., jacking force).

Meschke et al. (2019) implemented Page-Hinkley Test (Hinkley, 1971), an unsupervised learning algorithm, to detect a sudden change in the ground during TBM tunneling. However, they also only used univariate data set as the input, i.e., the specific torque, which is computed based on the torque and penetration rate. Cao et al. (2021) implemented a drift detection method (Saadallah et al., 2019), also to detect sudden changes in the ground during TBM tunneling. Different from the previous studies, they used a multivariate data set as the input. However, as a trade-off, it can be difficult to interpret the model. Furthermore, the study was carried out to detect TBM advances through a concrete wall. No discussion on the performance of detecting less sensitive changes such as geologic variability.

8.1.3 Objectives

This study aims to propose a more interpretable unsupervised anomaly detection method to infer changing ground conditions in real-time during tunneling based on multivariate TBM data. The proposed method uses a dimensionality reduction method, i.e., the Principal Component Analysis (PCA), to (i) extract information on the ground conditions from selected TBM variable features, and (ii) project the lower dimension data set on a geometrical space. Subsequently, the proposed method uses a density-based outlier detection method, i.e., the Local Outlier Factor (LOF), to measure the degree of the anomaly of the projected data points. The combination of the geometrical space projection and the density-based clustering was expected to enable the geometrical interpretation of the model.

8.2 Data

8.2.1 Tunneling Case and Geologic Conditions

This study used a data set of the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA. Along the tunnel horizon, the geologic conditions were dominated by overconsolidated glacial and non-glacial pre-Vashon geologic units. More discussion of this tunneling case and the geologic conditions can be found in Chapters 5 and 7, respectively.

8.2.2 Data Description and Preparation

Similar to Chapter 7, this study used ring-length aggregate spatial series as the observation data points. A set of variable features was selected from the earth pressure balance (EPB) TBM data set as the input for the proposed unsupervised anomaly detection method. The input consisted of 13 features related to the cutter, thrust, and ground conditioning systems (Table 8.1). As concluded in Chapter 7, these features contain larger weights in the soil classification model. Therefore, relationships among these features are expected to contain information on the ground conditions. The data preparation and cleaning were done in accordance with the steps discussed in Chapter 7.

Table 8.1: List of EPB TBM features used in the unsupervised geologic anomaly detection method.

No.	Feature	Unit	Group
1	Cutter torque	kNm	Excavation
2	Cutter rotation speed	rpm	
3	Copy cutter position	deg	
4	Copy cutter stroke	mm	
5	Thrust force	kN	Advancing
6	Thrust stroke	mm	
7	Advance rate	mm/min	
8	Foam volume	m^3/m	Ground conditioning
9	Foam pressure	kPa	
10	Polymer volume	m^3/m	
11	Polymer pressure	kPa	
12	Slurry volume	m^3/m	
13	Slurry pressure	kPa	

8.3 Methods

The proposed unsupervised anomaly detection method was developed under several hypotheses, as listed below.

1. Some variable features in EPB TBM data contain information on the ground conditions. Therefore, EPB TBM data with properly selected features can be used to identify the ground conditions of a chainage location.

2. By projecting the data on a geometrical space, data points (i.e., chainage locations) with similar ground conditions will cluster together. A dimensionality reduction method such as PCA can be used to extract the information from the original data set into a data set with a lower dimension. This enables the projection of the data into a geometrical space.

3. Data points that do not belong to any clusters can be identified as potential outliers (i.e., anomalies) and may have a unique ground condition. The degree of the anomaly for every data point on the projection space can be measured using density-based outlier detection methods. LOF may be suitable for this purpose, considering an EPB TBM data set is a non-parametric data set where clusters of the projected data points may have different density levels.

4. The PCA projection of EPB TBM data can be made dynamically in a sequence during the tunneling (in an online manner). Meaning that a new projection can be created every time the EPB TBM advance forward and produce new streams of data. The LOF of the newest data point on the projection can be used to identify the degree of the anomaly of the ground condition at the EPB TBM location. A new data point with LOF ¿ 1.0 indicates that the EPB TBM is entering a new ground condition that is not similar to the previous ground conditions. The higher the LOF, the more the ground condition potentially changes.

8.3.1 Model Setup

These hypotheses can be derived into an unsupervised anomaly detection method to infer changing ground conditions in real-time during tunneling. Table 8.2 presents the algorithm of the proposed method. In this method, PCA was used to transform the original data set into PC scores at every TBM advance in an online manner. In other words, PC scores at chainage i were obtained by transforming the data set with all observation data points from the beginning of the tunneling to chainage i. Subsequently, PC scores at chainage i+1 were obtained by transforming the data set with all observation data points from the beginning of the tunneling phase to chainage i+1. This process was repeated until the end of the tunneling sequence.

Note that PCA transforms an original data set into PC scores based on variances of the data variables. Since variable features of EPB TBM data are in various units, the magnitude

and distribution of every feature can be significantly different. Therefore, the EPB TBM variable features should be standardized before performing PCA. Without standardization, the PC scores will be governed by the features that the units are in higher magnitude. The standardization can be done by transforming every feature to have zero mean and a unit standard deviation.

This study selected minimum points k-nearest neighbors of 10 for the LOF analysis. This value is the minimum k value as recommended by Breunig et al. (2000). Based on their experiment, any k value below this threshold may produce unwanted statistical fluctuations that can result in insensible LOF.

Table 8.2: Algorithm of the unsupervised geologic anomaly detection method.

1: data = TBM data with selected features
2: **for** *chainage* in StartCainage to EndChainage **do**
3: Scale data from StartChainage to *chainage*
4: Compute PCA on the scaled data
5: Get PC scores
6: Select k value for LOF
7: Compute LOF on PC scores (e.g., PC1 and PC2)
8: Get LOF at *chainage* data point
9: **end for**

8.3.2 Principal Component Analysis

PCA is an unsupervised learning algorithm to reduce the dimensionality of a large data set. The algorithm can find representative variables in a lower dimension that contain most of the variability (i.e., information) in the original data set (Hastie et al., 2009; James et al., 2013). The algorithm converts measured variables $(X_1, X_2, \ldots, X_p)$ to derived variables called principal component (PC) scores $(Z_1, Z_2, \ldots, Z_p)$. The amount of stored information is ordered. The first PC contains the most information (i.e., the largest variance in the original data set), and followed by the second PC.

The first PC and its relationship to the measured variables can be written as

$$Z_1 = \phi_{11}X_1 + \phi_{21}X_2 + \cdots + \phi_{p1}X_p, \tag{8.1}$$

where coefficients of ϕ are the PC loadings, which essentially are the eigenvectors. Typical data sets consist of n observations on each measured variable $p(n \times p)$. Thus, the k^{th} PC of data with i^{th} observation data points and j^{th} measured variables can be generalized with the equation

$$z_{ik} = \sum_{j=1}^{p} \phi_{jk}x_{ij}. \tag{8.2}$$

PCA can be solved using the eigendecomposition of the covariance matrix or singular value decomposition (SVD). For computation, the SVD solver is typically more efficient and stable.

8.3.3 Local Outlier Factor

LOF is an algorithm that measures the degree of being an outlier by comparing the local density of a data point to the local densities of the neighboring data points (Breunig et al., 2000). Conceptually, a data point is considered to be an outlier if it has a substantially lower density than its neighbors. LOF can be computed by first determining k-nearest neighbors and measuring the reachability distance (RD) of data points in a given data set. RD of a point p with respect to a point o is defined as the larger value between the distance of the two data points and the k-nearest neighbor distance of point o

$$RD_k(p, o) = \max\{distance(p, o), k\ distance(o)\}. \tag{8.3}$$

Local reachability density (LRD) measures how far a point is from the nearest cluster. A low value of LRD implies that the point is far from the closest cluster. LRD is defined as the inverse of the average RD of a data point p from its neighbors

$$LRD_k(p) = 1/\left(\frac{\sum_{o \in N_k(p)} RD_k(p, o)}{|N_k(p)|}\right), \tag{8.4}$$

where $N_k(p)$ is the minimum points k-nearest neighbors of p.

Finally, LOF score can be computed as a ratio of the average LRD of the neighbors and the LDR of the data point p,

$$LOF_k(p) = \sum_{n \in N_k(p)} \frac{LRD_k(o)}{LRD_k(p)} \times \frac{1}{|N_k(p)|}. \tag{8.5}$$

LOF values of 1 or less indicate that the data point is similar or inlier to its neighbors. In contrast, values significantly larger than 1 indicate outliers. LOF can identify outliers in a data set that has clusters with different densities.

8.4 Results and Discussion

8.4.1 PCA Projection

Figure 8.1 presents PCA projections of the EPB TBM data set at nine chainage locations along the tunnel alignment. The figure shows examples of two-dimensional (2D) projections of the first two PCs, i.e., PC1 on the x-axis and PC2 on the y-axis. Each data point represents EPB TBM data at a chainage location, hence, the number of data points increases as the chainage increases. To interpret the projection, the data points were colored based on

available boreholes according to the largest percentage of soil units at each of their location. Data points without any borehole information were colored in grey. The size of the data points represents the percentage of the dominant soil unit. Larger data points indicate that the soil unit was dominant in the tunnel horizon, while smaller data points indicate more mixed ground conditions.

Several findings can be drawn from Figure 8.1. First, the projected data points shifted dynamically as the data increased. It can be observed that every chainage plot had different patterns of data points. This is true since every new observation data point added to a data set can change its variance. The change in data variance affects the PCA transformation and alter the produced PC scores. Second, it appears that data points with similar ground conditions tended to cluster together. Interestingly, the dynamic behavior of the projected data points did not substantially affect the soil clusters. Data points with similar ground conditions were continually clustered. These findings confirm the first two hypotheses of this study. This suggests the potential application of PCA projection to cluster data points with similar ground conditions, dynamically.

Better visualization of the clusters can be seen in Figure 8.2. This figure presents the three-dimensional (3D) projection (i.e., PC1, PC2, and PC3) of all observation data points in the data set, which is comparable to Figure 8.1(i). The 3D projection shows better cluster separations that cannot be clearly seen in the 2D projection. This is true since the three PCs contain more information than the two PCs. As shown in the scree plot (Figure 8.3(a)), the first two PCs capture about 42% of the information in the original data set. While the first three PCs capture about 55% of the information. The figure also shows that the first eight PCs can represent about 90% of the information in the original data set. This demonstrates the capability of PCA to reduce the dimensionality of high-dimensional data. The dimensionality reduction enables the projection and visualization of the data without losing too much of its information.

Figure 8.3(b) presents the PCA biplot, which shows the contribution of each original variable feature to the first two PCs. As mentioned, these PCs stored most of the information in the data set. The biplot shows that many variable features shared large contributions to PC1 and PC2, for instance, the cutter torque, cutter rotation speed, thrust force, advance rate, and features related to ground conditioners. The biplot also shows that three variable features had insignificant contributions to the PCs, i.e., the thrust stroke, copy cutter stroke, and copy cutter position. These are sensible since the thrust stroke is specified by the target advancement length, which is typically determined based on tunnel segment length. Similarly, the copy cutter features are used to create over-excavation for shield maneuvers (Chanchaya & Suwansawat, 2014). These features are specified based on the tunnel alignment design and may not relate to the ground conditions.

8.4.2 Local Outlier Factors

It has been demonstrated that the PCA projection can dynamically cluster EPB TBM data points with similar ground conditions. This implies that a data point that does not belong

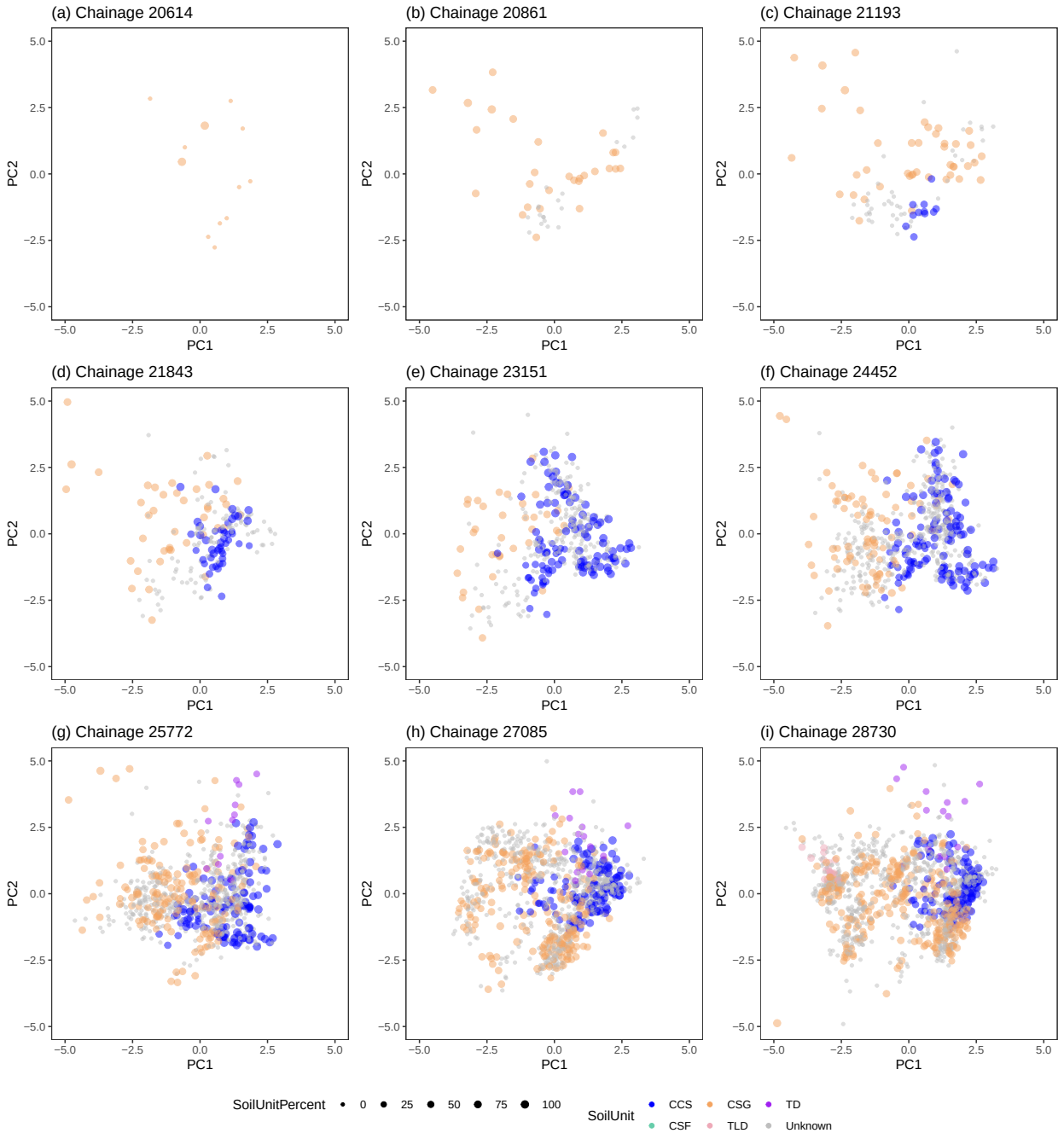

Figure 8.1: Examples of 2D PCA projections of the EPB TBM data at different chainages along the tunnel alignment.

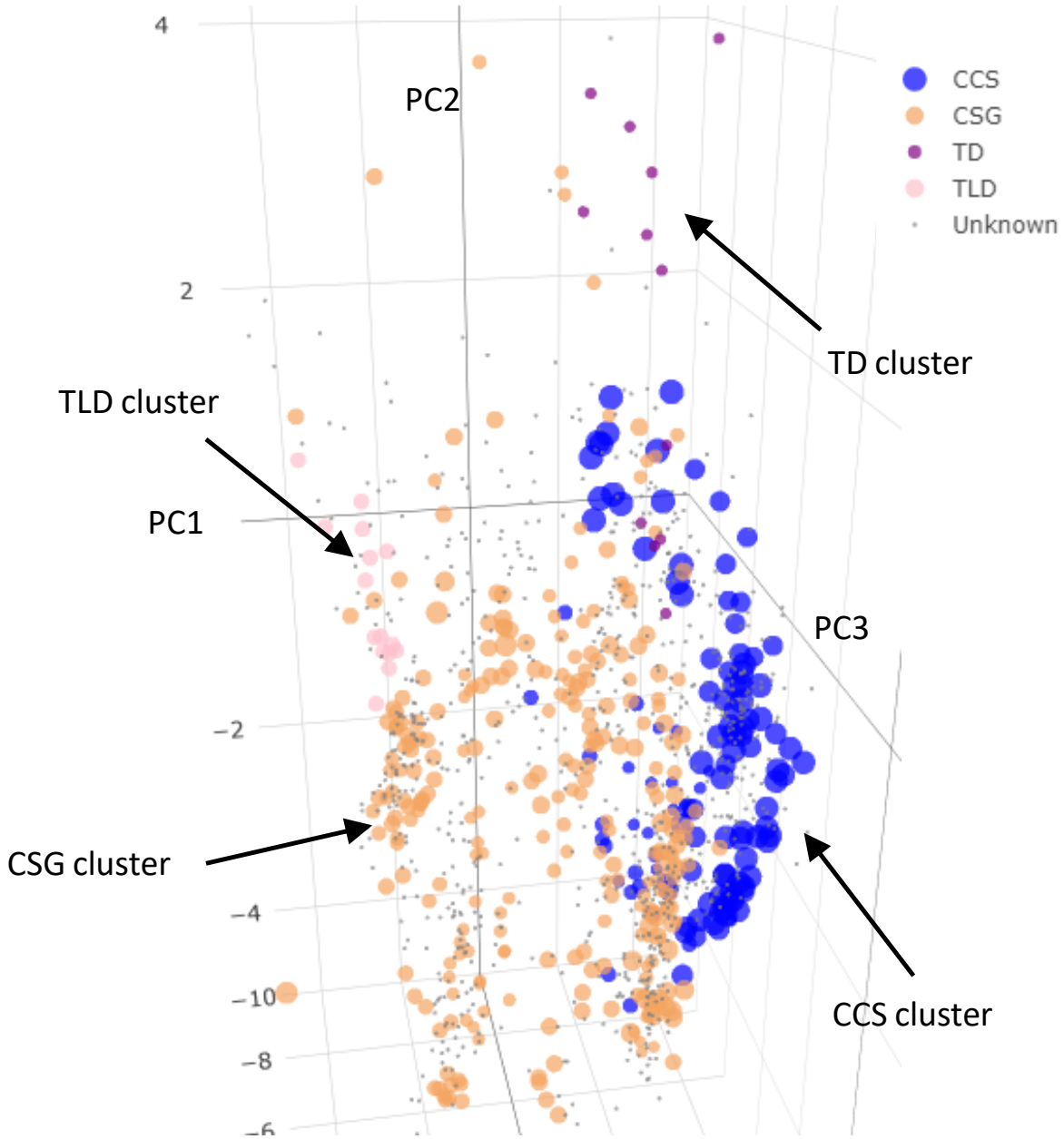

Figure 8.2: Example of 3D PCA projection of the EPB TBM data (at Chainage 28730).

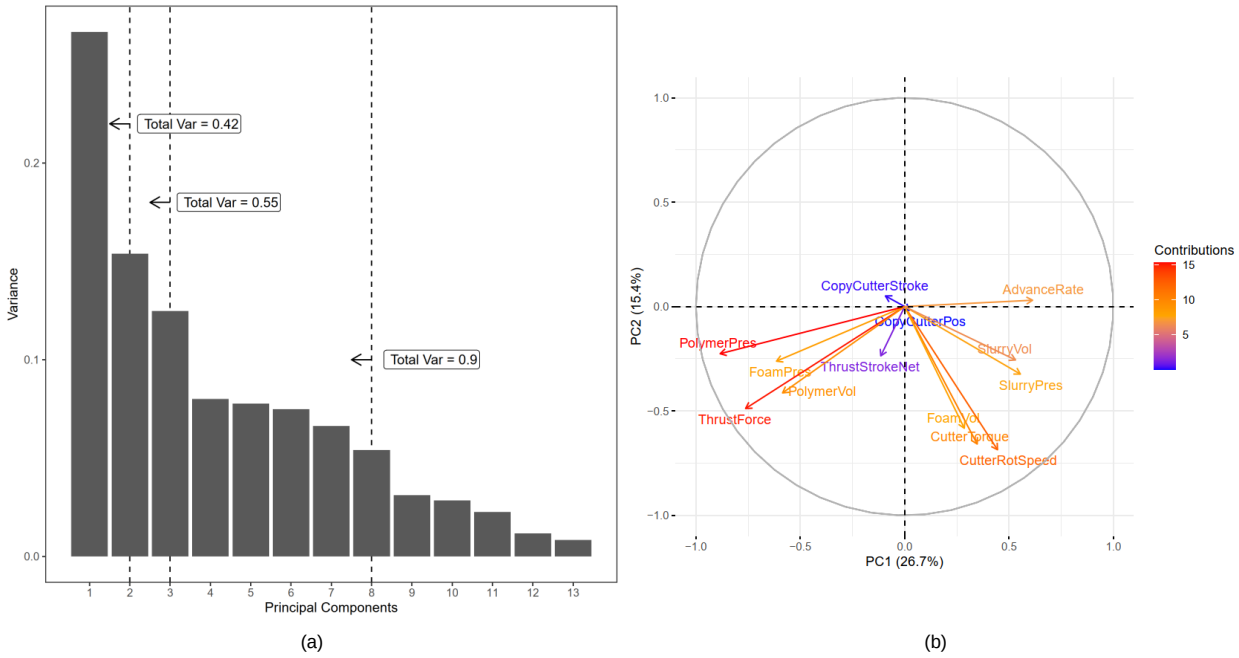

Figure 8.3: Interpreting PCA results: (a) Scree plot to visualize the information in each PC score, (b) biplot to visualize the contribution of each variable feature to PC1 and PC2.

to any clusters may have a unique ground condition. This concept can be used to indicate anomalies in ground conditions during tunneling. If a new data point is projected far from any existing clusters, the EPB TBM may be entering a new ground condition. The degree of the anomaly can be measured using LOF.

Figure 8.4 presents 2D PCA projections together with the LOF plots, which represent the degree of the anomaly of the data points. The 2D projection plot was still used (instead of the 3D projection) for easy visualization and interpretation. The LOF are shown in red circles. The larger the LOF of a data point, the larger the circle that surrounds the data points. Figure 8.4(a) shows an example of a representation when the EPB TBM appears to be located in a similar ground condition to the previous ground conditions. The new projected data point is located inside clusters. This results in a low LOF (small red circle). In contrast, Figure 8.4(b) shows an example of a representation when the EPB TBM enters a new ground condition that differs from the previous ground conditions. This is shown by the new projected data point, which is located outside any clusters on the projection. This results in a high LOF (large red circle).

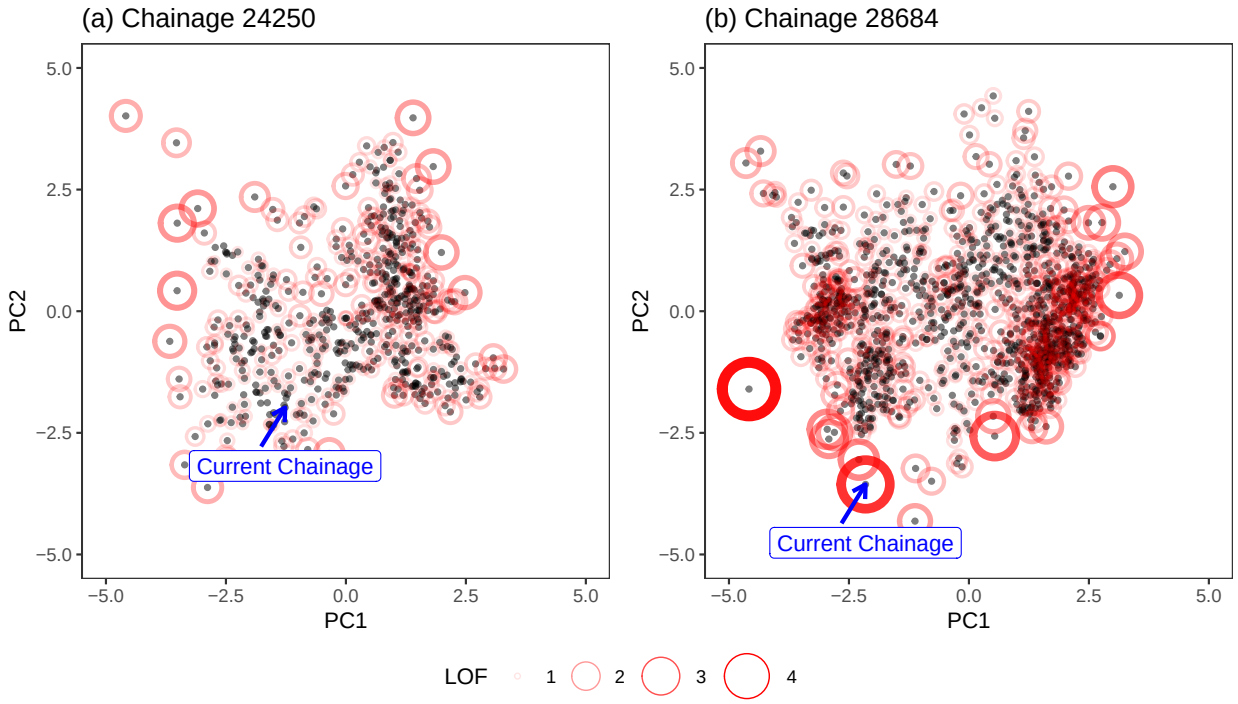

Figure 8.4: Examples of degree of anomaly of projected data points: (a) current EPB TBM location inside clusters, (b) current EPB TBM location outside clusters (i.e., potentially at new ground conditions).

8.4.3 Anomaly Detection

Figure 8.5 presents the ground anomaly level plot, i.e., the LOF of every new EPB TBM data point projection along the tunnel alignment. In more detailed explanation, a PCA projection was created for every chainage along the tunnel alignment. Subsequently, for each PCA projection, the LOF of the latest data point in the projection (i.e., the current EPB TBM chainage) was obtained and stored to measure the anomaly of ground conditions. The longitudinal geologic map of the tunnel was plotted as a background on the figure to provide a qualitative comparison between the LOF and the interpreted ground conditions.

This investigation was performed to simulate a real-time anomaly detection of ground conditions during tunneling. The results indicated that locations with high LOF values were associated with changes in ground conditions. The LOF values spiked when the TBM entered mixed soil regions and fully cohesive soils at approximate chainages of 20700, 21000, and 22200, respectively. Additionally, spikes at chainage 21500 might suggest the presence of a sandy soil bump in predominantly cohesive soils. The LOF values also spiked when the EPB TBM encountered mixed soils with substantial Till-like and Till deposits at chainages

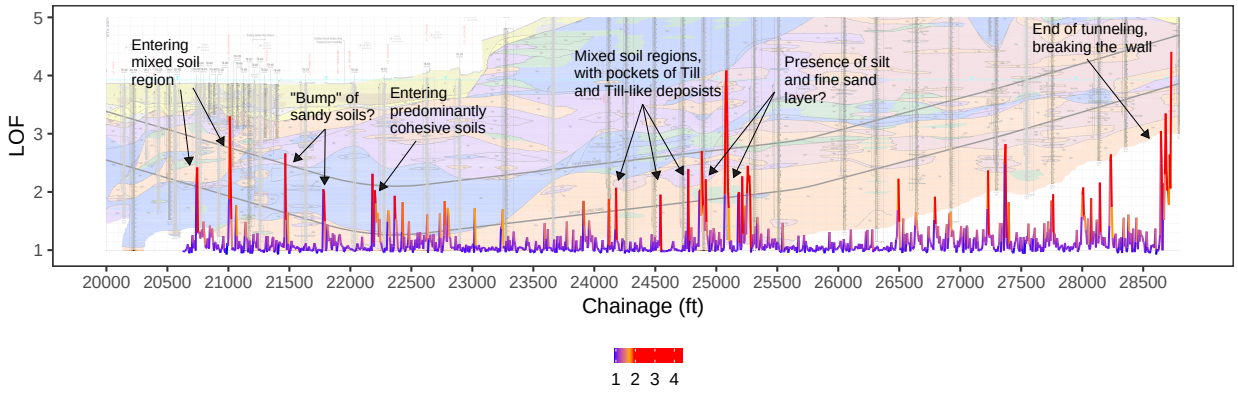

Figure 8.5: Longitudinal plot of the geologic anomaly measurement: LOF of every new EPB TBM data point projection along the tunnel alignment.

24100 and 24700, respectively. LOF peaks after chainage 25000 may indicate the presence of mixed soil regions, specifically silt and fine sand layers. Finally, several LOF spikes occurred after chainage 26500, as the EPB TBM encountered irregular mixed soils until it reached the end of tunneling. These results confirm the third and fourth hypotheses of the study.

This result suggests the potential application of PCA projection combined with LOF for anomaly detection of ground conditions during tunneling. This method allows real-time interpretation of TBM data without any predefined soil labels that are required in supervised learning-based interpretation methods (e.g., Apoji et al. (2022b)). The development of this method is expected to be able to assist TBM operators in detecting the changing ground conditions systematically and quantitatively.

8.5 Conclusions

This study has introduced a more interpretable unsupervised anomaly detection method to infer changing ground conditions in real-time during tunneling based on multivariate EPB TBM data. The method combines PCA as a tool to project the data into a lower dimension space and LOF to measure the degree of the anomaly of the projected data points. This study has demonstrated that, with appropriately selected EPB TBM features, PCA could dynamically cluster EPB TBM data according to the ground conditions. Interestingly, the dynamic behavior of the projected data points did not substantially affect the soil clusters as they were always grouped together. This study has also demonstrated that LOF could be a sensible measure to detect changing ground conditions.

Further investigations can be performed to improve the performance of the proposed method and to ensure its reliability and generalizability. (i) Applying a nonlinear dimension-

ality reduction method may be worth to explore to better cluster the data points, especially since interactions among EPB TBM variable features can be complex and nonlinear (e.g., Apoji et al. (2022a)). (ii) Investigating the effect of selected k in LOF analysis, which can affect the degree of anomaly (Breunig et al., 2000). (iii) Using finer observation data points as the input to evaluate its reliability against data noises. (iv) Testing the method with more tunneling cases to evaluate its generalizability.

Chapter 9

Connecting TBM to the Ground Responses

Some contents of this chapter have been accepted in Apoji, D., Ning, Z., and Soga, K. (2023). *Connecting EPBM Data to Ground Movement Data using Machine Learning.* Geo-Congress. American Society of Civil Engineers.

9.1 Introduction

9.1.1 Background

Tunneling-induced ground movements constitute a major concern in tunneling projects. The induced ground movements are mainly governed by (i) the tunnel spatial geometries, (ii) the geologic conditions, and (iii) the tunneling processes, i.e., the tunnel boring machine (TBM) behaviors. The movements can be estimated using various methods, for instance, empirical (Mair & Taylor, 1999; Peck, 1969), analytical (Loganathan & Poulos, 1998; Pinto & Whittle, 2014; Pinto et al., 2014), and numerical methods (Avgerinos et al., 2018; Kasper & Meschke, 2004; Komiya et al., 1999). Every method offers different approaches, but the key input parameters are always dominated by the tunnel geometries and the geologic conditions. Limited attention has been given to the effects of TBM operations.

9.1.2 Related Works

Maximum Ground Movements. Some studies have been conducted to include more information related to TBM operations in tunneling-induced ground movement models. Most of the studies used a data-driven method by utilizing TBM operation data and machine learning algorithms. Shi et al. (1998) proposed a method to estimate tunneling-induced ground movements using Artificial Neural Networks (ANN). The study used a data set from the Brasilia Tunnel project in Brazil. Two types of ANN were examined, i.e., general and modular ANNs. They concluded that the modular ANN model produced better prediction

accuracy and computation convergence in the training process. This is an early implementation of data-driven methods for estimating tunneling-induced ground movements.

Suwansawat and Einstein (2006) examined the performance of various ANN architectures in estimating the maximum surface settlements. The study used a data set from the Bangkok MRTA project in Thailand. The best ANN architecture was evaluated by developing 18 ANN models with different numbers of hidden layers (1, 2), hidden nodes (10, 15, 20), and training epochs (1000, 2000, 3000). They reported that the best ANN architecture was achieved using one hidden layer, 20 nodes, and 2000 training epochs.

Boubou et al. (2010) proposed a method to estimate the cross-section profile of tunneling-induced ground movements. The study used a data set from the subway line B tunnel project in Toulouse, France. Two algorithms were used to develop the models, i.e., ANN and least square approximation. The most influential TBM features were evaluated using a feature elimination procedure. They concluded that the models could produce more realistic ground movements since the estimated cross-sectional ground settlements were not restricted to the typical Gaussian shape. They also concluded that the three most influential parameters from models were the same, i.e., the advance rate, hydraulic pressure used for the cutting wheel, and TBM vertical guidance parameter. Furthermore, the sensitivity analysis results indicated that the ANN model could not be trained with less than 40% of the total data set and in optimal performance when trained using at least 60% of the total data set.

The implementation of data-driven methods to estimate tunneling-induced ground movements has become more prevalent in the past decade. Different machine learning algorithms have been used as prediction models. For example, Bouayad and Emeriault (2017) implemented the Adaptive Neuro-Fuzzy-based Inference System algorithm (ANFIS) to develop a prediction model for ground settlement data from the subway line B tunnel in Toulouse, France. Goh et al. (2018) implemented Multivariate Adaptive Regression Splines (MARS) to develop a prediction model for ground settlement data from three different earth pressure balance (EPB) TBM tunneling projects in Singapore. Su et al. (2022) implemented Extreme Gradient Boosting (XGBoost), an ensemble supervised learning algorithm, to develop a prediction model for 533 cases of ground surface settlements from an urban shield tunneling project.

These studies have demonstrated the opportunity of estimating tunneling-induced ground movements using data-driven methods. Machine learning algorithms also allow more realistic estimation results since tunneling-induced ground movement profiles may not always follow the assumed Gaussian type. Furthermore, the algorithms enable flexible mapping between the input features (TBM operation parameters) and the response (the estimated ground movement). The model could produce good prediction performance, although they were built using different input features.

Comparative Studies. Some studies have been conducted to compare and evaluate the performance of different machine learning algorithms in estimating tunneling-induced ground movements. These studies include Mahmoodzadeh et al. (2020), Ocak and Seker (2013), Tang and Na (2021), P. Zhang, Wu, Chen, and Chan (2020), and W. G. Zhang et al. (2021). These studies involved various machine learning algorithms, from regression-based

and decision trees-based algorithms to neural network-based algorithms. The algorithms used in studies and the best performer algorithm are summarized in Table 9.1.

Table 9.1: Comparison of machine learning algorithms for estimating tunneling-induced ground movements from previous studies.

Reference	Models	Best Performer
Ocak and Seker (2013)	ANN, SVM, GPR	GP
Mahmoodzadeh et al. (2020)	LSTM, DNNs, kNN, GPR, SVM, DT, LR	DNN
P. Zhang, Wu, Chen, and Chan (2020)	BPNN, GRNN, ELM, SVM, RF	RF
W. G. Zhang et al. (2021)	XGBoost, ANN, SVM, MARS	XGBoost
Tang and Na (2021)	SVM, RF, BPNN, DNN	RF

ANN: Artificial Neural Networks
BPNN: Backpropagation Neural Networks
DNN: Deep Neural Networks
DT: Decision Tree
ELM: Extreme Machine Learning
GPR: Gaussian Process Regression
GRNN: General Regression Neural Networks
kNN: k-Nearest Neighbor
LR: Linear regression
LSTM: Long Short-Term Memory
MARS: Multivariate Adaptive Regression Spline
RF: Random Forest
SVM: Support Vector Machine
XGBoost: Extreme Gradient Boosting

The table shows that almost every study produced a different best-performer algorithm. This means no solid conclusion can be taken to decide the best algorithm for estimating tunneling-induced ground movements. The difference can be caused by various reasons, such as different tunneling data sets, data preparation methods, feature selection, and hyperparameter methods. These factors are critical in building a machine learning model and, therefore, might have substantial effects on the prediction performance.

Longitudinal Ground Movements. The discussed studies focused on connecting TBM operation data to the maximum surface settlements. It should be noted that tunneling-

induced ground movement is a function of the TBM position. The ground movement evolves as the TBM advances toward and passes the location. This progression can be related to different TBM operation parameters (JSCE, 2016; Suwansawat, 2002). F. Wang et al. (2013) reported an early development of the study that connected TBM operation parameters to the longitudinal ground movements. They used a wavelet smooth relevance vector machine algorithm to model longitudinal ground movement progression during EPB TBM tunneling. The algorithm was selected as it requires simpler hyperparameter tuning than other machine learning algorithms such as SVM and ANN.

P. Zhang, Wu, Chen, Dai, et al. (2020) investigated interactions between a shield TBM and the ground in cross-sections and longitudinal sections. They used two machine learning algorithms to develop the prediction models, i.e., RF and LSTM. The performance of the two models was compared using a data set from the Changsha metro project. The generalizability of the models was evaluated using a data set from the Zhengzhou metro project. They concluded that the RF model performed the best when predicting discrete output (i.e., maximum settlement) with limited data. In contrast, the LSTM model performed best when predicting sequential form (i.e., the longitudinal settlement profile) and multioutput (i.e., TBM operation parameters). The LSTM model was also less sensitive to the input parameters and performed better when tested using a different data set. These results demonstrated the advantages of data-driven methods to model complex relationships between TBM operations and the progression of ground movements in the longitudinal section.

9.1.3 Objectives

The review demonstrates the opportunity to utilize data-driven methods for estimating tunneling-induced ground movements in the cross and longitudinal sections. Machine learning algorithms have been shown to be capable of finding complex interactions between the input features (e.g., the TBM operation) and the response (i.e., the estimated ground movement) without predefined assumptions. This is important in estimating tunneling-induced ground movements since they are affected by numerous factors and may not always follow the typical Gaussian shape assumption. Nevertheless, the review also reveals some gaps, as listed below.

- Most previous studies focused on predicting the maximum ground movements. Limited studies have been performed on modeling the progressions of tunneling-induced ground movements in the longitudinal section. This progression should result from interactions between different TBM parameters at different phases of ground movements (i.e., before, during, and after passing). These complex interactions are difficult to be explored without the help of machine learning algorithms.

- No solid conclusion can be drawn on the most appropriate machine learning algorithm for estimating tunneling-induced ground movements. Previous studies indicate that

most machine learning algorithms may produce reasonably good prediction performance if they are appropriately built and trained. This also implies that the selection of algorithms and their predictability may not be the central issue in developing a data-driven method for estimating tunneling-induced ground movements.

- Furthermore, many studies still reported conflicting results on the feature importance analysis. More efforts should be given to understanding the proper way to build and interpret the model, e.g., setting data set requirements, unifying data preparation methods and input features, and developing robust validation and evaluation methods. This may lead to a more generalizable model.

This study aims to develop a supervised artificial intelligence (AI) system that (i) connects TBM operation data to ground monitoring data and (ii) estimate various tunneling-induced ground movements solely based on the TBM data in real-time during tunneling. Segmentation and feature importance analyses were performed to investigate different TBM-ground interaction mechanisms and TBM control parameters that may affect ground movements during tunneling.

9.2 Data

9.2.1 Tunneling Case and Geologic Conditions

This study used a data set of the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA. More discussion of this tunneling case and the geologic conditions can be found in Chapters 5 and 7, respectively.

9.2.2 Ground Monitoring Data

The tunneling-induced ground responses along the tunnel alignment were measured using various monitoring instruments (Ning et al., 2019). Surface settlement monitoring plates and interferometric synthetic aperture radar (InSAR) were employed to monitor the surface movements. Multi-point borehole extensometers (MPBX) were employed to monitor the underground movements. To limit the scope, this study only considered underground movement obtained from the MPBX data, i.e., records at 5 ft (approx. 1.5 meters) and 10 ft (approx. 3 meters) above the tunnel crown. To limit the data size and reduce the noise, this study used the median values of the daily records. Note that the measured underground movements are the relative movement of these elevations to the movement at the ground surface. This selection resulted in a total of 159 MPBX location points. The MPBX locations along the tunnel alignment are shown in Figure 9.1.

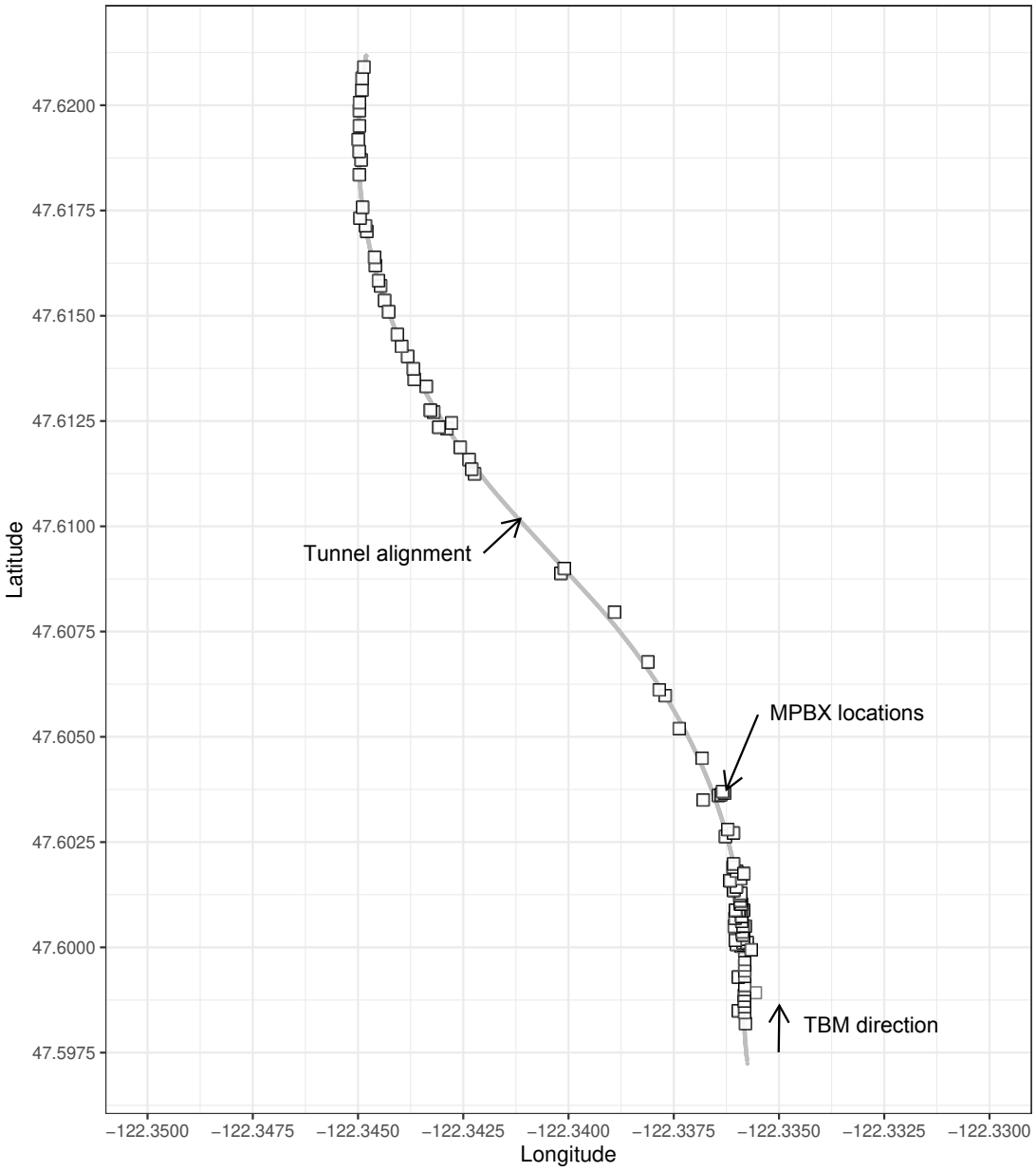

Figure 9.1: Locations of MPBX instrumentation along the tunnel alignment

9.2.3 TBM Data

Similar to Chapter 7, this study used ring-length aggregate spatial series as the observation data points and focused on the continuous data of the primary EPB TBM system, i.e., features related to (i) excavation, (ii) advancing, (iii) steering, (iv) ground conditioning, (v) earth pressure balancing, (vi) muck discharging, and (vii) tail grouting. List of EPB TBM features used in the tunnelling-induced ground movement estimation model is presented in Table 9.2. The features were selected based on the hierarchical feature selection discussed in Chapter 6. The data preparation and cleaning were done in accordance with the steps discussed in Chapter 7. This resulted in a total of 28 EPB TBM features and 1253 observation data points.

9.2.4 Data Integration

The main data frame used for the modeling was constructed by integrating the TBM and MPBX data frames. The TBM data frame consisted of the chainage head position, the corresponding time record, and the TBM operation features. The MPBX data frame consisted of the MBPX identity number, the location (latitude and longitude, chainage point), the distance to the tunnel alignment (to tunnel crown and the center of alignment), the ground movement value, and the corresponding time record. The two data frames were connected by the time record columns. This means one row of TBM data at a particular chainage head position can correspond to ground movement records from several MPBX locations. In this arrangement, the distances between the TBM head to every MPBX location can be measured, and the tunneling-induced ground movement at different locations relative to the TBM head can be obtained (Figure 9.2).

9.3 Methods

Conceptually, a model of ground movement at a particular location point can be estimated as a function of (i) the tunnel spatial geometries, (ii) the geologic condition of that point, and (iii) the TBM behaviors while passing that point. In this study, the tunnel spatial geometries consisted of the tunnel depth and distances between the estimated location point to the TBM head, the center of tunnel alignment, and the tunnel crown (Figure 9.3). The effects of tunneling on ground movements were considered negligible at 50 m ahead of the TBM face and 100 m behind the TBM tail. The TBM behaviors were represented by the 28 selected features of TBM operation data. The geologic conditions of the estimated point were not explicitly provided since this information has been implicitly contained in the interactions among the TBM features (Chapter 7). This resulted in a total of 32 input features for the models.

Table 9.2: List of EPB TBM features used in the tunnelling-induced ground movement estimation model.

No.	Feature	Unit	Sub-Component	Group
1	Cutter torque	kNm		Excavation
2	Cutter rot. speed	rpm		
3	Copy cutter stroke	mm		
4	Copy cutter position	deg		
5	Thrust force	kN		Advancing
6	Thrust stroke	mm		
7	Advance rate	mm/min		
8	Shield roll	deg	Front shield	Steering
9	Shield pitch	deg	Front shield	
10	Shield yaw	deg	Front shield	
11-12	Thrust stroke difference	mm	Left-Right, Top-Bottom	
13-14	Deviation at head	mm	Vertical, horizontal	
15	Foam volume	m^3/m		Ground conditioning
16	Foam pressure	kPa		
17	Polymer volume	m^3/m		
18	Polymer pressure	kPa		
19	Slurry volume	m^3/m		
20	Slurry pressure	kPa		
21	Additive volume	m^3/m		
22	Chamber pressure	kPa		Earth pressure balancing
23	Screw rot. speed	rpm		Muck discharging
24	Screw pressure	kPa		
25	Belt muck volume	m^3/m		
26	Belt muck weight	ton		
27	Grout volume	m^3/m		Backfill grouting
28	Grout pressure	kPa		

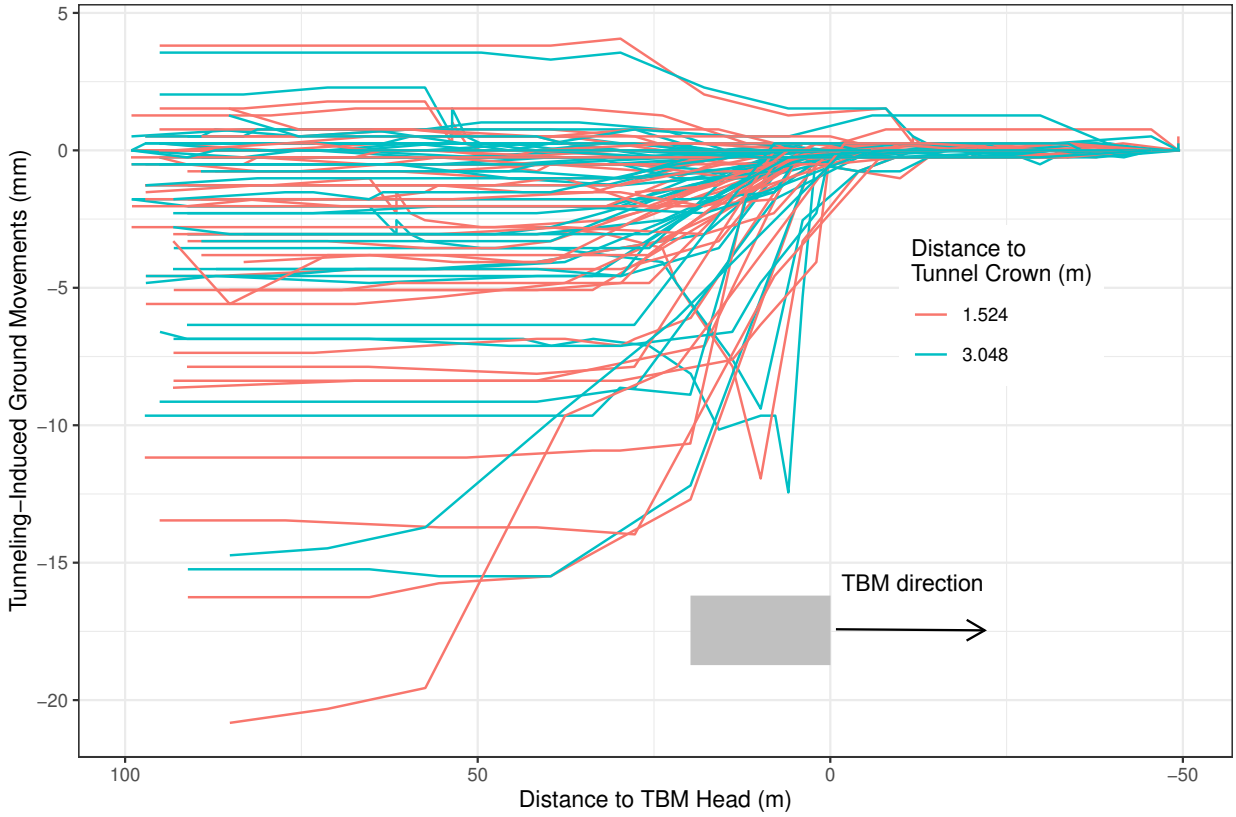

Figure 9.2: Tunneling-induced ground movement data relative to the TBM head position.

9.3.1 Model Setup

The tunneling-induced ground movement estimation system was developed using two data-driven models with different data-splitting schemes, i.e., (i) static random model and (ii) dynamic sequential model.

(i) **Static Random Model**. Similar to the geologic interpretation model (Chapter 7), the static random model was developed as a benchmark model that follows the standard practice in machine learning analysis. In this model, the integrated data frame was split randomly into training and testing data sets in a ratio of 70:30. The prediction results were evaluated using the mean absolute errors (MAE).

(ii) **Dynamic Sequential Model**. The dynamic sequential model was developed to be a model that can be implemented in an actual tunneling process, where the ground movements were estimated sequentially based on the relationship between the previously recorded ground movements and TBM operation data. The predictions were

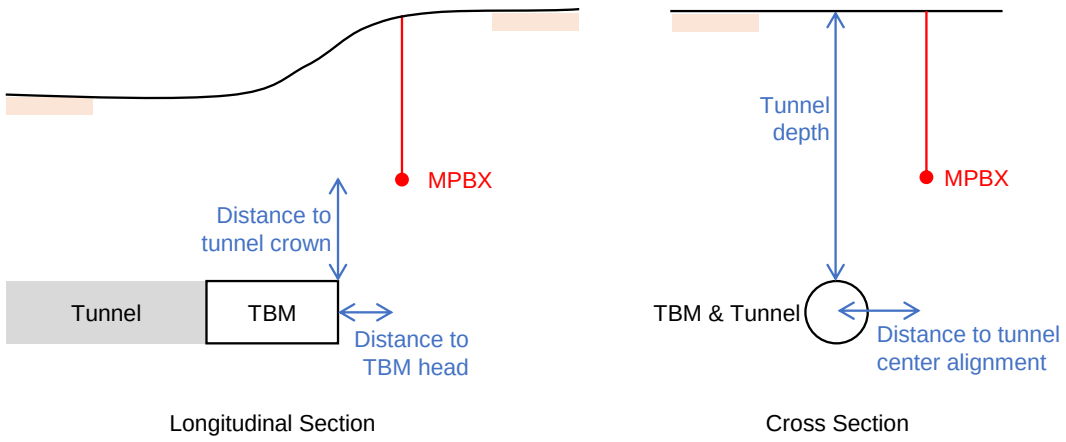

Figure 9.3: Schematic illustration of tunnel spatial geometries.

continuously performed in an area adjacent to the TBM, i.e., between 50 m ahead and 100 m behind the TBM face. This area was envisaged to be the area that could be affected by the TBM behaviors. A schematic diagram of the static and dynamic models is shown in Figure 9.4.

9.3.2 Prediction Methods

The ground movement model was developed using two prediction methods, i.e., the Ordinary Least Squares (OLS) and Random Forests (RF). Similar to Chapter 5, OLS was selected to represent parametric linear prediction methods, and RF was selected to represent nonparametric nonlinear machine learning methods. This study used the RF fast implementation in C++ and R (Wright & Ziegler, 2017). More discussion of these prediction methods and their mathematical bases can be found in Chapter 4.

9.3.3 Model Hyperparameters

Similar to Chapter 7, hyperparameter analysis was performed on the static random model. The analysis was performed to investigate the effects of the hyperparameters on the model performance and to select the best hyperparameter configuration. The analysis was also conducted in five times repeated 10-fold cross validation (CV), but with only three RF hyperparameters, i.e., (i) ntrees, (ii) mtry, and (iii) minimum node size. Since ground movement estimation is a regression problem, the estimated response variance was used as the

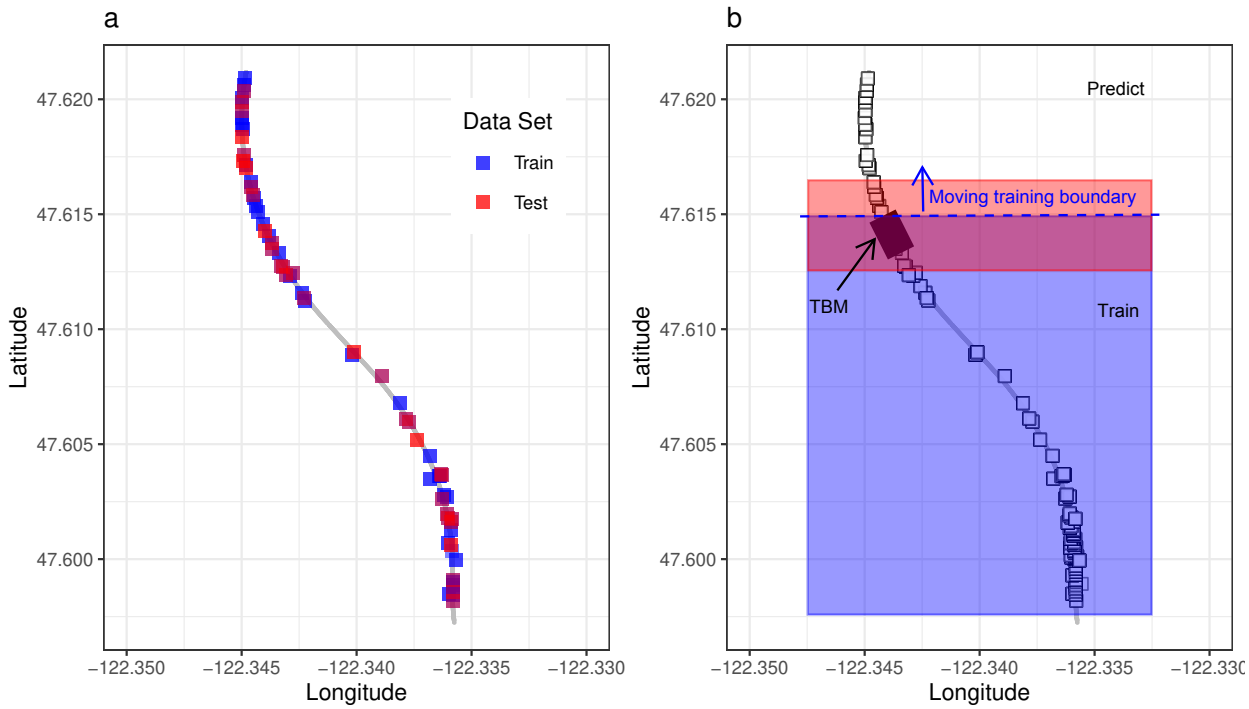

Figure 9.4: Schematic diagram of the (a) static random and (b) dynamic sequential models.

split rule (Wright & Ziegler, 2017). More discussion on the hyperparameters can be found in Chapter 6. The model performance was evaluated using the out-of-bag samples (OOB).

9.3.4 Feature Importance

Feature importance analysis was performed to investigate the role of TBM features in tunneling-induced ground movements. The analysis was done using the permutation-based feature importance. More discussion on the permutation feature importance can be found in Chapter 5.

9.3.5 Segmentation Analysis

Segmentation analysis was performed to investigate different ground response mechanisms relative to TBM positions. The algorithm of the segmentation analysis is presented in Table 9.3. In principle, the analysis was conducted by dividing the ground movement response into

several segments. A segment was defined as a distance where an RF model produced negligible prediction errors. The segment was analyzed using a moving boundary line throughout the ground responses. For example, the first segment was from 50 m ahead of the TBM face to a boundary line where errors within this segment were negligible. The next segment started from the end of the previous segment's boundary. This analysis was repeated until the final segment was obtained, where the end boundary of the segment was 100 m behind the TBM face.

Table 9.3: Algorithm of the segmentation analysis.

```
Input = training and testing data from randomly split instrument IDs
Set StartBoundary, EndBoundary, and ErrorThreshold
while StartBoundary < EndBoundary do
    Initialize SegmentMAE
    for b in StartBoundary to EndBoundary do
        Slice data from StartBoundary to b
        Get training and testing data within the slice
        Train model
        Initialize InstrumentMAE
        for each instrument in the testing data do
            Get prediction
            Get InstrumentMAE and store the result
        end for
        Get SegmentMAE as InstrumentMAE at boundary b
    end for
    Get SegmentBoundary as the last b where MAE < ErrorThreshold and store the result
    Update StartBoundary as SegmentBoundary at the current iteration
end while
```

The mean absolute error (MAE) of 0.2 mm was selected as the error threshold, where the model's error was assumed to be negligible. This threshold was selected to allow better interpretability of the model. Note that an ideal threshold of 0 mm MAE resulted in a very sensitive analysis that produced too many ground response segments. This might be due to various factors, such as noises in the measurement data.

9.4 Results and Discussion

9.4.1 Effects of Hyperparameters

Figure 9.5 presents the effects of RF hyperparameters on the OOB predictions. The left panel presents the effect of the ntrees parameter on MAE in different mtry parameters, with a constant minimum node size parameter of 1. The figure shows that a low number of ntrees (e.g., ntrees < 200) produced substantially low prediction performance. Increasing the ntrees improved the prediction performance until it reached the threshold value and stabilized. This result agrees with previous studies, either using the EPB TBM data set (Apoji et al., 2022b) or other data sets (Probst et al., 2019). In this study, ntrees of 500 were selected for the model to ensure high prediction performance and reasonable computation costs.

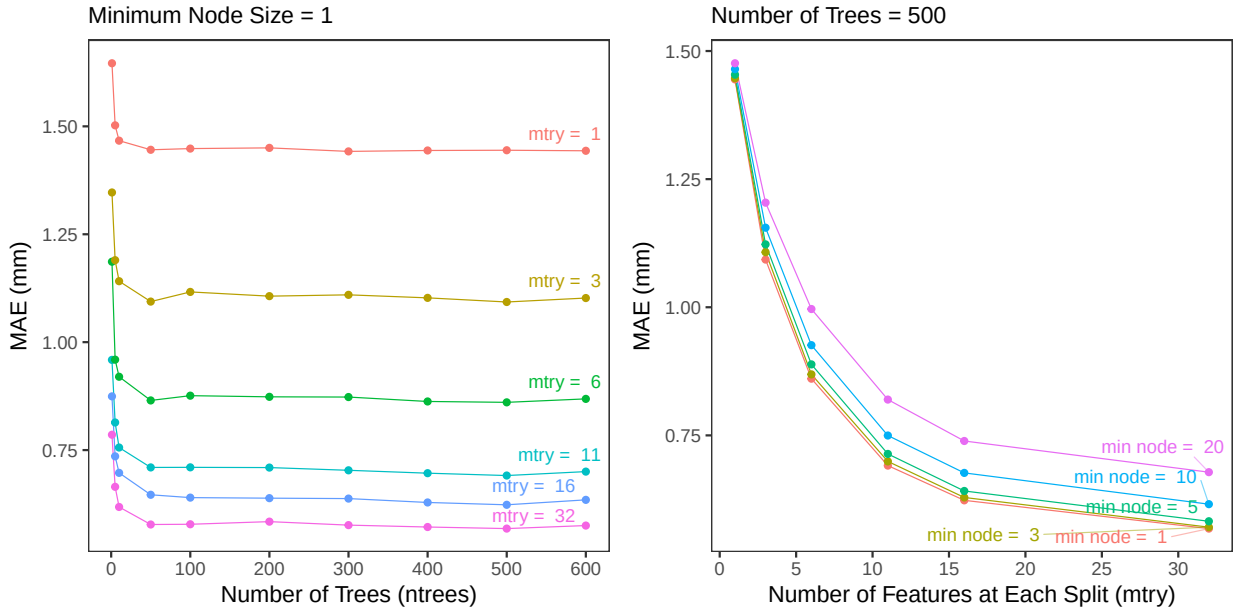

Figure 9.5: Effects of RF hyperparameters on the OOB samples of the training data set.

The right panel presents the effect of the mtry parameter on the MAE in different minimum node size parameters, with a constant ntrees parameter of 500. This figure shows that higher numbers of mtry produced lower MAE, which means higher prediction performance. The best prediction performance was produced when the mtry was equal to the number of predictors (i.e., p = 32). This is an interesting finding since the best prediction performance for an RF regression is typically produced by mtry = p/3 (e.g., Apoji et al. (2022b) and Probst et al. (2019)). The minimum node sizes produced fewer effects on the prediction performance. The best prediction performance was produced at the minimum node sizes 1

and 3. The large mtry and small minimum node size values indicate that more complex trees were required for the predictions. This hyperparameter configuration reveals the complexity of the relationship between TBM and ground responses. In this study, mtry of p = 32 and the minimum node size of 1 were selected for the models.

9.4.2 Static Random Model Predictions

Figure 9.6 presents tunneling-induced ground movement predictions at several selected locations from the testing data set. The x-axes show the distance from the measurement points (MPBX locations) to the TBM head (in meters). The y-axes show the induced ground movements (in mm). The measured ground responses are shown in black. The OLS and RF predictions are shown in red and blue, respectively. The figure presents different types of induced ground responses, i.e., settlement (left column panels), heaving (middle column panels), and relatively stable responses (right column panels). The selected responses show the variability of tunneling-induced ground movements. Note that the simplified model, such as the Gaussian settlement profile-based models, may not capture this variability.

These figures show that the RF model could predict tunneling-induced ground movements solely based on the TBM data. The predicted ground responses are in reasonably good agreement with the measured responses. In contrast, the OLS model could not reconstruct the measured responses, with substantial discrepancies in both pattern and magnitude of the ground movements. This result indicates the presence of nonlinear interactions between the TBM and the ground responses. This nonlinearity could not be captured by OLS, which is a parametric model that strictly constrains the fitting to be linear. This result suggests the value of nonparametric machine learning methods to model data sets with complex interactions.

Figure 9.7 presents the absolute error (y-axis, in mm) of every monitoring point relative to the TBM head distance (x-axis, in meters) from all predictions in the testing data set. The absolute errors of the OLS and RF models are shown in red and blue, respectively. The Locally Estimated Scatterplot Smoothing (LOESS) lines are shown to visually represent the scattering data points. Note that LOESS is essentially a generalization form of moving average and polynomial regression (Garimella, 2017). The figure shows that the overall RF model performance was better than the overall OLS model. The RF models produced relatively small errors at points ahead of the TBM (before passing).

In contrast, the OLS model produced substantial errors in this segment. It even failed to capture zero ground movements at 50-meter distances ahead of the TBM. Note that all the training data at this point were set to zero movements. Both models' errors increased during TBM passing and stabilized after the passing. This indicates different mechanisms of TBM and ground interactions before, during, and after TBM passing.

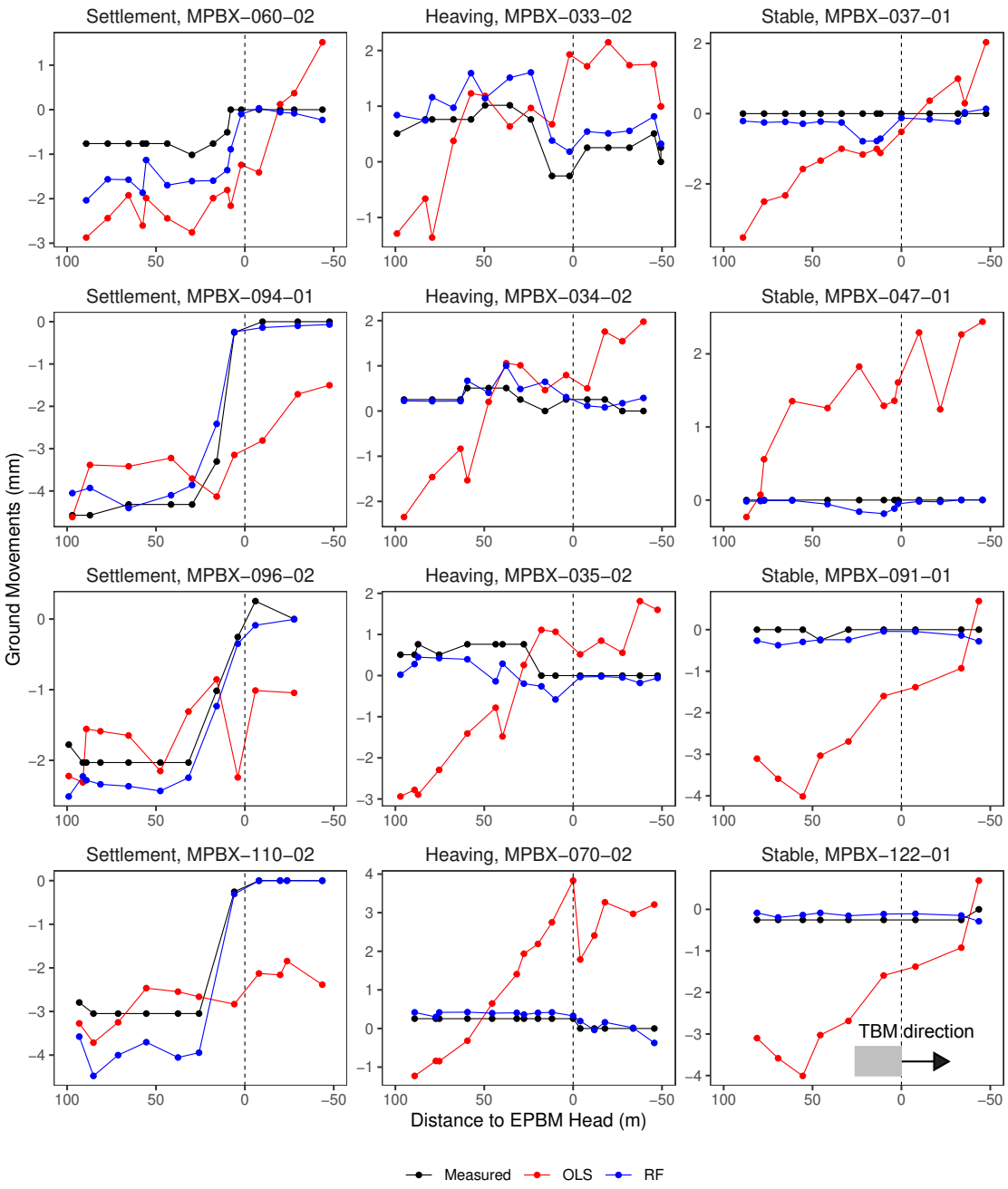

Figure 9.6: Tunneling-induced ground movement predictions with different types of responses at selected MPBX testing data set: settlement (panels in the left column), heaving (the middle column), and relatively stable responses (the right column).

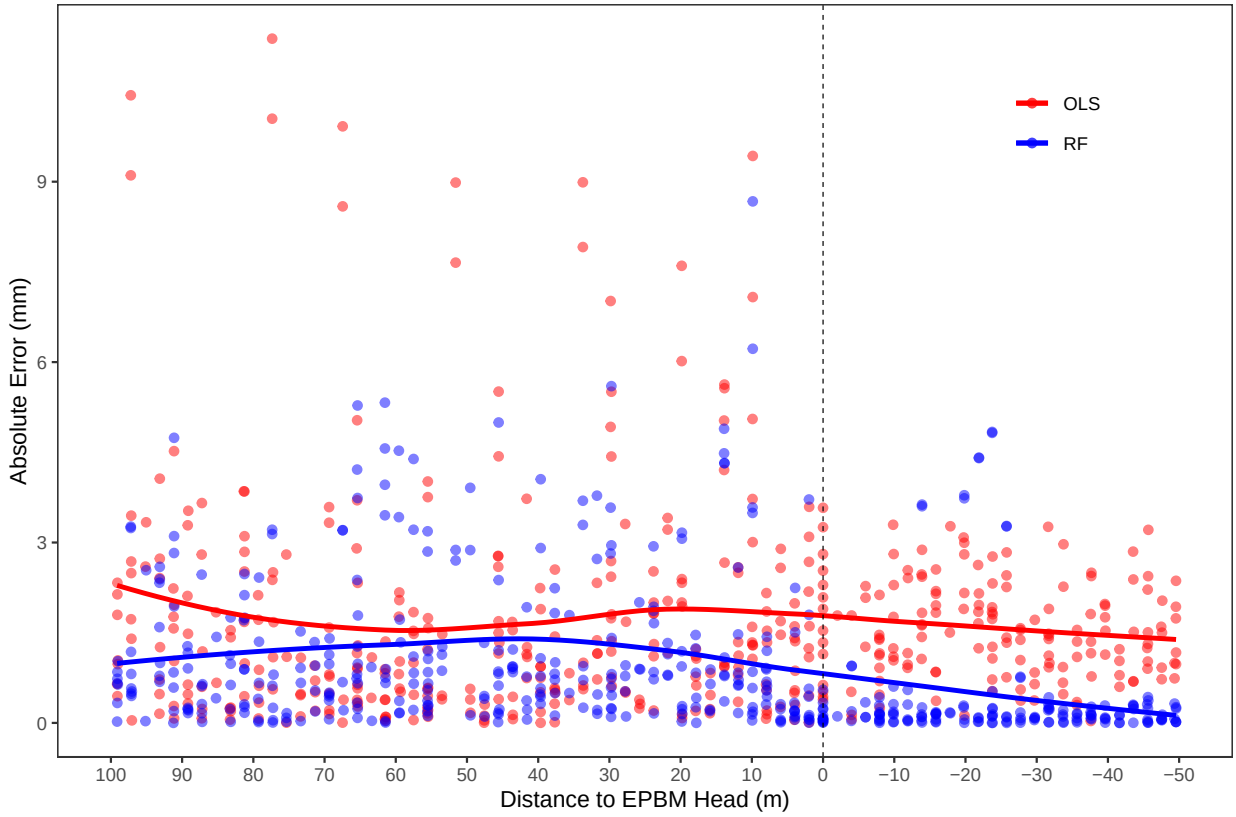

Figure 9.7: Absolute error of every monitoring point relative to the TBM head distance from all predictions in the testing data set.

9.4.3 Dynamic Sequential Model Predictions

The static random model can provide some understanding of how the models perform in different ground responses and using different prediction methods. However, the model cannot be implemented in real tunneling cases. This is where the dynamic sequential model is required. Figure 9.8 presents a simulation of real-time tunneling-induced ground estimation based on TBM data. The simulation was performed from chainage 26000 to 28000 ft, and at about 3 m distance above the tunnel crown. The colored points on the top panel represent the measured ground movements obtained from MPBX records. The colored lines on the bottom panel show the TBM operation data during tunneling at a particular TBM location. These data were connected by the proposed AI system (i.e., the dynamic sequential model) to estimate the ground movements along the longitudinal section. The estimated ground movements are represented in the black points. The red-shaded area indicates the region in active predictions (i.e., 50 m ahead and 100 m behind the TBM face).

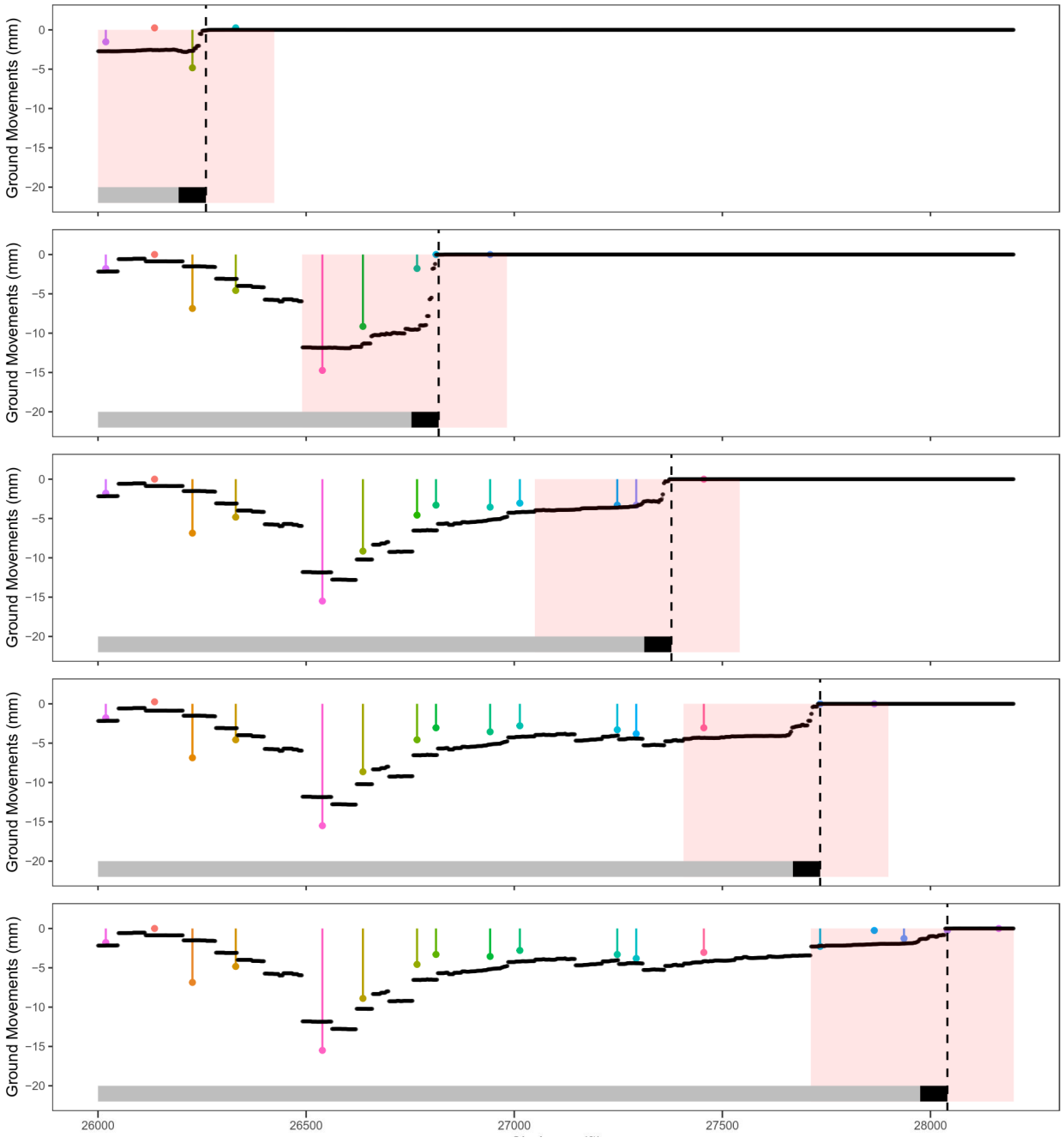

Figure 9.8: Simulation of real-time tunneling-induced ground movement predictions using the dynamic sequential model.

The simulation demonstrates that the system could produce a continuous estimation of ground responses along the chainage, covering the gaps of ground movement information provided by the discrete monitoring instruments. Furthermore, the system could also model the progression of longitudinal ground movements during TBM passing, taking into account the changing TBM control parameters and the measured ground monitoring. Figure 9.9 shows the average mean absolute errors (MAE) of the predictions at every TBM location along the chainages. The figure shows that the MAE tends to decrease along the chainages. This might be due to more training data being available for the model, suggesting that the system can perform better along the tunnel chainages during the TBM advance.

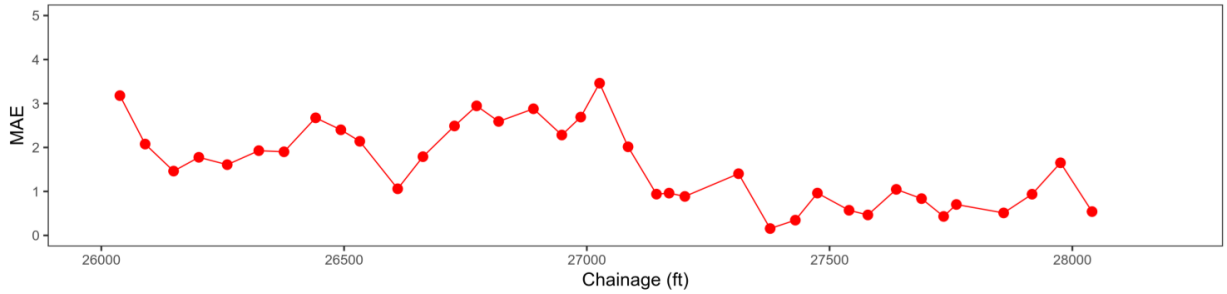

Figure 9.9: Mean absolute error of predictions along the tunnel alignment.

9.4.4 Three-Dimensional Expansion

Figure 9.10 presents a three-dimension (3D) ground movement estimation at about 3 m distance above the tunnel crown, when the TBM was at chainage 28000 ft. This prediction was produced by setting two-dimension (2D) horizontal grids at the specified elevation as thrprediction points. The figure shows reasonable ground movement results around the TBM region, indicating the largest ground settlement occurred on top of the TBM centerline and diminished with more distance. This suggests another advantage of machine learning, where the prediction can flexibly be expanded to more dimensions.

9.4.5 Ground Response Segmentation

Figure 9.11 presents ground response segments based on the segmentation analysis. Each segment is shown as a distance between two dashed boundary lines (red), where the model produces negligible errors (MAE $<$ 0.2 mm). In this case, the first segment was produced from -50 to 3 m to the TBM head. This segment represents the ground response ahead of the TBM. The second, third, and fourth segments were produced from 3 to 7 m, 7 to 13 m, and 13 to 21 m to the TBM head, respectively. These segments represent the ground response

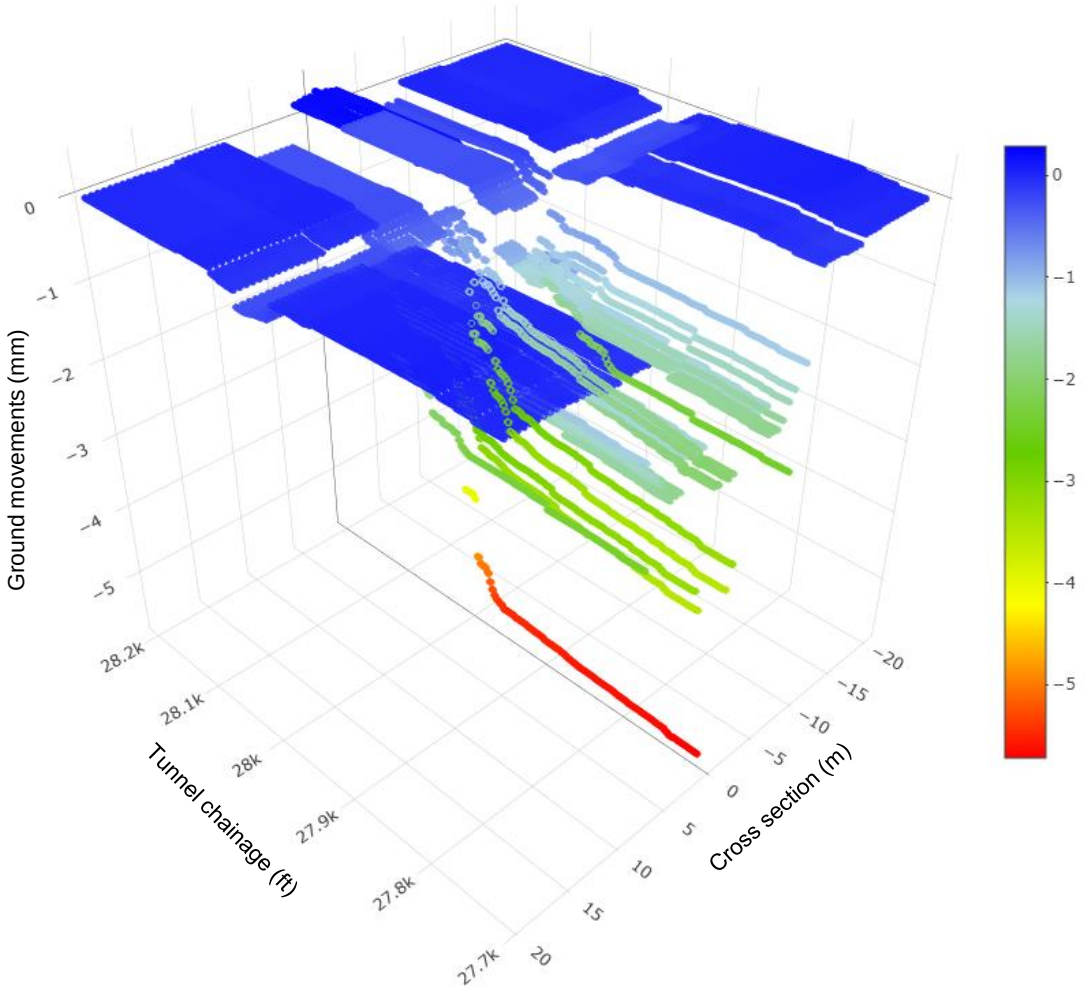

Figure 9.10: Three-dimensional expansion of the tunnelling-induced ground movement predictions.

during the TBM passing. The fifth segment was produced from 21 to 31 m, representing the post-tunneling ground response over the newly constructed tunnel lining. The sixth and seventh segments were from 31 to 95 m and 95 to 100 m to the TBM head, respectively. These segments represent post-tunneling ground responses in the longer term. The smaller segment size during TBM passing may indicate a more complex ground response mechanism during this period.

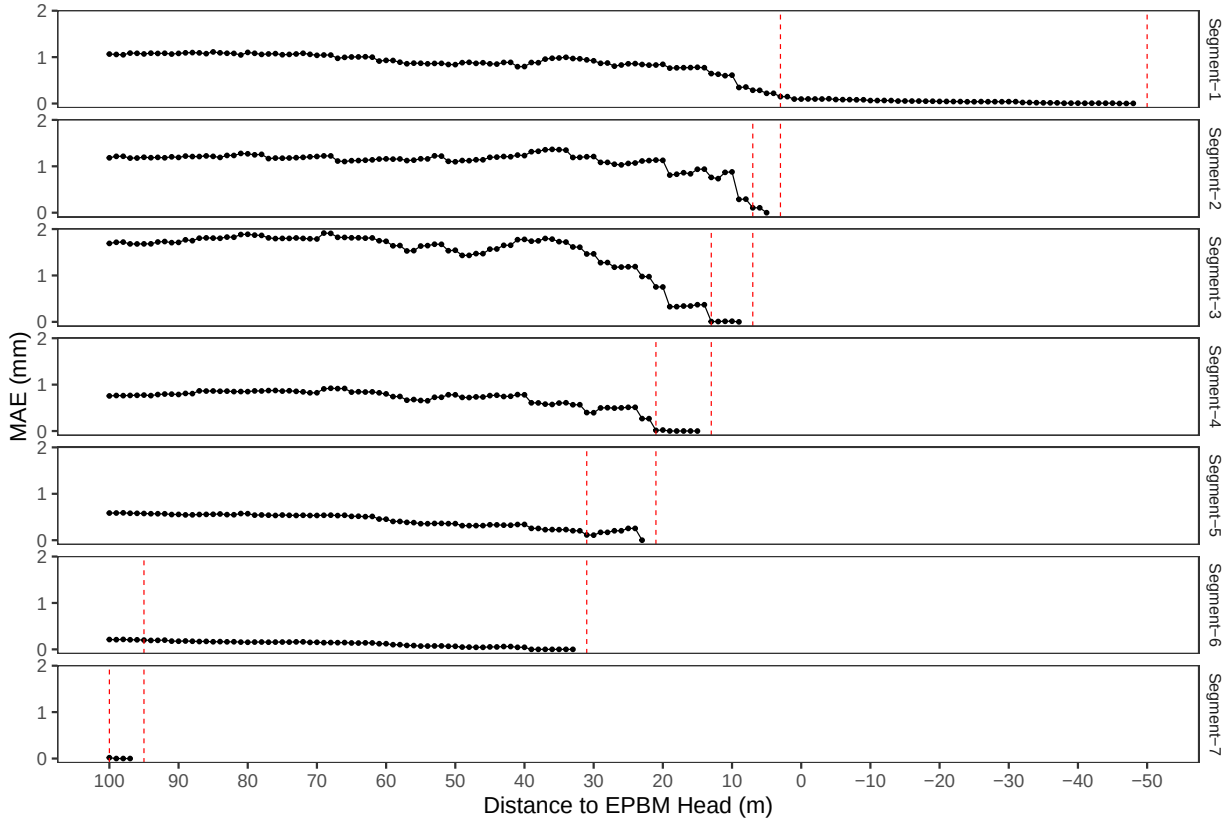

Figure 9.11: Ground response segments based on the segmentation analysis. Each segment produces MAE < 0.2 mm and is shown as a distance between the two red dashed boundary lines.

9.4.6 Feature Importance for Overall Response

Figure 9.12 presents the feature importance rank of the overall ground response model, covering from -50 m to 100 meters distance to the TBM head. The rank shows that spatial geometries are the key parameters in estimating the ground responses (i.e., the distance

from the point of interest to the center of tunnel alignment, the tunnel depth, and the distance from the point of interest to the TBM head and crown). This is expected as the conventional tunneling-induced ground movement prediction methods have also considered the geometrical information in the models.

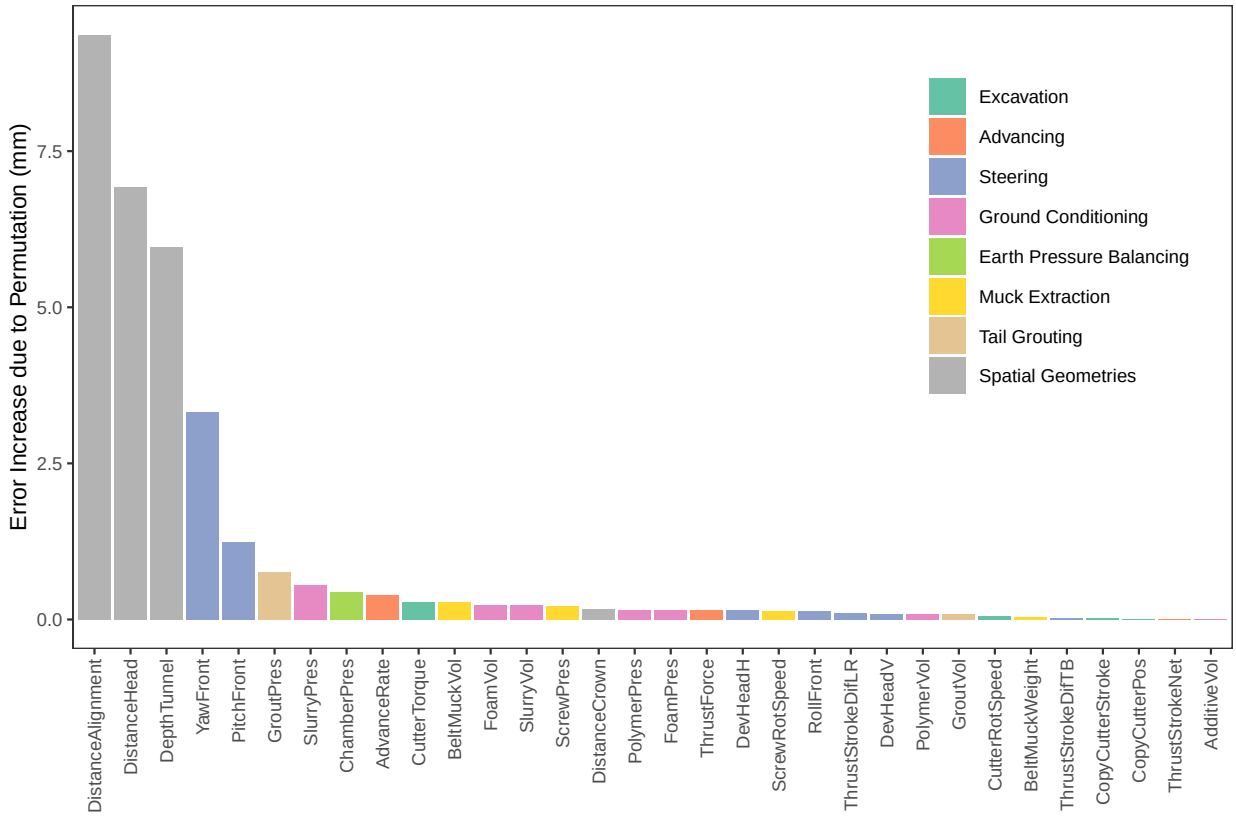

Figure 9.12: Feature importance rank of the overall ground response model

However, interestingly, the figure shows that features related to steering control (i.e., pitch, yaw, and deviation) and chamber pressure are in high-importance ranks. Note that these features are not commonly included as specific parameters in conventional prediction methods. This suggests that TBM control parameters substantially govern the tunneling-induced ground movement and should be considered more carefully in the prediction models.

9.4.7 Feature Importance for Segmental Response

Figure 9.13 presents simplified ground response segments (top panel) and the feature importance ranks of each segment (bottom panels). The simplified ground response segments

were approximated based on the segmentation analysis result. The feature importance ranks only show the top 8 features, to focus on the most critical features in the predictions. The figure shows that each ground response segment produced different feature ranks, indicating different TBM-ground interaction mechanisms. The simplified ground response segments are described below.

- Segment 1 represents the ground response ahead of the TBM. This segment appears to be dominated by geometrical parameters. This suggests that the ground response mainly depends on the distance to the TBM.
- Segment 2 represents the ground response above the TBM front shield. This segment indicates a strong influence of the screw-related features. Note that these features can be related to how the operators regulate the chamber pressure. This suggests the critical role of pressure control on the ground response.
- Segment 3 represents the ground response above the TBM articulation and rear shield. This segment shows domination from features related to the steering control (i.e., deviation and thrust stroke difference). This suggests that appropriate shield attitude controls can limit the induced ground movements. This means erratic shield movements may contribute to excessive ground movement responses.
- Segment 4 represents the post-tunneling ground response above the newly constructed tunnel lining. This segment appears to be dominated by the geometrical parameters and the muck volume. This may be related to ground loss. The more the ground loss, the larger the tail void between the excavated ground and the segment. And the larger the tail void, the more the induced ground movements.
- Segment 5 represents the longer-term post-tunneling ground behavior. This segment is strongly dominated by features related to the ground conditioning system, i.e., the polymer volume. Note that polymer volume is typically injected to minimize the stickiness of clayey soils (Todaro et al., 2021). This indicates that the ground response mainly depends on the soil type, e.g., more long-term settlement is expected on clayey soils due to the soil consolidation.

9.5 Conclusions

This study has proposed an AI system to connect TBM operation data to the ground monitoring data. The proposed system can estimate various types of tunneling-induced ground movements in real-time during tunneling solely based on the EPB TBM features and the tunnel spatial geometries. Thus, it can be employed as a tool to control the TBM and limit the induced ground movements. This also enables quantitative investigation of the interactions between TBM control parameters and the ground responses at different locations (i.e.,

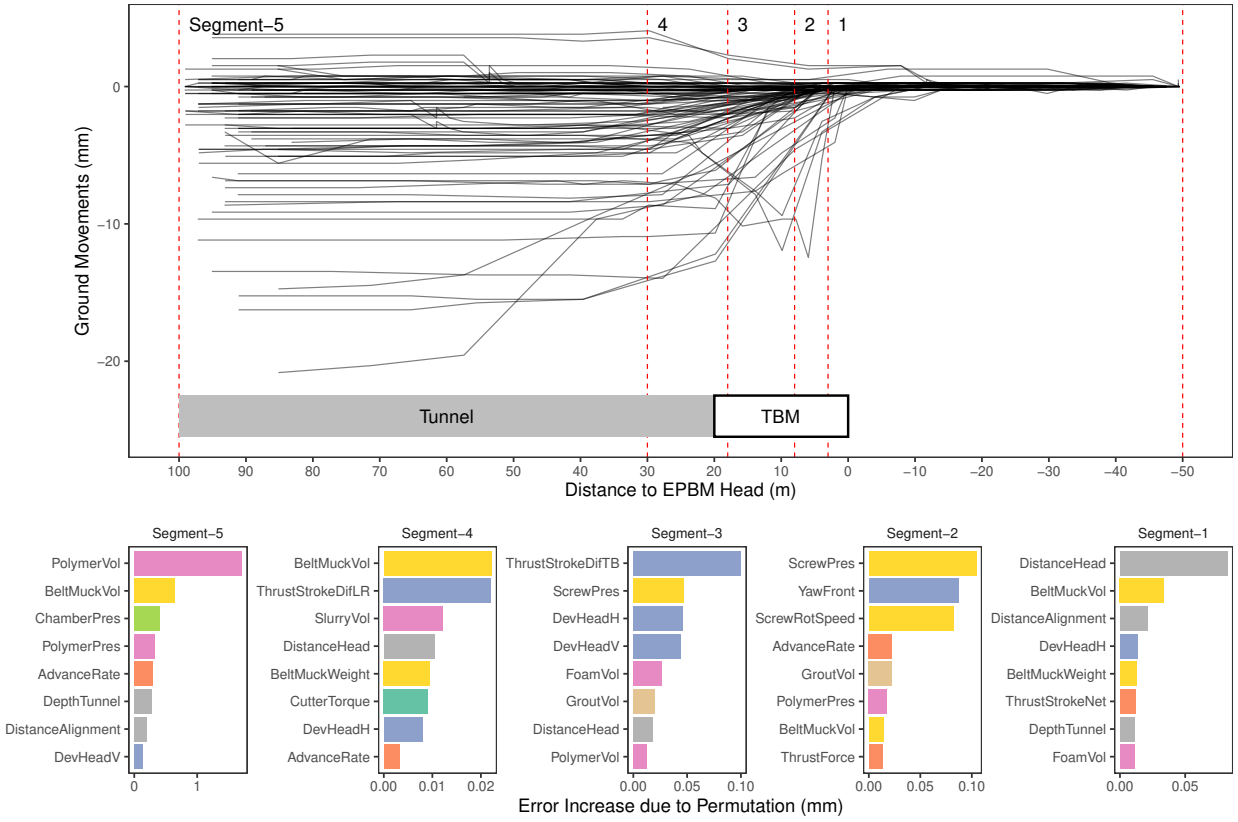

Figure 9.13: Feature importance rank of each ground response segment.

before, during, and after TBM passing). The main findings of this study are summarized in the following.

- **Estimating ground movements**. The proposed system can estimate any shape of ground movements solely based on TBM operation data and tunnel spatial geometries without any prior assumption on the ground movement shape, geologic material parameters, and expected ground loss. Furthermore, unlike most previous data-driven models, the model can perform real-time estimation of tunneling-induced ground movements during tunneling in both longitudinal and cross sections. The nonparametric and nonlinear machine learning prediction model produces better estimations than the parametric linear regression model, indicating the complexity and nonlinearity of TBM-ground interactions.

- **Segmentation analysis**. The segmentation analysis shows tunneling-induced ground movements can be divided into several ground response segments relative to the TBM

positions, indicating different TBM-ground interaction mechanisms at each ground response segment. This result quantitatively confirms the theoretical segmentation proposed in previous literature (e.g., JSCE (2016)).

- **Feature importance**. The feature importance analysis reveals that each segment may have different controlling parameters. This information can be useful to determine which control parameters need to be specified to limit ground movements during tunneling. Features related to the steering and pressure controls appear to influence the induced ground movements during TBM passing strongly. These features are not typically considered in conventional tunneling-induced ground movement estimation methods.

Based on these findings, more work can be done to further develop the proposed tunneling-induced ground movement estimation system. This may include as the following. (i) Incorporating records from multi-instruments with a finer resolution of both TBM and ground monitoring data. (ii) Examining the effects of the threshold value in the segmentation analysis. (iii) Investigating the effects of different machine learning algorithms and their behavior in extrapolation problems (as discussed in Chapter 7). (iv) Investigating the change of hyperparameter configuration along the tunnel alignment (as discussed in Chapter 7), and (v) Exploring more model interpretation tools to obtain a more solid conclusion on the role of TBM features in tunneling-induced ground movements (as discussed in Chapters 5 and 6).

Chapter 10

Modeling TBM Excavation Process

Some contents of this chapter have been published in Apoji, D., Fujita, Y., and Soga, K. (2022). *Exploring Interactions among EPBM Features using Bayesian Networks.* World Tunnel Congress. International Tunneling and Underground Space Association.

10.1 Introduction

10.1.1 Background and Related Works

Tunneling is a complex process due to unpredictable interactions between geologic conditions and tunnel boring machine (TBM) behaviors. To control a TBM, the operators have to continuously make real-time interpretations, judgments, and decisions. For example, in soft ground tunneling using an earth pressure balance (EPB) TBM, the operators have to interpret the encountered geologic conditions, decide the excavation control parameters, inject foam to condition the encountered ground, balance the chamber pressure, and navigate the shield to achieve the correct tunnel position (JSCE, 2016; Maidl et al., 2013). Therefore, tunneling performance, quality, and safety are strongly reliant on the skill and experience of the TBM operators.

Typically, TBM operators are guided by instruction sheets that are prepared by tunnel engineers prior to each tunneling excavation phase. However, during the excavation phase, the operators have to make judgments and adjustments according to the numerous operational data collected by TBM sensors. For humans, continuously making real-time interpretations, judgments, and decisions in various control tasks based on numerous streaming sensor data is not straightforward and may lead to inconsistent results. A more systematic approach is required to utilize TBM operation data to support the tunneling process (Garcia et al., 2021).

Studies to address this challenge have been emerging in recent years. Especially due to the growing number of data collected by TBM sensors, the increase in computing power, and the emergence of machine learning techniques. For example, many studies have been attempted

to predict or forecast the ahead geologic conditions (e.g., Apoji et al. (2022b), Erharter et al. (2020), Sousa and Einstein (2012), Q. Zhang et al. (2019), and Zhao et al. (2019)), predict the tunneling performance (e.g., Martins and Miranda (2013) and Mokhtari and Mooney (2020)), and predict the tunneling-induced settlement (e.g., R.-P. Chen et al. (2019) and W. G. Zhang et al. (2021)). However, most past studies focus on prediction performance. Less focus has been paid to understanding the features (measured variables used as predictors), the causal inference, and the generalization of the data (Sheil, Suryasentana, Mooney, et al., 2020).

10.1.2 Objectives

This study aims to introduce a method to explore interactions and causal relationships of EPB TBM features in a systematic and quantitative approach. Bayesian Networks (BN) and a structure learning algorithm were used to model the interactions of EPB TBM features along the tunnel alignment. The relationship between the feature dependencies and EPB TBM control mechanisms was explored. The effects of providing explicit information about geologic conditions (i.e., soil labels) in the interactions were also evaluated.

10.2 Data

10.2.1 Tunneling Case and Geologic Conditions

This study used a data set of the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA. More discussion of this tunneling case and the geologic conditions can be found in Chapters 5 and 7, respectively.

Similar to Chapter 7, the data was labeled according to soil types based on the borehole information and physical characteristics described as engineering soil units (ESU) in the geotechnical baseline project reports (WSDOT, 2010a, 2010b, 2010c). The boreholes were envisaged to represent about 25 feet (about 7.62 meters) radius distance of their vicinity. Five labels were created to represent mixtures of the ESU at each borehole location, namely,

(i) cohesive clay and silt (CCS),

(ii) predominantly cohesive clay and silt with layers of cohesionless silt, fine sand, sand, and gravel (CCS-CSGCSF),

(iii) predominantly cohesionless silt, fine sand, sand, and gravel with layers of cohesive clay and silt (CSGCSF-CCS),

(iv) predominantly cohesionless silt, fine sand, sand, and gravel with layers of till/till-like deposits (CSGCSF-TLTD), and

(v) mix of cohesionless silt, fine sand, sand, and gravel; cohesive clay and silt; till/till-like deposits (TLTD-CSGCSF-CCS).

10.2.2 Data Description and Preparation

To capture the key interactions, this study focused on features related to excavation processes, i.e., features related to cutter head, thrust, ground conditioning, and earth pressure balancing systems (Table 10.1). To allow more generalization and interpretability of the results, data from the same sensor types were summed (e.g., for sensors related to volume, force, and pressure) or averaged (e.g., for sensors related to speed, and length) together.

This study used statistical ring data as the observation data points. In other words, observation records of a feature within a ring were represented by a data point. Depending on the data characteristics, the data point was obtained by taking the final observation (e.g., for features related to volume, length), or taking the average observation (e.g., for features related to force, pressure, speed) along the ring.

The data was prepared by removing (i) observations of non-advancing stages, (ii) observations with an incomplete set of features, i.e., missing values, (iii) observations with other erroneous records such as data outliers and duplications. The data outliers were determined using a modified interquartile range (Maher, 2015).

Table 10.1: List of EPB TBM features used in the tunneling excavation process modeling

No.	Feature	Unit	Group
1	Cutter torque	kNm	Excavation
2	Cutter rotation speed	rpm	
3	Cutter head force	kN	
4	Thrust force	kN	Advancing
5	Thrust stroke	mm	
6	Advance rate	mm/min	
7	Penetration rate	$mm/rot.$	
8	Foam volume	m^3/m	Ground conditioning
9	Air volume	m^3/m	
10	Foam liquid volume	m^3/m	
11	Foam agent volume	m^3/m	
12	Chamber pressure	kPa	Earth pressure balancing
13	Screw rotation speed	rpm	Muck discharging

10.3 Methods

10.3.1 Bayesian Networks

Bayesian network (BN) is a probabilistic graphical model that uses a directed acyclic graph (DAG) to compactly represent the joint distributions of a set of variables and their conditional independencies (Koller & Friedman, 2009). In a BN graph G, the joint probability distribution P over variables $x_1, \ldots, x_n$ can be expressed as a product of the individual conditional probability distributions (CPDs),

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | Pa_{x_i}^G). \tag{10.1}$$

$Pa_{X_i}^G$ denotes the parents of x_i in G. And G encodes the local independencies $I_l(G)$ which state that each node x_i, given its parent nodes Pa, is conditionally independent of its non-descendants ND, as expressed by

$$For\ each\ x_i : (x_i \perp ND_{x_i} | Pa_{x_i}^G). \tag{10.2}$$

By using this method, the global complexity of a high-dimensional distribution can be reduced into local probabilistic models and dependencies.

10.3.2 Structure Learning

BN graphs can be constructed manually based on domain knowledge or can be inferred based on data using Bayesian network structure learning (BNSL) algorithms. In general, BNSL can be divided into classes of (i) constraint-based method, (ii) score-based method, and (iii) hybrid method.

In this study, the BN graphs were constructed using the score-based method. Tabu search algorithm (Glover & Laguna, 1993) was used to iteratively construct and modify the graphs toward an optimum network model. The score of the constructed graphs was computed based on the Bayesian information criterion (BIC), defined as

$$score\ BIC(G) = \log L(\hat{\theta}_G) - 1/2 \log(n)\, dim(G), \tag{10.3}$$

Where L is the likelihood function, $\hat{\theta}_G$ is the parameter of the graph, n is the sample size, and $dim(G)$ is the model dimension or the number of parameters of the whole network. This study was performed using a BN framework implementation by Scutari and Ness (2020).

10.3.3 Model Setup

First, a simple BN graph was constructed using the structure learning algorithm to model the relationship of ground conditioning volumes, i.e., foam volume, foam liquid (solution) volume,

and foam agent (surfactant) volume. This was performed to demonstrate the capability of BNSL to model a true physical mechanism that can be easily validated.

Subsequently, more complex BNSL graphs were constructed to model interactions of EPB TBM excavation features. To evaluate the effect of providing explicit information about geologic conditions, the graph models were constructed with and without soil labels included as a feature. To evaluate the stability and the change of feature interactions along the tunnel alignment, the graphs were sequentially retrained for each ring chainage. Meaning that a graph was constructed at each ring chainage based on data from all the previous ring chainages.

No prescribed graph structure was provided to the algorithm. The strength of the node connections (links) and the directions were ensured probabilistically by repeating each graph construction 200 times.

10.4 Results and Discussion

10.4.1 Ground Conditioning System

A simple BN graph was constructed to examine the capability of the structure learning algorithm to model the true physical mechanism of ground conditioning volumes. The graph was constructed by the algorithm without any prescribed network skeleton. Figure 10.1 shows the constructed BN graph at the end of chainage (using all observations in the dataset). This figure shows that BNSL successfully modeled the true physical mechanism of the ground conditioning volumes. The graph shows that the foam volume had conditional dependencies on the foam liquid volume and the air volume. And the foam liquid volume had a conditional dependency on the foam agent volume. This represents the true ground conditioning physical mechanism where foam is a product of foam liquid plus injected air, and foam liquid is a product of foam agent plus water.

To evaluate the stability of these interactions, the graph was sequentially retrained for each ring chainage location using all the previous ring chainages. Normalized data (for visualization) of the ground conditioning volumes, probabilistic strength of the network links, and probabilistic direction of the links along the tunnel alignment are shown in Figure 10.2(a), (b), and (c), respectively. Probabilistic strength of 1 means a link was always connected in 200 repetitions of graph construction. The probabilistic direction of 1 and 0 means a link always had direction as shown in Figure 10.1 and always had its reverse direction, respectively, in the repetitions.

This figure shows that the links were consistently connected along the tunnel alignment. This indicates the robustness of this method in capturing the interactions of these features, even when the data distributions were altered as the data were continuously added during tunneling. However, the probabilistic direction of the links shows more mixed results. The figure shows that the link directions could be switched. It seems the algorithm could not decide which feature was the parent or the child node since the foam agent volume can be

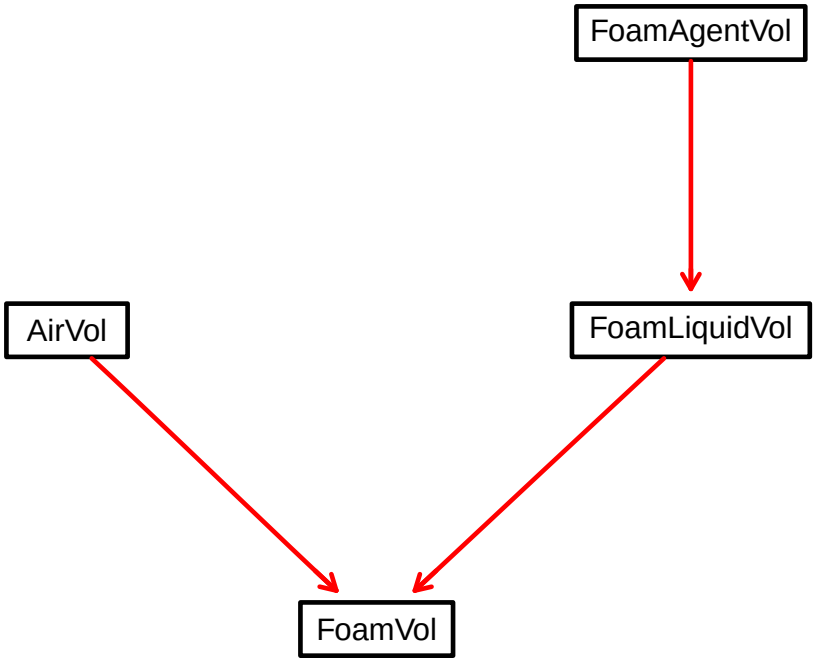

Figure 10.1: Interactions of features related to ground conditioning volumes constructed by the structure learning algorithm. Red links denote true relationship.

estimated based on foam liquid volume, but foam liquid volume can also be estimated if the foam agent volume is known. This might be due to the score equivalence of the graph (Koller & Friedman, 2009) and further investigation should be perform to draw a stronger conclusion (Chickering, 2002).

10.4.2 Interactions of Excavation Features

More complex BN graphs were constructed to model the interactions of EPB TBM excavation features. Similar to the ground conditioning system model, the graphs were constructed by the structure learning algorithm without any prescribed network skeleton. Examples of BN graph iteration during the structure learning process is presented in Figure 10.3. In this process, the algorithm re-iterated the graph configuration to find the highest BIC score.

Figure 10.4 shows two samples of BN graphs that model the interactions of EPB TBM excavation features at ring chainage of 23000 ft (predominantly cohesive soils and till deposits)

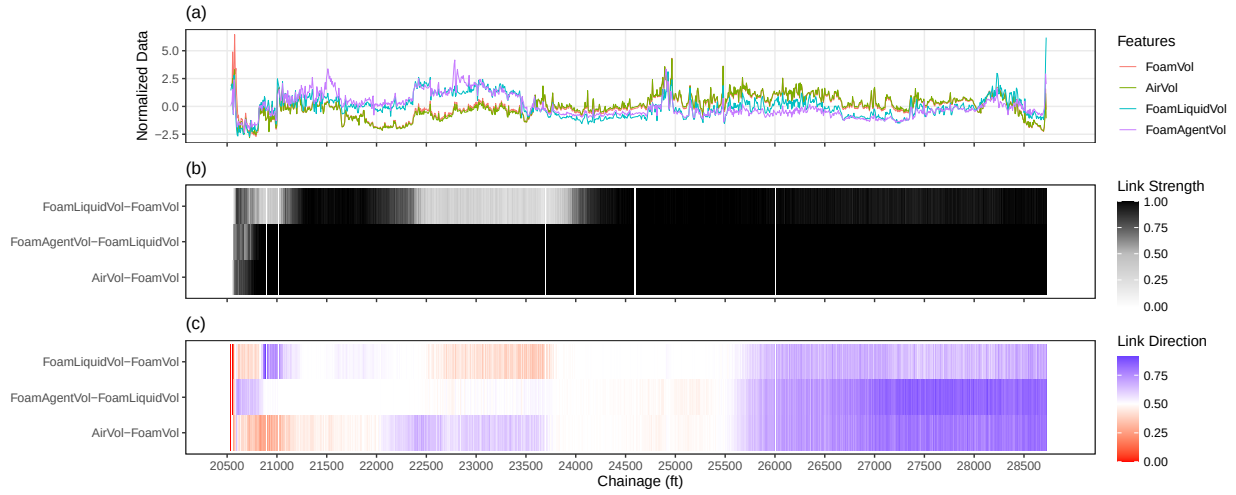

Figure 10.2: Ground conditioning volumes along the tunnel alignment: (a) normalized data, (b) probabilistic strength of the links, and (c) probabilistic direction of the links.

and 28000 ft (predominantly cohesionless soils). To observe more possible interactions of the features, probabilistic strength of 0.5 was used as the link threshold (the link appeared at least in 50% of the 200-graph construction repetition).

Both graphs share several identical results. First, both graphs show that the thrust stroke was independent of other features. This result is sensible since thrust strokes are typically pushed to their length capacity according to the subsequent tunnel lining width and, therefore, should not have any dependencies on excavation conditions. Second, both graphs show that the penetration rate had conditional dependencies on the advance rate and the cutter rotation speed. This constructed interaction successfully models the true mechanism since, in this case, the penetration rate (mm/rotation) was computed as the advance rate (mm/min) divided by the cutter rotation speed (rpm).

Both graphs also show an interaction between the foam volume and the penetration rate, as well as between the foam volume and the advance rate. This is sensible since some studies have shown that ground conditioning can be associated with tunneling excavation performance (Mokhtari et al., 2020; Roby & Willis, 2014). Furthermore, to maintain the expected foam injection ratio (FIR) during each excavation step, foam flow is typically regulated automatically to correspond to the advance rate (Maidl et al., 2013).

Besides the above-mentioned results, both graphs show that interactions of the features might change during tunneling. The graph at chainage 23000 ft shows that the chamber pressure had a conditional dependency on the advance rate, whereas no noticeable link to the screw rotation speed. In contrast, the graph at chainage 28000 ft shows that the screw rotation speed had a conditional dependency on the chamber pressure, whereas no noticeable

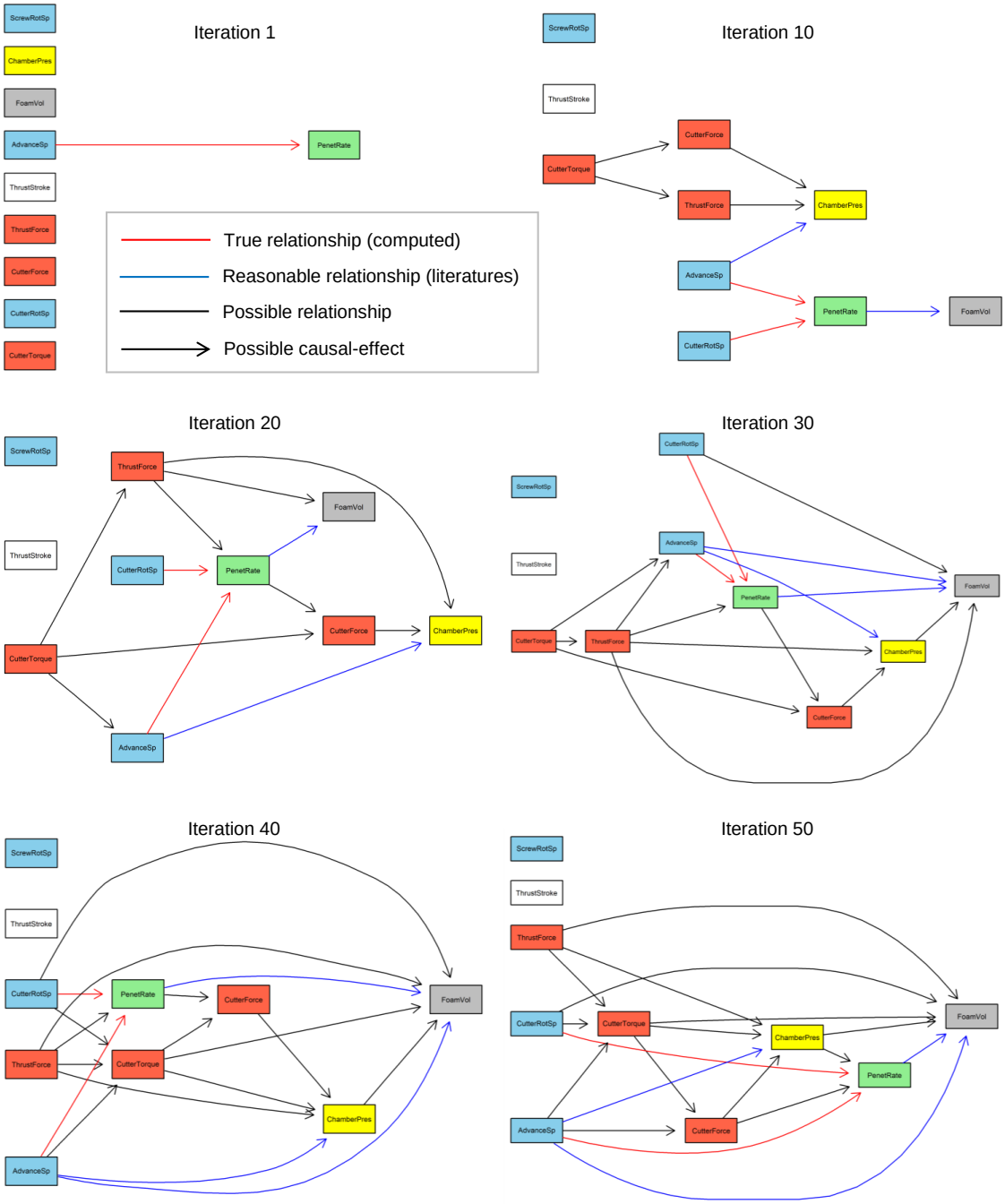

Figure 10.3: Examples of iteration during structure learning to find a BN graph configuration with the highest BIC score.

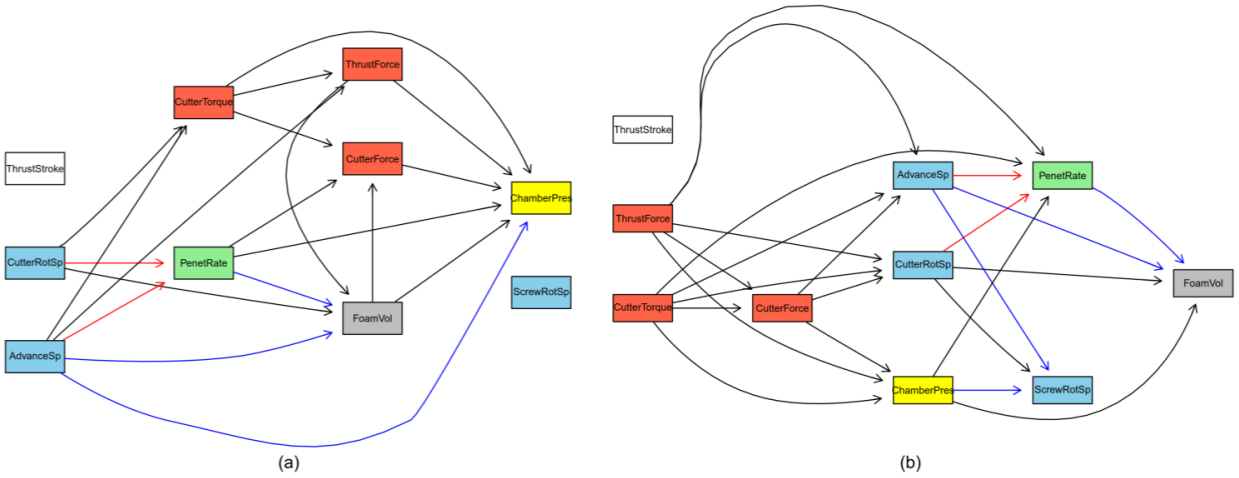

Figure 10.4: Interactions of features related excavation constructed by the structure learning algorithm: (a) chainage 23000 ft, (b) chainage 28000 ft. Red links denote true relationship as calculated; blue links denote sensible interactions based on available literatures; black links denote possible interactions.

link between the chamber pressure and the advance rate.

Furthermore, the graph at chainage 23000 ft shows that speed features (i.e., cutter rotation speed and advance rate) were the root parents of the interactions. In contrast, at chainage 28000 ft, it seems force features (i.e., cutter torque, thrust force, and cutter head force) were the root parents. Typically, advance rate and cutter rotation speed are the key excavation control parameters used by EPB TBM operators. After specifying these control parameters, the operators need to check whether the produced forces and torques are still within the tolerable limits. Therefore, these graphs may indicate that at chainage 23000 ft (predominantly cohesive soils), the operators controlled the speeds without any significant intervention was taken regarding the forces and torques limits. In contrast, at chainage 28000 ft (predominantly cohesionless soils and till deposits), the operators had to keep adjusting the speeds according to the forces and torques limits.

10.4.3 Effects of Soil Labels

To evaluate the effect of providing explicit information about geologic conditions, the BN graphs were also constructed with soil labels included as a feature. Figure 10.5 shows two samples of BN graphs that model the interactions of EPB TBM excavation features with soil labels included as a feature. Similar to the previous graphs, probabilistic strength of 0.5 was used as the link threshold. By incorporating the soil labels, the dataset contained both

numerical and categorical variables. In a conditional linear gaussian BN model with a mixed variable dataset, the categorical features will always be the parent nodes for the numerical features (Scutari & Denis, 2015). Hence, the soil label node was the parent node of almost all features on these graphs.

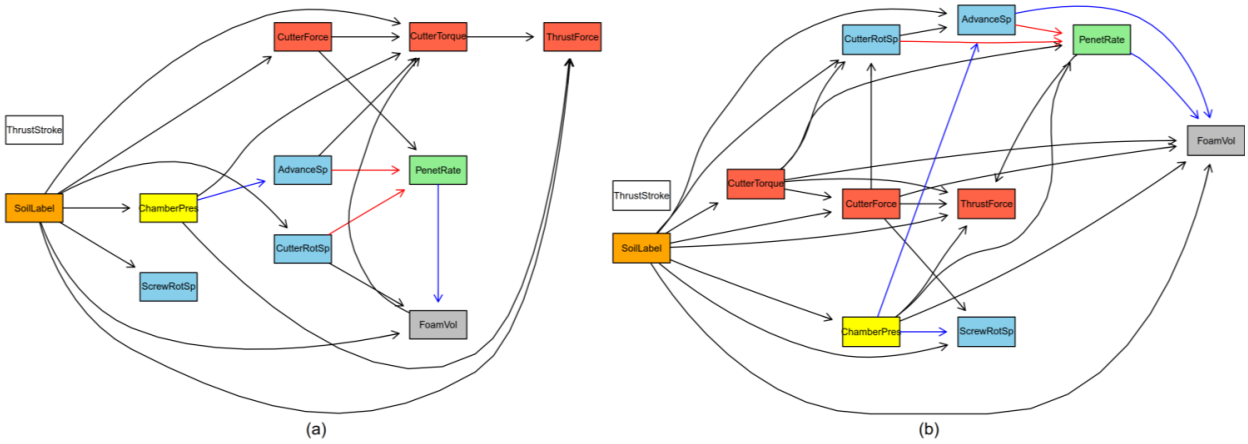

Figure 10.5: Interactions of features related excavation with soil labels constructed by the structure learning algorithm: (a) chainage 23000 ft, (b) chainage 28000 ft. (See Figure 10.4 caption for links color legend.)

Nevertheless, several key interactions in these graphs were similar to the previous graphs without the soil labels. The thrust stroke was independent of other features. The penetration rate always had conditional dependencies on the advance rate and the cutter rotation speed. The foam volume always interacted with the penetration rate. At chainage 23000 ft, the chamber pressure interacted with advance rate, but not with the screw rotation speed. At chainage 28000 ft, the screw rotation speed had a conditional dependency on the chamber pressure, but in this case, the chamber pressure still interacted with the advance rate.

In general, the graphs with soil labels produced fewer interactions between speed features (i.e., cutter rotation speed and advance rate) and force features (i.e., cutter torque, thrust force, and cutter head force). In contrast, features in the graphs without soil labels had more dependencies to force features. This might indicate that the information about the geologic conditions (i.e., the soil types) was stored and represented by these force-related "reaction" features.

10.4.4 Feature Interactions along Tunnel Alignment

To evaluate the stability and the change of feature interactions along the tunnel alignment, the BN excavation graphs were sequentially retrained for each ring chainage location using all the previous ring chainages.

Penetration Rate. Figure 10.6 shows the interactions of features related to the penetration rate along the tunnel alignment. This figure shows that along the tunnel alignment, links between the advance rate and the penetration rate, as well as between the cutter rotation speed and the penetration rate, were consistently strong. The direction of the links was also consistent. These interactions successfully modeled the true mechanism of penetration rate interactions along the tunnel alignment.

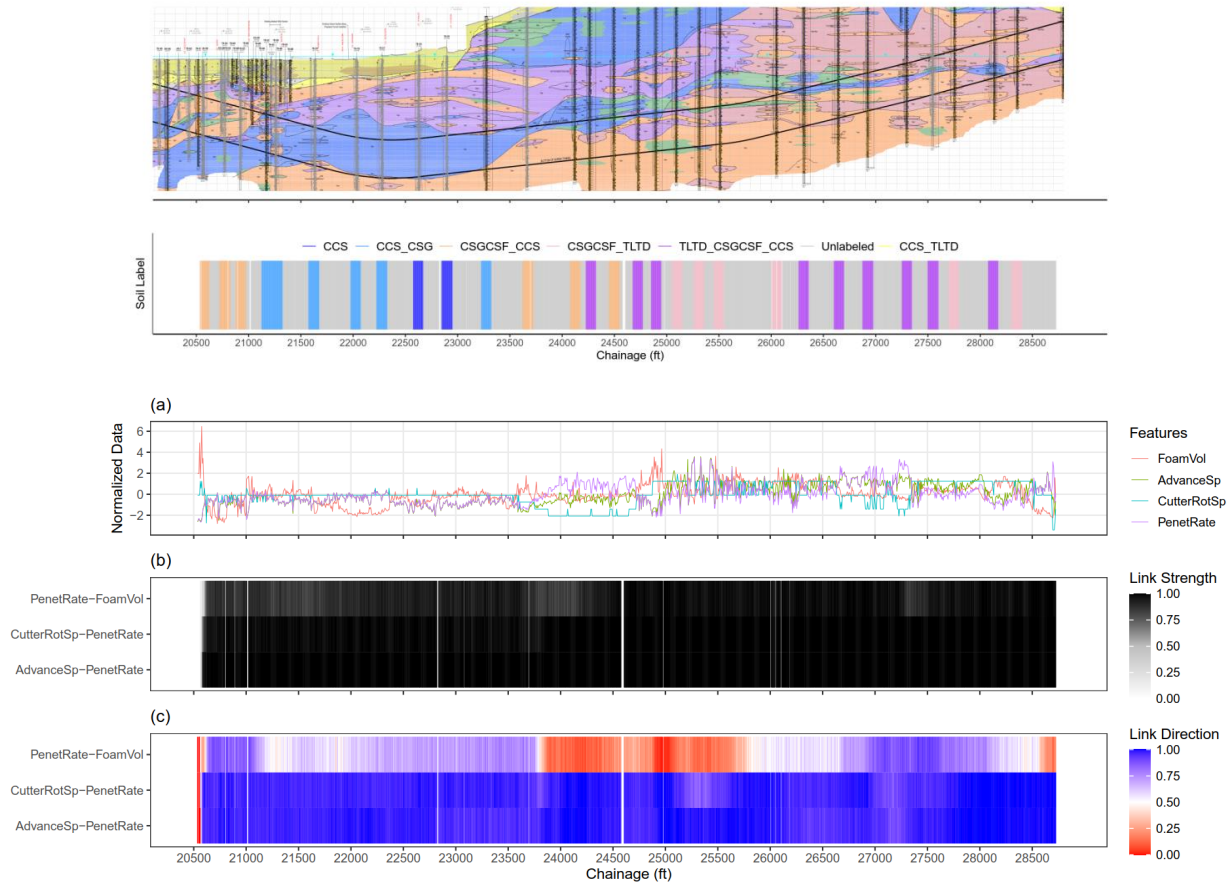

Figure 10.6: Features related to penetration rate along the tunnel alignment: (a) normalized data, (b) probabilistic strength of the links, and (c) probabilistic direction of the links.

The interaction between the penetration rate and the foam volume was also consistent

along the tunnel alignment. This interaction was stronger especially in the end half of the alignment. However, the link direction was inconsistent. Again, this might be due to the score equivalence of the graph. Or this might indicate the tendency on how the operators controlled the advance rate and the injected foam volume. As previously discussed, ground conditioning can be associated with tunneling excavation performance. Injected foam volumes may affect the excavation performance (Mokhtari et al., 2020; Roby & Willis, 2014). Reciprocally, engineers and operators estimate the injected foam volume for the subsequent excavation step based on the previous excavation performance. further investigations should be carried out to obtain more conclusive results on this interaction.

Chamber Pressure. During tunneling, chamber pressures must be regulated to balance the earth pressure. Typically, a target earth pressure is specified according to the overburden pressure and the estimated lateral earth pressure coefficient at a chainage location. Operators regulate the chamber pressure by adjusting the advance rate, which dictates the speed of the chamber to contract or expand, and/or adjusting the screw rotation speed, which dictates the depressurization of the chamber pressure (JSCE, 2016; Maidl et al., 2013).

Figure 10.7 shows the interactions of features related to the chamber pressure along the tunnel alignment. This figure shows that the link between the advance rate and the chamber pressure was stronger during the early half of the tunnel alignment. During this period, the direction of the link was mostly consistent. In contrast, the link between the chamber pressure and the screw rotation speed was stronger during the end half of the tunnel alignment. The direction of the link was also relatively consistent during this period.

This graph might reveal how the chamber pressure was regulated by the operators during tunneling. These results might indicate that no significant intervention on the screw rotation speed was required to regulate the chamber pressure during the early half of the tunnel alignment when the geology was dominated by predominantly cohesive soils. In contrast, more intervention on the screw rotation speed was required to regulate the chamber pressure during the end half of the tunnel alignment when the geology was dominated by mixtures of cohesionless soils and till/till-like deposits. As evidence, it can be observed in Figure 10.7(a) that the screw rotation speeds highly fluctuated during the end half of the tunneling, which indicates stronger control by the operators.

10.5 Conclusions

This study has demonstrated that BN graphs can potentially be utilized to systematically model the interactions of EPB TBM features in a compact and interpretable representation. This study has also shown that the score-based BNSL algorithm could successfully capture several true and sensible mechanisms of the feature interactions based on data, both with and without soil labels. Furthermore, causal relationships between features could be exploited by examining their dependencies and therefore might provide some indication on how the operators controlled the EPB TBM.

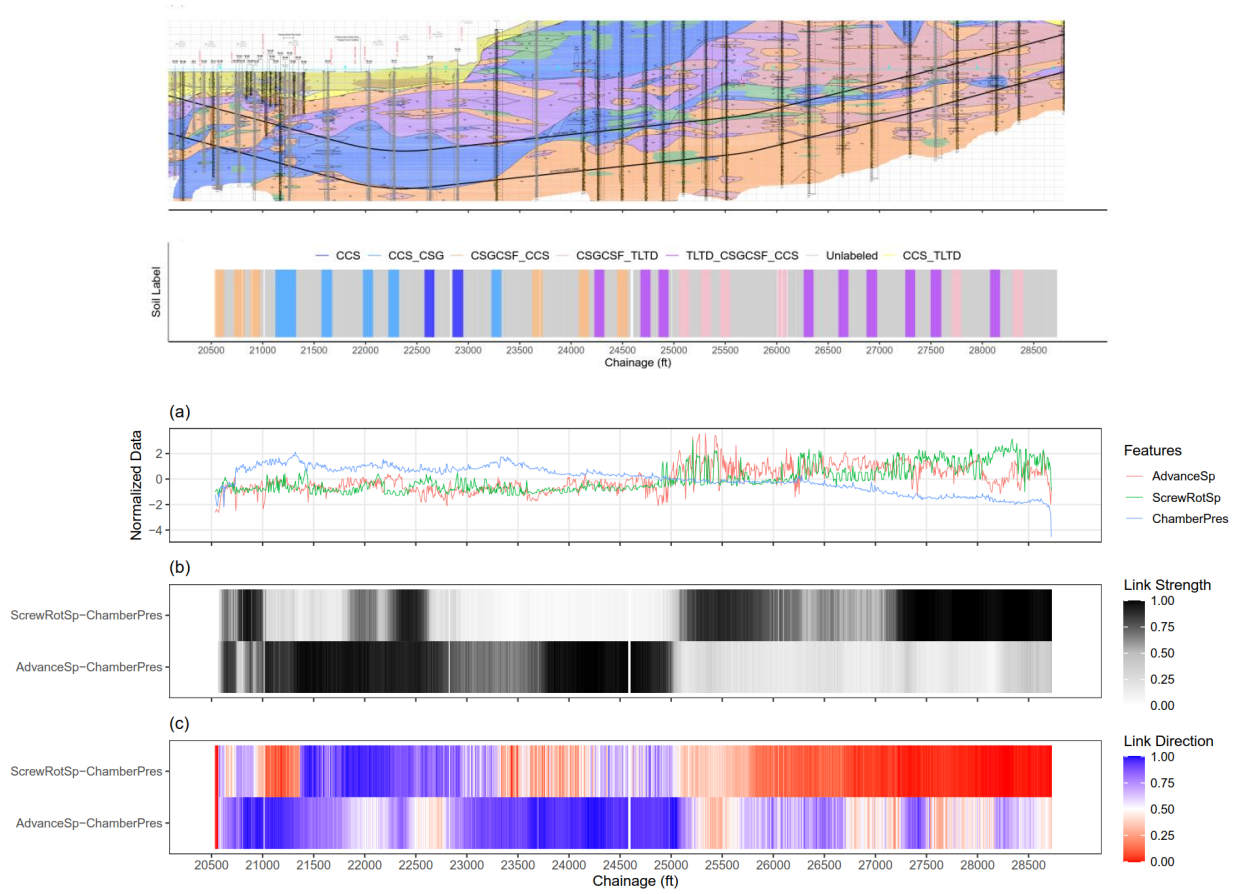

Figure 10.7: Features related to chamber pressure along the tunnel alignment: (a) normalized data, (b) probabilistic strength of the links, and (c) probabilistic direction of the links.

Detailed investigations should be carried out to obtain more conclusive results about the interactions, e.g., by introducing more features to add more complexity, by utilizing probability distribution in a finer spatial distance to capture detailed operator decisions, and by implementing dynamic BNSL to explicitly capture the sequential decision-making processes.

Chapter 11

Modeling TBM Steering Control Decisions

11.1 Introduction

11.1.1 Background and Related Works

The tunneling industry has been aiming to develop autonomous tunnel boring machine (TBM) steering control systems in the past few decades. The early development started more than two decades ago when Uematsu et al. (1996) proposed a rule-based steering control system for a TBM. In principle, the system was executed if the TBM position exceeded a specified deviation limit. The steering control was performed by specifying the thrust jack stroke difference, which was estimated based on the statistical relationship from past data. The system was implemented at the Tsukui headrace tunnel project in Japan by a Joint venture of Kumagai, Goyo, and Dainippon.

In the following decade, the development was dominated by numerical and experimental simulations. The popular control framework was the feedback control system. Many studies implemented this system in different settings, such as the two closed-loop control structures (Yue et al., 2012), cascade strategy with fixed-value feedforward (Xie et al., 2012), and multi-cylinder control system (L. Wang et al., 2018). These feedback systems used different types of controllers. However, most of them were in the family of PID controllers, such as the basic PID controller (Huayong et al., 2009), variable PID controllers (Xie et al., 2012), and fuzzy PID controller (Wu et al., 2022).

The development of autonomous TBM systems in real tunneling projects emerged again in the past few years. This emergence is most likely due to the availability of massive TBM operation data, the advancement of machine learning, and the growth of computing performance. Hoshino et al. (2020) proposed a force-based steering control system. Unfortunately, the paper does not explicitly describe the developed steering control system. However, it appears that the system was developed using conventional control systems such as PID controllers or rule-based algorithms without any machine learning algorithms. The system was

implemented by Obayashi Corp. at multiple tunneling sites in Japan.

Xiong et al. (2020) proposed an autonomous steering control system that could operate with minimal human intervention. In principle, the system determined the optimal thrust jack pressure assignments to maneuver the machine toward the design tunnel alignment. The jack pressure assignments were adjusted automatically and continuously in real-time as the data was fed into the system every 3 seconds. Again, the paper does not clearly describe the predictive system inside the controller. However, it appears to be some rule-based algorithms embedded in a feedback control system. The system was implemented by MMC-Gamuda KVRMT(T) Sdn., Bhd. at the Klang Valley MRT Line 2 tunnel in Kuala Lumpur, Malaysia.

Wada et al. (2021) proposed a machine learning-based steering control system. In principle, the system guided the TBM maneuver by estimating the optimal resultant force point produced by the thrust jacks. The system used a gradient boosting type of algorithm as a predictive model and utilized 14 TBM features as the input predictors. These predictors included features related to the machine coordinates, attitudes, deviations, and advancing features (i.e., thrust force, cutter torque, advance rate). The system was implemented by Shimizu Corp. at an undescribed tunneling project.

Hu et al. (2022) proposed an autonomous earth pressure balance (EPB) TBM control system that could operate without human intervention. The system was developed based on the human performance model (Rasmussen, 1983) and used intelligent modules to access and control various set points, such as cutter rotation, thrust speed, jack selection, earth pressure, screw rotation, gate opening, as well as tail grease, and grouting quantity. Unfortunately, the paper does not provide detailed descriptions of each intelligent control module. The system was implemented by Shanghai Tunnel Engineering Co., Ltd. (STEC) at the intercity railway tunnels between Hangzhou and Shaoxing, China.

11.1.2 Objectives

From the discussion above, it can be noticed that most of these developments were backed by the industry. This allows the implementation of the developed systems in real tunneling projects, which is immensely valuable for advancing this technology. However, on the other hand, this may be the reason why many of the published works provide obscure descriptions of the developed systems (e.g., due to proprietary-related issues). Nevertheless, it should be noted that most of the proposed systems were developed based on conventional control systems such as rule-based algorithms and PID controllers. These systems may not be fully intelligent to learn and adapt to changing trajectories and ground conditions, which are the main challenge in tunneling. These lead to problems with generalization. Wada et al. (2021) mentioned that the performance of their guidance system deteriorated when the tunnel alignment rate changed. Hu et al. (2022) admitted that their system needs to be readjusted and retrained in different ground conditions.

This study aims to develop an artificial intelligence (AI) system that is capable of making steering control decisions in changing trajectories and external disturbances. The system was developed in a feedback control system framework to ensure its robustness. The system's

controller unit was built using machine learning. This enables the controller to (i) make concurrent decisions in multiple TBM steering control parameters and (ii) adapt to the changing target trajectories and ground conditions.

11.2 Data

11.2.1 Tunneling Case and Geologic Conditions

This study used a data set of the State Route 99 (SR99) highway tunnel in Seattle, Washington, USA. This data set was considered suitable for the steering control simulation since the tunnel alignment has different trajectories in various ground conditions. More discussion of this tunneling case and the geologic conditions can be found in Chapters 5 and 7, respectively.

11.2.2 Data Description and Preparation

This study used the 20 mm-length aggregate spatial series as observation data points. This data resolution was considered adequate for simulating detailed TBM movements and behaviors with sensible data noises and computation costs. Also, this study used 54 features (excluding ring numbers and the chainages) that were obtained from the original data set. These features were selected based on the EPB TBM taxonomy and the hierarchical feature selection discussed in Chapter 6. The selected features are listed in Table 11.1. Note that the coordinate positions were transformed as the model's input. This transformation is discussed later in the model setup section.

After the initial selection, the data sets were prepared by (i) normalizing volume-related features with the stroke length, (ii) removing non-excavation observation data points, and (iii) creating features that represent the difference between the current and the next TBM coordinate positions, and (iv) removing rows with missing data. This resulted in 138,926 observation data points from ring numbers 0 (chainage 19496.93 ft) to 1418 (chainage 28722.99 ft).

11.3 Methods

11.3.1 Steering Control Parameters

The EPB TBM used in the SR99 tunneling project could be maneuvered using active and passive articulation systems. The passive system is the primary steering control system of a TBM. In this system, the steering is performed by regulating the patterns and pressures of thrust jacks to achieve a specified jack stroke difference (JSD). In contrast, the active system is essentially articulation jacks that connect the front and rear shields of a TBM. This system allows the TBM to maneuver in a smaller curve radius. The steering is performed

Table 11.1: List of EPB TBM features and positions used in the steering control decision model.

Group	Component	Sub-component	Unit	Num. Feature	Remark
Shield Position	Coordinates	Northing	m	3	Head, artic., tail
		Easting	m	3	Head, artic., tail
		Elevation	m	3	Head, artic., tail
	Deviations	Horizontal	mm	3	Head, artic., tail
		Vertical	mm	3	Head, artic., tail
		Articulation	mm	2	Left-Right, Top-Bottom
Excavation	Cutter head	Cutter rot. speed	rpm	1	
		Cutter torque	kNm	1	
Advancing	Thrust	Advance rate	mm/min	1	Average
		Force	kN	1	
		Stroke	mm	1	Net stroke
Steering	Thrust	Pressure	bar	8	Thrust jack groups
	Articulation	Stroke	mm	4	Artic. jack groups
	Attitude	Pitch	deg	2	Front, rear
		Roll	deg	2	Front, rear
		Yaw	deg	2	Front, rear
Ground conditioning	Foam	Volume	m^3/m	1	Total of 32
		Pressure	bar	1	Average of 32
	Polymer	Volume	m^3/m	1	Total of 15
		Pressure	bar	1	Average of 15
	Slurry	Volume	m^3/m	1	Total of 26
		Pressure	bar	1	Average of 26
	Additive	Volume	m^3/m	1	Total of 15

Earth pressure balancing	Excavation chamber	Pressure	bar	1	Average of 12
Muck discharging	Screw conveyor	Rot. speed	rpm	1	Average of 2
		Pressure	bar	1	Average of 4
	Belt conveyor	Muck weight	ton	1	Total of 2
		Muck volume	m^3/m	1	Total of 2
Backfill grouting	Tail grout	Pressure	bar	1	Average
		Volume	m^3/m	1	Total

by regulating jack strokes to achieve a specified articulation angle. The schematic diagram of these systems is shown in Figure 11.1.

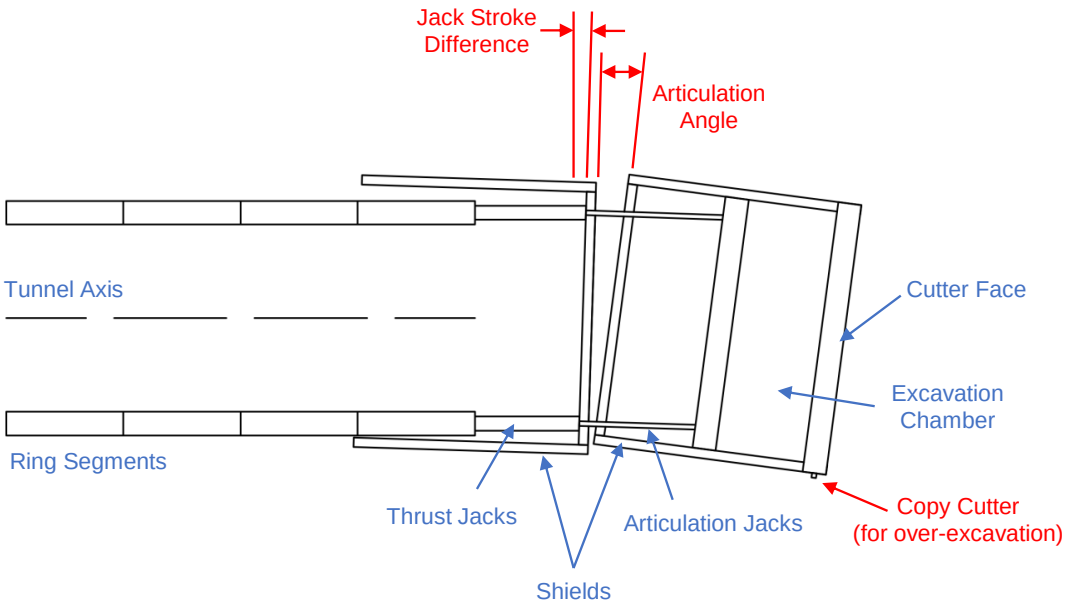

Figure 11.1: Schematic diagram of TBM steering systems.

In this study, the passive and active steering control systems of the EPB TBM were represented as thrust pressures and articulation strokes. The thrust jacks were arranged into eight pressure groups formed into a circle inside the cutter head periphery. The articulation jacks were also arranged into four articulation groups, i.e., the top, bottom, right, and left

strokes. The configuration of the thrust and articulation jack groups used in this study are presented in Figure 11.2.

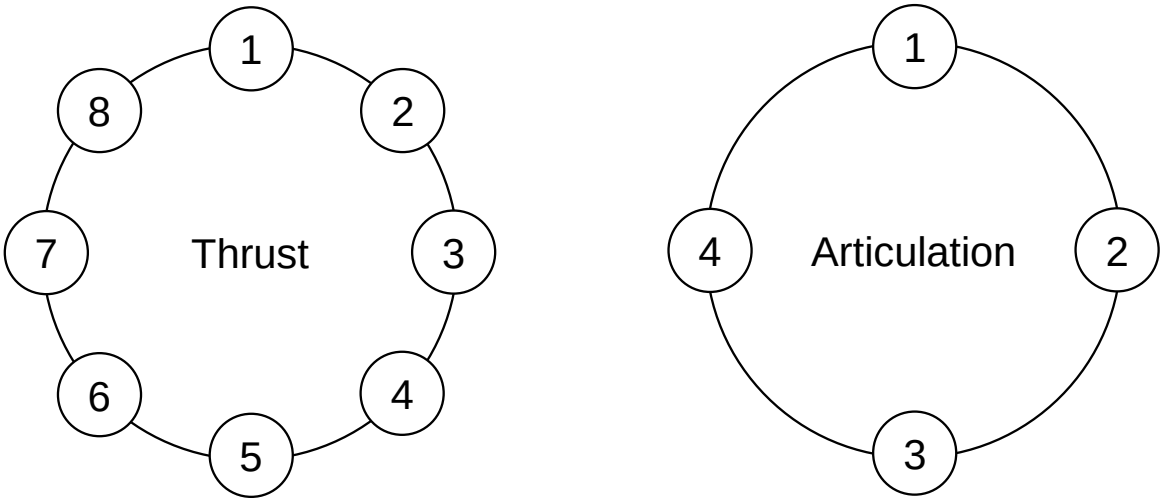

Figure 11.2: Configuration of the thrust and articulation jack groups.

11.3.2 Model for Steering Control Decisions

Feedback control systems have been popularly used to build robust autonomous systems. The configuration of this system enables the controller to adjust the input by injecting the system's errors, i.e., the difference between the desired output (reference point) and the actual output. In this study, the steering control system was developed in the framework of a feedback control system (Figure 11.3). In this framework, TBM behaviors and movements correspond to the system's process. The ground conditions correspond to the external disturbances that affect the system's process. The tunnel design alignment and the actual position correspond to the reference point and the process output, respectively. The difference between the target and actual TBM positions corresponds to the system's error. This error can be identified using the system's measurement unit to be injected back into the controller unit. In manual operation, the human operator represents the controller unit. This unit is responsible for making the steering control decisions based on the given reference point and measured errors. This unit is also responsible for activating the steering control actuators, such as the thrust and articulation jacks.

Essentially, the system's process (i.e., TBM movements) can be seen as a dynamical system that transforms the TBM position, given the actuated steering control parameters and

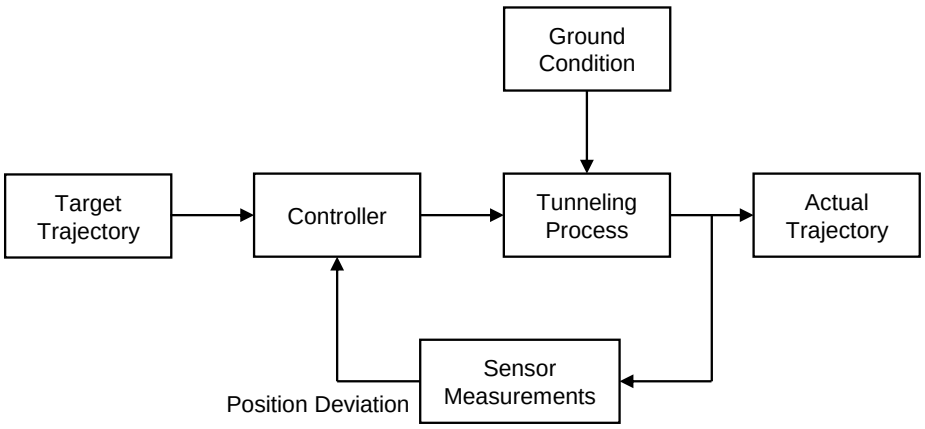

Figure 11.3: Schematic diagram of the feedback system for TBM steering control.

the external disturbance, such as the ground conditions (Figure 11.4). In this perspective, the system's process can be inversed into a model that outputs the steering control parameters based on the TBM current position, the TBM target position (with an expected deviation of 0), and the external disturbances (Figure 11.5). Data-driven techniques can be used to model this process. This data-driven approach was chosen due to the complexity of the process and the uncertainty of the external disturbance. These conditions are difficult to be modeled using analytical or numerical approaches.

11.3.3 Model Setup

The steering control decision model was simulated using the SR99 data set. This model was developed based on the premise that there must be a relationship between the steering control parameters and the three other factors of the TBM movement process, i.e., the TBM current position, the TBM next position, and the external disturbances. In this real data set, the TBM positions were represented by the TBM delta coordinates (i.e., the difference between the current and next coordinates) and the deviations. Note that, transforming the current and the next target coordinates to the delta are required. Without this transformation, the machine learning model will always extrapolate the predictions since the next coordinate will always be outside the training boundary (all previous coordinates). As discussed in Chapter 7, machine learning may perform poorly in extrapolation problems.

The external disturbances should correspond to information about the ground conditions. This information was represented by the TBM excavation data at observation data point i+1. Previous studies have shown that TBM excavation operation data contain information that

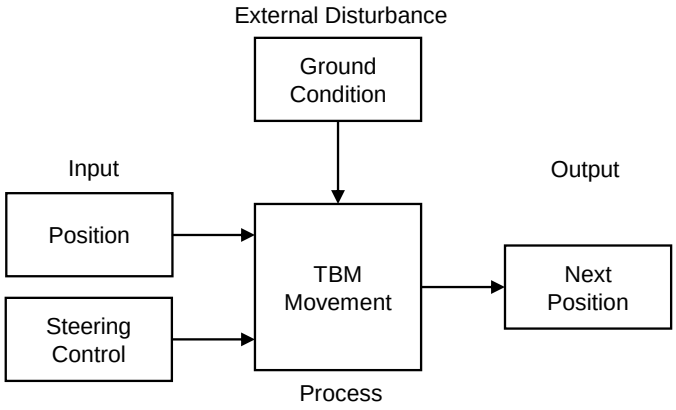

Figure 11.4: TBM movements as a dynamical system.

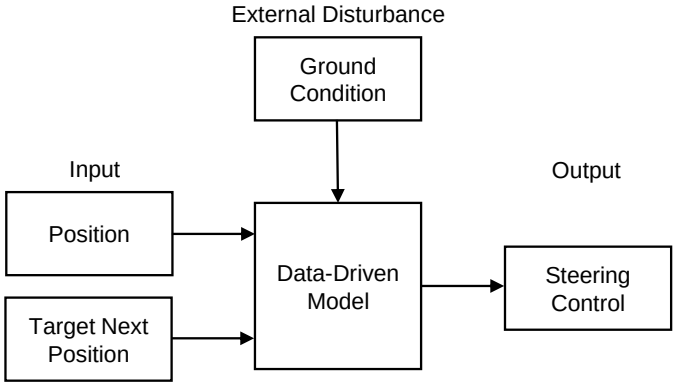

Figure 11.5: Inverse of TBM movement process that outputs the steering control parameters.

can characterize the ground conditions and their interactions with TBM behaviors (Chapters 7, 8, and 10). Machine learning was used to map these three factors to the steering control parameters at observation data point i+1, i.e., thrust pressures and articulation strokes. The schematic diagram of the model input-output mapping in the training phase is presented in Figure 11.6.

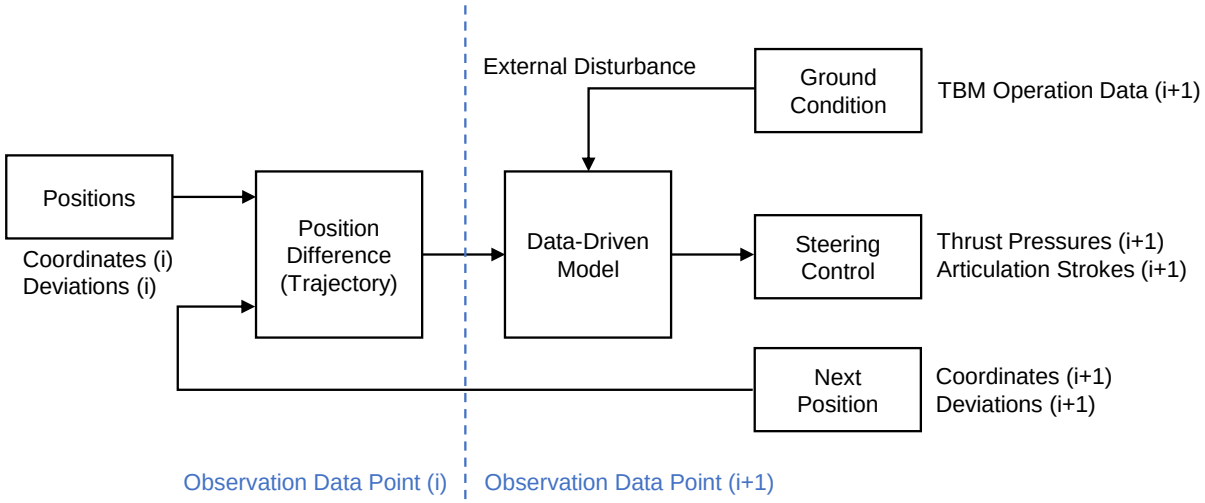

Figure 11.6: Schematic diagram of the model input-output mapping (training phase).

The steering control decision model was developed in the training phase by accurately mapping the input and output based on the real physical process. However, the model's input must be modified in the testing (or prediction) phase. In real tunneling, the external disturbances (i.e., the TBM excavation data) at the next observation data point (i+1) do not exist at the current observation data point i. Thus, the model must use the latest available external disturbance (at observation data point i) as a proxy for the external disturbance at the following observation data point (i+1). This substitution was made by presuming that the external disturbance (i.e., ground conditions) at the current and following observation data points is not substantially different. This assumption should be valid in fine-level observation data points (see Chapter 5), where there is no substantial time and space difference between observation data points i and i+1. The schematic diagram of the prediction model input and output (testing phase) is presented in Figure 11.7.

The models were trained and tested in a sequential scheme. This scheme was selected to represent the real tunneling sequence. In this scheme, the following steering control parameters were determined by the model trained using all data from the previous tunneling sequence. This training process was performed at every tunneling sequence simulation. This sequential training allows the model to learn incrementally during tunneling.

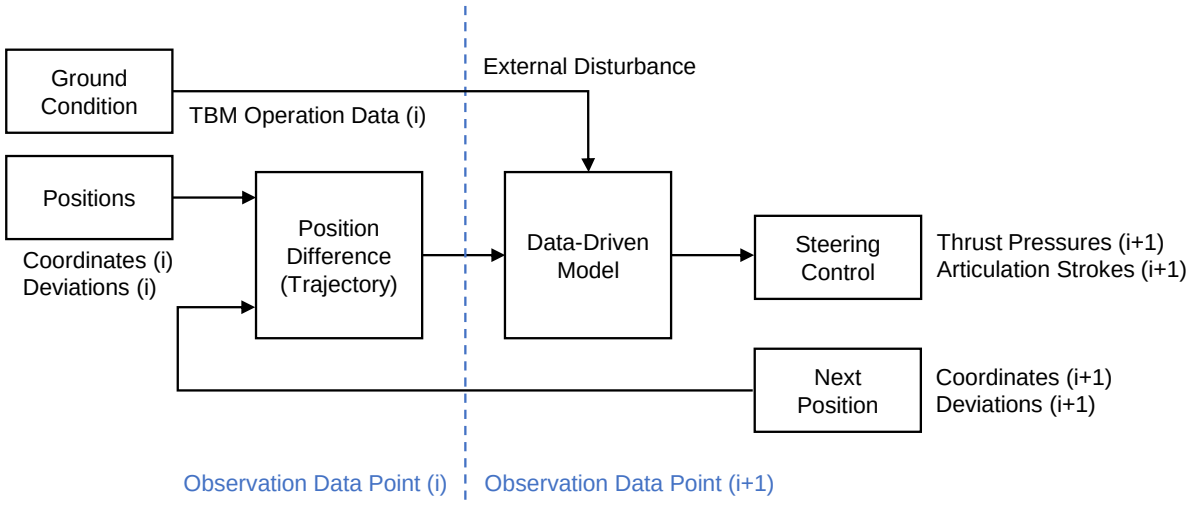

Figure 11.7: Schematic diagram of the prediction model input and output (testing phase).

11.3.4 Prediction Method

Multivariate Random Forest (MRF) was selected as the machine learning algorithm that maps the model's inputs and outputs. MRF is a supervised learning algorithm that can produce multiple prediction outputs simultaneously (Ishwaran et al., 2021). Similar to random forests (Breiman, 2001), MRF is a nonparametric nonlinear prediction method, where the prediction results are produced by bootstrapping the training data, constructing a number of decision trees, and aggregating all the single decision tree results. This algorithm was also selected due to its robustness to overfitting and straightforward hyperparameter tuning. The selection of the hyperparameters is discussed in the following.

- Number of trees = 500. This number of trees was considered adequate to ensure decent predictability of the model (see Chapters 6, 7, and 9).
- Number of randomly selected features to split in each node = number of predictors divided by 3. This value was selected based on the default number recommended by Probst et al. (2019) and the previous chapters (see chapters 6 and 7).
- Minimum size of terminal nodes = 5. This value was selected based on the default number recommended by Probst et al. (2019).
- Number of random splits for splitting a feature = 10. This stochastic approach was selected to reduce the computation cost. The value was selected according to a recommendation by including stochastic (Ishwaran et al., 2021).

- Splitting rule = Mahalanobis distance. This rule was selected since this rule takes into account the correlation over the response feature coordinates, as proposed by Ishwaran et al. (2021). By using this splitting rule, the multivariate outputs do not need to be standardized to avoid unbalanced influence from outputs with a larger range of values (Ishwaran et al., 2021).

More discussion and mathematical background of this algorithm can be found in Chapter 4. This study used the MRF implementation by Ishwaran and Kogalur (2022).

11.3.5 Model Evaluation

The model was evaluated by comparing the steering control parameters produced by the model and the actual steering control parameters determined by human operators. The simulation was carried out at five selected ring segments, i.e., ring numbers 251, 501, 751, 1001, and 1251. These ring segments were selected to represent various geologic conditions and TBM trajectories along the tunnel alignment. Subsequently, another simulation was performed to investigate the effects of incremental learning on the model predictability. In this simulation, the model was used to estimate the steering control parameters of the following six ring segments, i.e., ring numbers 1009 to 1014. In this case, the training was done one-time using data from the beginning of the tunneling to ring number 1008. This simulation was then compared to another model at the same simulation condition but was retrained in every ring segment advancement.

11.4 Results and Discussion

11.4.1 Steering Control Decisions: Human Operator versus Machine Learning

Figure 11.8 presents the result of the steering control simulations at the five selected ring segments. The figure compares the average steering control parameters of each ring segment that were determined by the human operator (from the real TBM operation data) and the machine learning model (from the simulation). As discussed previously, the steering control parameters were represented as the assigned pressures on eight thrust jack groups and stroke lengths of four articulation jack groups. The human operator and machine learning control decisions are shown in red and blue bars, respectively.

Ring number 251 represents segments with a straight trajectory and increasing depth. For this prediction, the model was trained using all previous operation data until ring number 250. It can be observed that the human operator assigned almost the same pressures on all the thrust jack groups and slightly longer strokes for the top articulation jack group. It should be noted that no dominant thrust pressure is required when a TBM increases the depth as gravity tends to drag the shield to pitch downward. The simulation shows that the

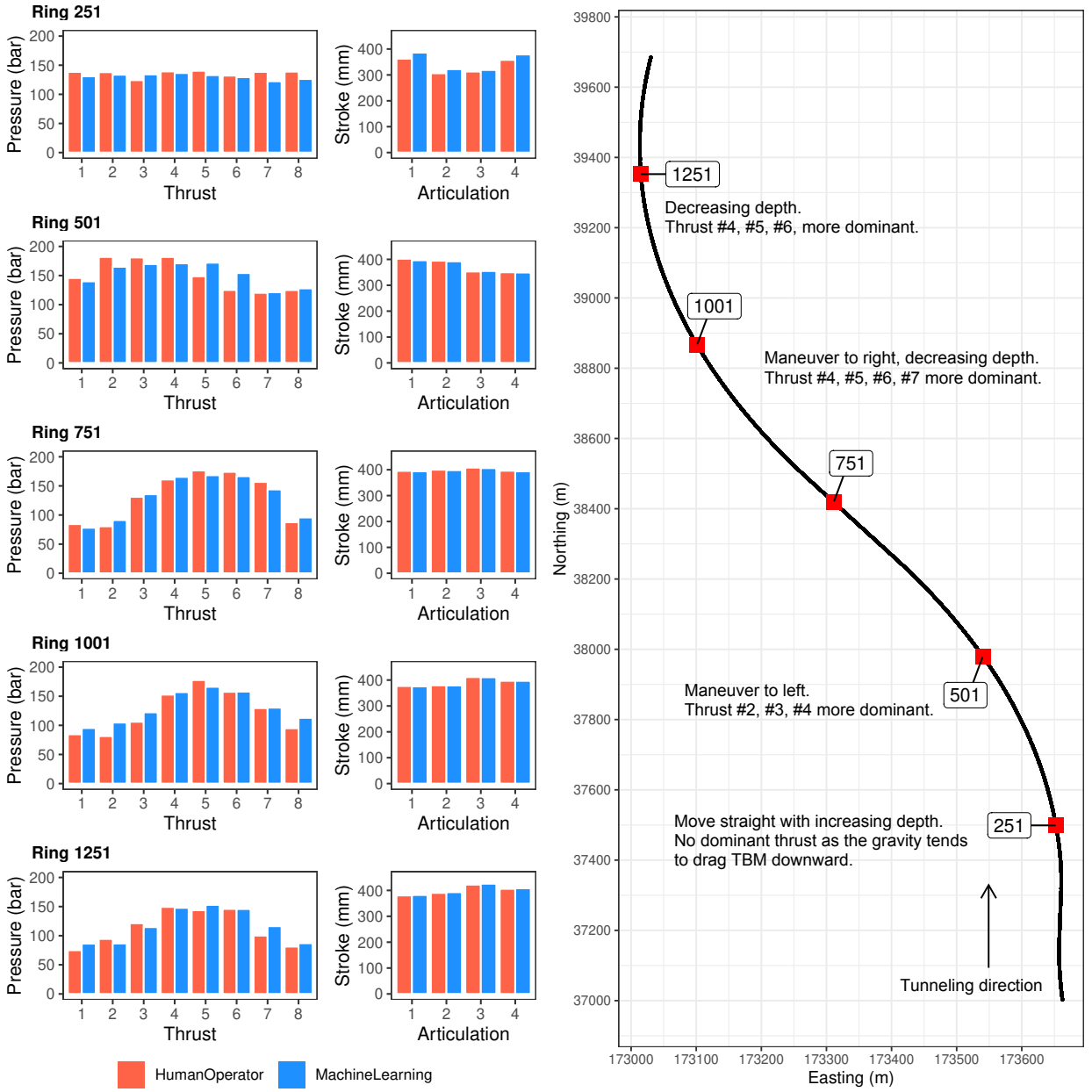

Figure 11.8: Simulation of steering control decisions at five selected ring segments: human operator vs. machine learning.

machine learning model could produce reasonable steering control parameters with similar thrust pressure and articulation stroke patterns to the human operator decisions.

Ring number 501 represents segments that curve to the left. At this ring segment, the human operator maneuvered the TBM to follow this trajectory by assigning larger pressures on the right side thrust jack groups, i.e., thrust group numbers 2 to 5. The operator also specified longer strokes at the top and right articulation jacks, indicating the TBM was still moving down to a greater depth. Interestingly, the machine learning model could capture these steering control decisions. The model produced similar thrust pressure and articulation stroke patterns to the human operator. Minor discrepancies can be observed in the thrust group numbers 5 and 6. These discrepancies may be due to a bias in the training data, where the model was trained using more straight segments than curving segments. Note that the TBM just started to curve to the left side from a long straight trajectory at this chainage.

Ring numbers 751 and 1001 represent segments that curve to the right. The TBM climbed up at these ring segments to decrease the tunneling depth. To follow this trajectory, the human operator assigned higher pressures on the left and bottom thrust jack groups, i.e., thrust group numbers 4 to 7. The high pressures on the bottom side were required to push the TBM upward against gravity. The articulation jacks were in similar stroke lengths, with slightly longer strokes at the left articulation jack. This indicates that the radius of the curve was large enough, and adjusting the thrust jacks with only small intervention from the articulation jacks was adequate to maneuver the shield. These patterns apply to both ring segments and machine learning simulations. No obvious errors can be seen in every thrust and articulation group, indicating the increase in the model's predictability performance. This is expected since, at this point, the model had been trained using more data with various trajectories.

Ring number 1251 represents the end of the tunnel alignment. At this segment, the TBM continuously climbed up to reach the ground surface. It can be observed that both the human operator and machine learning assigned larger thrust jack pressures and longer articulation jack strokes at the bottom side of the shield, i.e., thrust group number 4 to 6 and articulation group number 3. These results demonstrate the capability of the machine learning model to mimic human operator decisions. The model successfully assigned appropriate steering control parameters based on the current and the next TBM position, given the TBM excavation data as the proxy of ground conditions.

11.4.2 Incremental versus Non-Incremental Learning

The discussion above shows that the machine learning model could reasonably determine the average steering control parameters of a ring segment. The following questions were, should the model be trained in every ring segment advancement? Could the model estimate the steering control parameters of the following ring segments without retraining? Figure 11.9 presents a simulation where the machine learning model was employed to determine the steering control parameters of six consecutive ring segments, i.e., ring numbers 1009 to

1014. In this simulation, the training was done one-time using data from the beginning of the tunneling to ring number 1008.

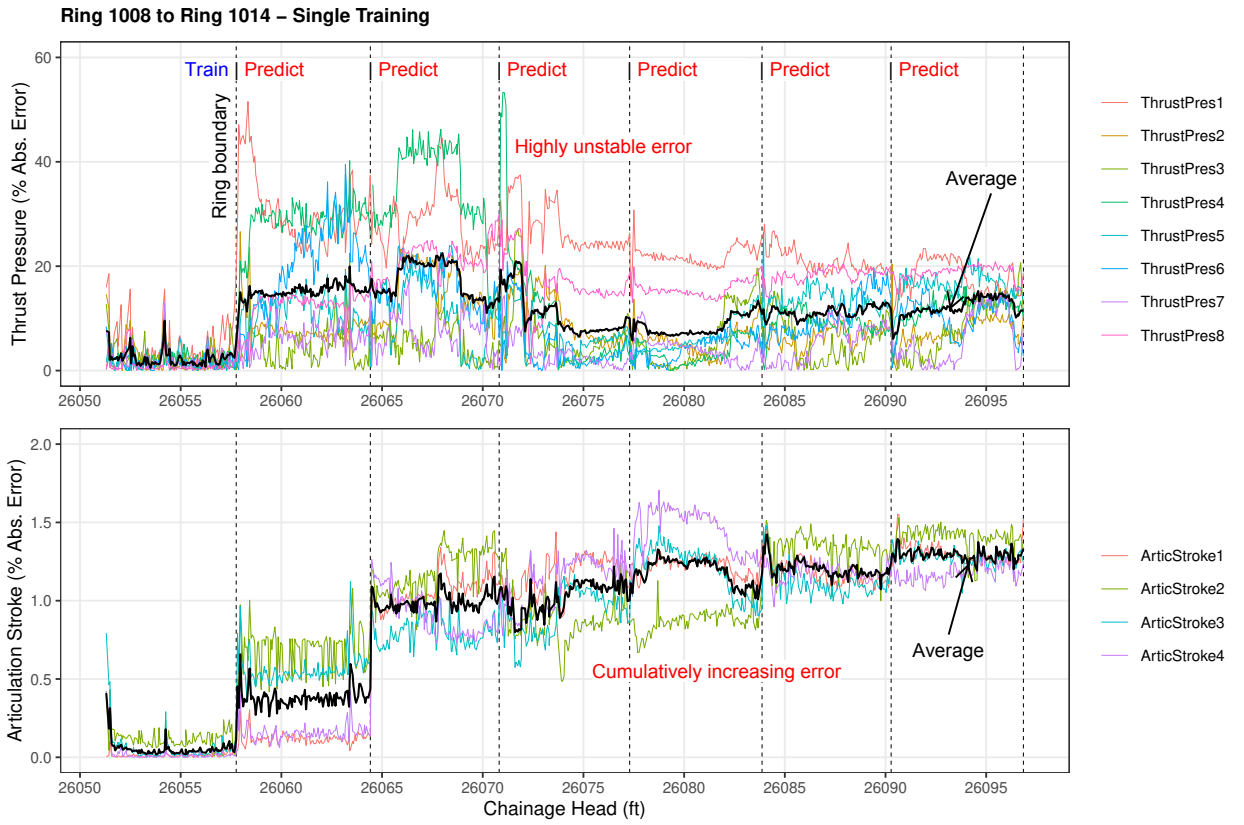

Figure 11.9: Simulation of steering control decisions at six consecutive ring segments without incremental learning.

The top panel presents the absolute errors of the thrust jack pressures in percentage to the actual values. The average error of all the thrust pressures is shown as the black line. Firstly, the model was employed to estimate the thrust pressures in ring 1008, which is inside the training boundary. As expected, the model resulted in a decent performance with very low errors. Then, the model was employed to estimate the following ring segment, i.e., ring number 1009. It can be observed that the errors were inflated with an average error of about 15%. Note that there is also a substantial range of errors for every individual thrust group. This may indicate different trajectory characteristics between ring number 1009 and the training data, which might force the model to do extrapolation. As widely known, machine learning models may have poor performance when predicting new data outside their training boundary (Dubois et al., 2020; H. Zhang et al., 2017). These errors

even increase in the following ring (ring 1010), with the average error reaching more than 20%. The average error drops to about 10% in rings 1011 and 1012 but increases again to about 15% in rings 1013 and 1014. This shows the instability of the model's prediction performance.

The bottom panel presents the absolute errors of the articulation jack strokes in percentage to the actual values. This figure shows that the articulation strokes fluctuated less than the thrust pressures, which can be seen from a smaller range of errors. The average error in predicting ring 1009 is only less than 0.5%. However, the average error continuously increases in the following ring segments, producing an average error of almost 1.5% at ring 1014. This shows that the model suffered more errors as the distance to the training data increased. This increase might be due to error accumulation since articulation stroke is a geometrical parameter, where the error of a prediction can affect the following prediction.

Figure 11.10 presents another simulation where the machine learning model was also used to estimate the steering control parameters of ring numbers 1009 to 1014. However, in this simulation, the model was retrained at every ring segment advancement. For example, to estimate the steering control parameters in ring 1014, the model was retrained using all previous operation data until ring 1013. Similar to Figure 11.9, the top and bottom panels present the absolute errors of the thrust pressures and the articulation strokes in percentage to their actual values.

This figure shows considerably lower errors than Figure 11.9, especially in rings 1010 to 1014, where the model had incrementally learned from the preceding ring segments. The average error of thrust pressures is stable at about 15% in rings 1009 and 1010 and decreases to about 10% or less in the following ring segments. Similarly, the average error of articulation strokes is about 5% in ring 1009, slightly increased to about 5 to 10% in rings 1010 and 1011, and finally stabilized at about 5% in the following ring segments. Note that no accumulation of errors can be observed in the articulation strokes.

The decent performance of this simulation highlights the importance of the sequential training scheme in developing machine learning models for steering control decision-making, especially since the steering process involves different ground conditions, trajectories, and shield attitude responses that may be outside the previous training boundaries. Updating the model by retraining enables incremental learning and helps the model overcome this extrapolation problem.

11.5 Conclusions

This study has proposed an AI system that is capable of making TBM steering control decisions in changing trajectories and external disturbances. The proposed system uses multivariate random forests (MRF), a multi-output supervised learning algorithm, to simultaneously determine the assignment of multiple steering control parameters, i.e., thrust pressure and articulation stroke arrangements. The machine learning model utilizes the current TBM position, the next step target position, and the current step TBM excavation

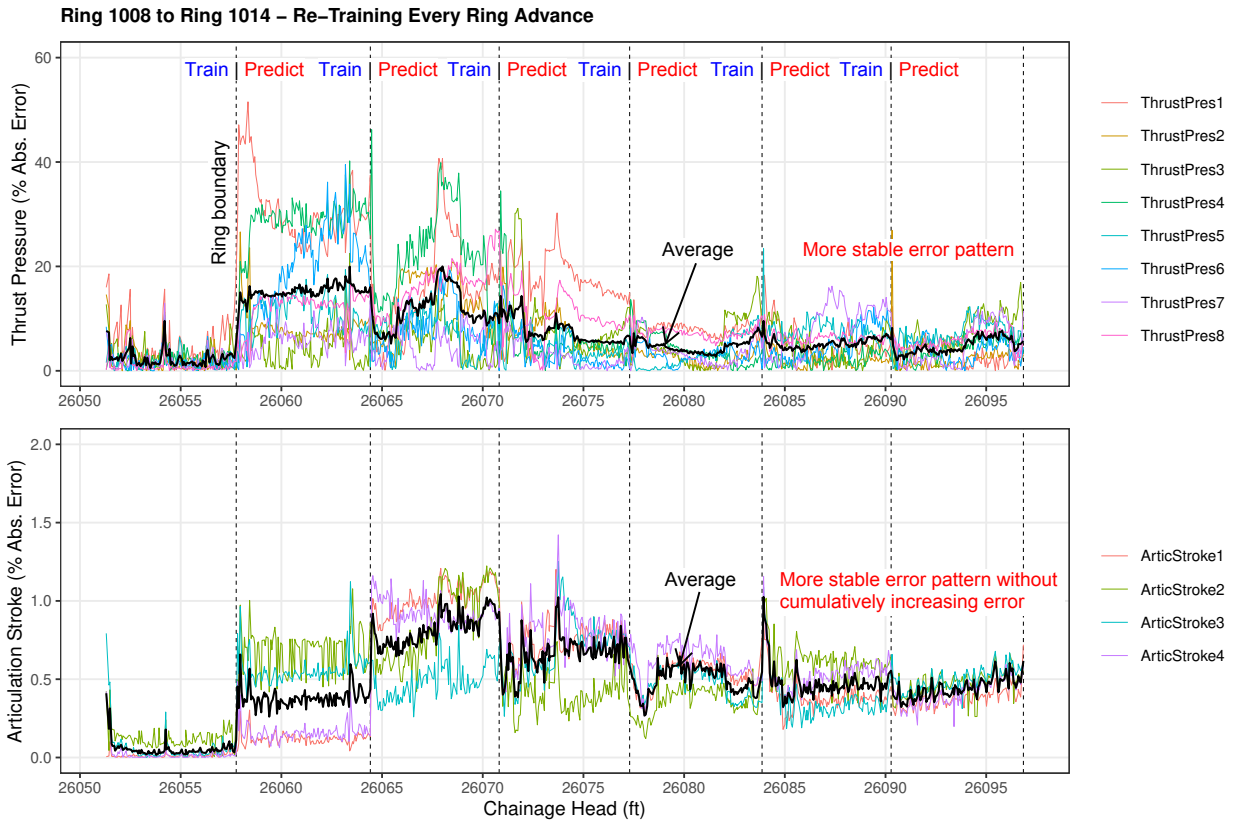

Figure 11.10: Simulation of steering control decisions at six consecutive ring segments with incremental learning.

data (a proxy of the ground conditions) as predictors of the steering control parameters. The main findings of the simulation are summarized in the following.

- **Simulating steering control decisions**. The proposed system successfully mimics human operator decisions by producing appropriate patterns of steering control parameters. This result demonstrates the possibility of using an end-to-end learning framework to develop a self-driving TBM system.

- **The importance of incremental learning**. This study demonstrates the importance of incremental learning (i.e., the sequential training scheme) in developing a machine learning model for steering control decisions. Without this learning scheme, the model may produce unstable prediction performance or accumulate the produced errors. The incremental learning helps the model overcome extrapolation problems, especially since the steering process may inherently involve prediction outside the training

boundaries.

These results indicate that the steering control model will be more accurate if the training is conducted in a smaller step, e.g., every few seconds during TBM advancement. This will substantially increase the computation demand since the steering control parameters must be decided continuously and the training must be completed immediately. This highlights the importance of computability and scalability in developing steering control decision models, which can be an interesting area for future work. In addition, similar to Chapters 7 and 9, more studies should be done to investigate the effects of different machine learning algorithms, their behavior in extrapolation problems, and the change of their hyperparameter configuration along the tunnel alignment.

Chapter 12

Conclusions

12.1 Summary and Conclusions

This study has proposed an integrated framework of artificial intelligence (AI) systems for earth pressure balance (EPB) tunnel boring machine (TBM) tunneling. The proposed framework was developed based on the feedback loop control system. The AI systems were developed using data-driven models by employing machine learning algorithms, and structured to follow the human cognitive model, i.e., sensing, perceiving, and decision-making. The findings of this study are summarized below.

- **Effects of data aggregation**. This study has identified that different aggregation levels do not substantially affect the central tendency of TBM data. However, it changes the data variance and distribution. Data with a coarser aggregation level tend to have lower variance. The effects on data distribution are less predictable. These characteristics may affect data-driven models. In the conditions of sufficient observation data and predictor features, models developed using different data aggregation levels produce comparable prediction trends. However, models with a coarser aggregation level enjoy higher prediction performance due to the lower variance. Interestingly, if the predictors are sufficient to represent the response feature, models with different aggregation levels may produce similar feature importance ranks. This stable feature importance rank may indicate the stability of the model.

- **Effects of feature selection**. This study has identified that TBM operation data contain severe multicollinearity. This demonstrates the importance of feature selection when developing TBM data-driven models. Models with selected features can produce comparable prediction performance to models with the original features, with less computation cost and without multicollinearity-related issues. Furthermore, this study has shown that knowledge-guided TBM feature selection offers benefits over embedded machine learning-based feature selections. The models can produce relatively consistent

feature importance in different tunneling cases. This indicates better generalizability of the model.

- **Perceiving geologic transitions**. This study has proposed a supervised AI system to interpret the encountered geologic conditions along the tunnel alignment based on TBM operation data. Different from most previous studies, the proposed system can infer the geologic transitions in real-time during tunneling. The geologic transition detection can be performed using two approaches: (i) classification, by representing geologic conditions as probabilities of several geologic labels, and (ii) regression, by representing the thickness of every geologic deposit at borehole locations as numerical percentages. The regression approach is computationally more expensive than the classification approach since it requires multiple regression models to be built. However, it enables better prediction performance from early tunneling phases, where the classification approach performs poorly due to limited available training data.

  The proposed system was developed using Random Forests (RF). This study shows the model can deliver stable and decent prediction performance with simple hyperparameter tuning. This is important considering the sequential process in tunneling, where the model needs to be dynamically retrained whenever new training data become available (i.e., arriving at a borehole location). This study has provided quantified evidence that TBM data contain information about the geologic conditions. Features related to the cutter, thrust, and ground conditioners appear to have relatively larger "weights" in the geologic interpretation model. However, information on a single feature is inadequate for the interpretation, even if it is from the high-rank features. This indicates that it is the interactions of the features that contain the information.

- **Perceiving geologic anomalies**. The unsupervised approach can be necessary for tunneling due to the limited labeled data for training (i.e., boreholes). This study has proposed an unsupervised AI system to detect the encountered geologic anomalies along the tunnel alignment based on TBM operation data. The proposed system offers a more interpretable anomaly detection approach by combining principal component analysis (PCA) to project the data into a lower dimension space and local outlier factor (LOF) to measure the degree of the anomaly of the projected data points in real-time during tunneling. This study has demonstrated that PCA can dynamically cluster TBM data according to the geologic conditions with an appropriate selection of TBM features. Interestingly, the dynamics of the projected data points may not substantially affect the clusters, as similar geologic conditions are always grouped together. This study has also demonstrated that LOF can be a sensible measure to detect changing geology.

- **Connecting multi-source data and perceiving ground responses**. This study has proposed a supervised AI system to connect TBM features to the ground monitoring data. Different from the conventional methods for estimating tunneling-induced

ground movements, the proposed system can estimate any shape of ground movements solely based on TBM operation data and tunnel spatial geometries. No prior assumptions on the ground movement shape, geologic material parameters, and the expected ground loss are required. Furthermore, unlike most previous data-driven models, the model can estimate tunneling-induced ground movements in real-time during tunneling, and in both longitudinal and cross sections. This enables quantitative investigation of the interactions between TBM control parameters and the ground responses at different locations (i.e., before, during, and after TBM passing).

The segmentation analysis has quantitatively corroborated that tunneling-induced ground movements can be divided into several segments relative to the TBM positions. This indicates different ground-machine interaction mechanisms at each ground movement segment. Furthermore, the feature importance analysis has revealed that each segment may be governed by different TBM features. Features related to the steering and pressure controls appear to strongly influence the induced ground movements during TBM passing. These features are not typically considered in conventional tunneling-induced ground movement estimation methods.

- **Modeling ground-machine-human interactions**. This study has introduced a combination of Bayesian Networks and structure learning algorithms (BNSL) as a tool to model and explore the causal effect interactions of ground conditions, TBM behaviors, and operator decisions contained in TBM operation data. This study has demonstrated that Bayesian networks can quantitatively and systematically model the interactions of TBM features based on data in a compact and interpretable representation. This integrated representation has never been investigated before and can be essential to understand TBM tunneling processes better and constructing a robust decision-making model.

- **Making decisions on control parameters**. This study has proposed an AI system to steer a TBM along the tunnel alignment. Different from most previous studies, the proposed system employed a multi-output supervised learning algorithm to simultaneously determine multiple steering control parameters based on TBM operation data. The capability of performing multi-output predictions is vital since TBM driving involves making concurrent decisions of multiple control parameters. This study has also shown the importance of the incremental learning scheme and the scalability of AI systems in tunneling.

In addition to the findings, this study has shown that TBM data contain valuable information that can be extracted to benefit the tunneling process. However, the data have complex and nonlinear relationships since it is produced by causal effect interactions during tunneling and thus cannot be handled using traditional data analysis techniques. This highlights the importance of machine learning in tunneling. In this complexity, the nonlinear and nonparametric machine learning models offer advantages over the conventional linear

and parametric models by producing not only better prediction performance but also better stability and generalizability.

These findings are expected to provide some answers for two main challenges in tunneling, i.e., (i) reducing the uncertainty by exploiting latent information in massive tunneling data that are not fully utilized in current practices and (ii) facilitating more quantitative and systematic perceiving and decision-making systems for complex tunneling processes. This will minimize laborious, subjective, and biased manual systems. This development is envisioned to be a building block that can increase tunneling performance, reduce risk and cost, and advance the fundamental development of autonomous TBM technology.

12.2 Future Works

It is evident that the implementation of AI and data-driven modeling in tunneling will continue to advance. This is motivated by the drive toward more autonomy in the tunneling process. The advancement will be boosted by the growth of data production in the tunneling process and the maturity of machine learning domain knowledge. In light of the findings from this study, insights for future challenges and opportunities are briefly discussed below.

- **Developing an integrated database for global tunneling operation data.** Tunneling data is not openly available and is often separately kept by different stakeholders. Furthermore, most published studies did not publish the data due to confidentiality issues from the data owners. Therefore, almost all previous studies were conducted using different data sets. Since data is one of the essential building blocks of data-driven models, this data availability issue limits the reproducibility and generalization of the developed models. Some questions may arise from this issue, will it be possible to create an integrated database for global tunneling operation data? Considering the massive data size, how should the data be structured? How to manage the computation issues? Governments and agencies that own the tunnel infrastructures may have an important role in this initiative. This database will enable everyone to contribute to advancing the development of AI in tunneling.

- **Developing a unified and reproducible modeling scheme.** Until now, there has been no unified procedure to build data-driven models for tunneling data. Every previous study was done using different modeling procedures, e.g., for their data preparation methods, data splitting schemes, selection of input features, hyperparameter tuning methods, validation methods, and model evaluation methods. This may be one of the reasons why many previous studies produced conflicting results. Furthermore, many of the studies provided obscure descriptions of their modeling procedures and did not publish the code openly. This causes reproducibility issues. Some questions may arise from this issue: Would it be possible to provide a unified scheme or guidance for developing data-driven models in tunneling? Note that the scheme should include not only the modeling part but also characterizing and preparing the data (Anik &

Bunt, 2021). The next question is, how to push reproducibility culture in tunneling communities? This is critical to provide solid foundations for scientific discoveries and engineering innovations in tunneling.

- **Developing a novel framework to evaluate the generalizability of machine learning models in dynamic sequential data environments.** The generalizability of machine learning models has been typically seen from the perspective of the bias-variance trade-off, i.e., an optimal spot between underfitting and overfitting. In this perspective, the generalizability of a model can be evaluated by finding minimum error in the validation data set. This approach is applicable to static data sets (Jankowsky & Schroeders, 2022). However, actual tunneling is done in dynamic sequential environments. Thus, the models should continuously be retrained dynamically in sequence using the generated streaming data. This leads to questions: How should model generalizability be seen in dynamic sequential training? How should it be measured? How to develop a generalizable model in this environment? This may require designing novel modeling and evaluation frameworks customized for data-driven modeling in tunneling.

- **Developing learning algorithms that can extrapolate.** Due to the changing ground conditions, a deployed machine learning model will be exposed to new environments at every advance during tunneling. This means the model will continuously make predictions in new ranges of data, which can be outside its training boundary. Supervised learning is a powerful interpolator tool that can find complex patterns within high-dimensional data without predefined physical laws and assumptions. However, it may perform poorly in extrapolation problems where the conditions are outside the training boundaries (Ebert et al., 2014; Hooker, 2004). This becomes a fundamental problem in deploying machine learning models in tunneling. This leads to questions: How can machine learning be reliable in making predictions for data outside its training boundary? How to evaluate this extrapolation reliability? Or, more radically, can we build machine learning algorithms that can extrapolate? Developing algorithms that can extrapolate remains an open question in other AI and machine learning research communities (Webb et al., 2020).

- **Developing tools to scrutinize machine learning models.** No matter how advanced the employed AI systems are, in the end, tunneling stakeholders hold the engineering responsibilities. These stakeholders may not be comfortable being responsible for something they cannot comprehend, especially since tunneling presents substantial risks. The consequences of wrong decisions can be catastrophically expensive and may involve casualties. Unfortunately, machine learning has been notoriously known as "black box" models (Rudin & Radin, 2019). This leads to questions: What tools are available to scrutinize machine learning models? Are these tools adequate to elucidate and understand how the model comes at the results? How to develop or customize the tools to be appropriate for data-driven models in tunneling? Interpretable machine learning is still an active research area, even in other AI and machine learning com-

munities (Linardatos et al., 2021; Ras et al., 2022). Tunnel engineers and researchers should engage more in this problem since it is critical to facilitate the deployment of AI systems in tunneling.

- **Developing scalable systems and computing strategies.** TBM operation data are generated every few seconds. This produces enormous training data for every batch in the dynamic sequential training process. This training must be completed instantaneously to be able to make decisions (predictions) for the subsequent TBM advance. Training machine learning models with massive streaming data in a restricted time constraint can be computationally challenging (Gomes et al., 2019). Therefore, computation bottlenecks will be an issue in deploying machine learning models in tunneling. This leads to questions: How should the massive streaming data be handled during tunneling? How should the computation be managed? How to increase the computation performance with limited computing resources owned by the tunneling contractors? Developing scalable AI systems will be a requirement in actual tunneling projects.

Bibliography

Alsahly, A., Marwan, A., & Meschke, G. (2019). TBM drive along curved alignments: Model based prognosis of shield movement. In D. Peila, G. Viggiani, & T. Celestino (Eds.), *Tunnels and Underground Cities: Engineering and Innovation meet Archaeology, Architecture and Art* (1st ed., pp. 1637–1646). CRC Press. https://doi.org/10.1201/9780429424441-173

Alsahly, A., Stascheit, J., & Meschke, G. (2016). Advanced finite element modeling of excavation and advancement processes in mechanized tunneling. *Advances in Engineering Software*, *100*, 198–214. https://doi.org/10.1016/j.advengsoft.2016.07.011

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347. https://doi.org/10.1093/bioinformatics/btq134

Amrhein, C. G. (1995). Searching for the Elusive Aggregation Effect: Evidence from Statistical Simulations [Publisher: SAGE Publications Ltd]. *Environment and Planning A: Economy and Space*, *27*(1), 105–119. https://doi.org/10.1068/a270105

Anagnostou, G., & Kovári, K. (1996). Face stability conditions with earth-pressure-balanced shields. *Tunnelling and Underground Space Technology*, *11*(2), 165–173. https://doi.org/10.1016/0886-7798(96)00017-X

Ang, K. H., Chong, G., & Li, Y. (2005). PID control system analysis, design, and technology [Conference Name: IEEE Transactions on Control Systems Technology]. *IEEE Transactions on Control Systems Technology*, *13*(4), 559–576. https://doi.org/10.1109/TCST.2005.847331

Anik, A. I., & Bunt, A. (2021). Data-Centric Explanations: Explaining Training Data of Machine Learning Systems to Promote Transparency. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3411764.3445736

Apoji, D., Fujita, Y., & Soga, K. (2022a). Exploring Interactions among EPBM Features using Bayesian Networks [type: article]. *World Tunnel Congress*. https://doi.org/10.31224/2141

Apoji, D., Fujita, Y., & Soga, K. (2022b). Soil Classification and Feature Importance of EPBM Data Using Random Forests [Publisher: American Society of Civil Engineers], 520–528. https://doi.org/10.1061/9780784484029.052

Åström, K. J., & Murray, R. M. (2008). *Feedback systems: An introduction for scientists and engineers* [OCLC: ocn183179623]. Princeton University Press.

Ates, U., Bilgin, N., & Copur, H. (2014). Estimating torque, thrust and other design parameters of different type TBMs with some criticism to TBMs used in Turkish tunneling projects. *Tunnelling and Underground Space Technology*, *40*, 46–63. https://doi.org/10.1016/j.tust.2013.09.004

Avgerinos, V., Potts, D. M., Standing, J. R., & Wan, M. S. P. (2018). Predicting tunnelling-induced ground movements and interpreting field measurements using numerical analysis: Crossrail case study at Hyde Park. *Géotechnique*, *68*(1), 31–49. https://doi.org/10.1680/jgeot.16.P.219

Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., Jesus, L., Berriel, R., Paixão, T. M., Mutz, F., de Paula Veronese, L., Oliveira-Santos, T., & De Souza, A. F. (2021). Self-driving cars: A survey. *Expert Systems with Applications*, *165*, 113816. https://doi.org/10.1016/j.eswa.2020.113816

Bernard, S., Heutte, L., & Adam, S. (2009). Influence of Hyperparameters on Random Forest Accuracy. In J. A. Benediktsson, J. Kittler, & F. Roli (Eds.), *Multiple Classifier Systems* (pp. 171–180). Springer. https://doi.org/10.1007/978-3-642-02326-2_18

Bhowmik, A. (2018). *Learning from Aggregated Data* (Doctoral dissertation). University of Texas at Austin.

Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, *25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7

Borghi, F. X. (2006). *Soil conditioning for pipe-jacking and tunnelling* (Doctoral dissertation). University of Cambridge.

Bouayad, D., & Emeriault, F. (2017). Modeling the relationship between ground surface settlements induced by shield tunneling and the operational and geological parameters based on the hybrid PCA/ANFIS method. *Tunnelling and Underground Space Technology*, *68*, 142–152. https://doi.org/10.1016/j.tust.2017.03.011

Boubou, R., Emeriault, F., & Kastner, R. (2010). Artificial neural network application for the prediction of ground surface movements induced by shield tunnelling [Publisher: NRC Research Press]. *Canadian Geotechnical Journal*, *47*(11), 1214–1233. https://doi.org/10.1139/T10-023

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. *ACM SIGMOD Record*, *29*(2), 93–104. https://doi.org/10.1145/335191.335388

Broere, W. (2016). Urban underground space: Solving the problems of today's cities. *Tunnelling and Underground Space Technology*, *55*, 245–248. https://doi.org/10.1016/j.tust.2015.11.012

Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *113*(15), 3932–3937. https://doi.org/10.1073/pnas.1517384113

Campello, R. J. G. B., Moulavi, D., & Sander, J. (2013). Density-Based Clustering Based on Hierarchical Density Estimates. In J. Pei, V. S. Tseng, L. Cao, H. Motoda, & G. Xu (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 160–172). Springer. https://doi.org/10.1007/978-3-642-37456-2_14

Cao, B.-T., Saadallah, A., Egorov, A., Freitag, S., Meschke, G., & Morik, K. (2021). Online Geological Anomaly Detection Using Machine Learning in Mechanized Tunneling. In M. Barla, A. Di Donna, & D. Sterpi (Eds.), *Challenges and Innovations in Geomechanics* (pp. 323–330). Springer International Publishing. https://doi.org/10.1007/978-3-030-64514-4_28

Chanchaya, C., & Suwansawat, S. (2014). Directional control of EPB shield in spiral curve. In C. Yoo, S.-W. Park, B. Kim, & H. Ban (Eds.), *Geotechnical Aspects of Underground Construction in Soft Ground* (pp. 561–566). CRC Press. https://doi.org/10.1201/b17240-105

Chen, B.-C., Chen, L., Ramakrishnan, R., & Musicant, D. (2006). Learning from Aggregate Views [ISSN: 2375-026X]. *22nd International Conference on Data Engineering (ICDE'06)*, 3–3. https://doi.org/10.1109/ICDE.2006.86

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. https://doi.org/10.1016/j.sandf.2018.11.005

Chen, R., Zhang, P., Wu, H., Wang, Z., & Zhong, Z. (2019). Prediction of shield tunneling-induced ground settlement using machine learning techniques. *Frontiers of Structural and Civil Engineering*, *13*(6), 1363–1378. https://doi.org/10.1007/s11709-019-0561-3

Chen, S., Liu, B., Qian, M., & Zhang, C. (2009). Kernel K-means Based Framework for Aggregate Outputs Classification [ISSN: 2375-9259]. *2009 IEEE International Conference on Data Mining Workshops*, 356–361. https://doi.org/10.1109/ICDMW.2009.33

Chickering, D. M. (2002). Learning Equivalence Classes of Bayesian-Network Structures. *Journal of Machine Learning Research*, *2*(Feb), 445–498. Retrieved September 1, 2021, from https://www.jmlr.org/papers/v2/chickering02a

Clark, W. a. V., & Avery, K. L. (1976). The Effects of Data Aggregation in Statistical Analysis [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1538-4632.1976.tb00549.x]. *Geographical Analysis*, *8*(4), 428–438. https://doi.org/10.1111/j.1538-4632.1976.tb00549.x

Dornfeld, A. (2014). Bertha Is Still Stuck But 'Everything's Good,' Official Says [Section: News]. Retrieved November 24, 2022, from https://archive.kuow.org/news/2014-09-19/bertha-is-still-stuck-but-everythings-good-official-says

Du, X., Xu, H., & Zhu, F. (2021). Understanding the Effect of Hyperparameter Optimization on Machine Learning Models for Structure Design Problems. *Computer-Aided Design*, *135*, 103013. https://doi.org/10.1016/j.cad.2021.103013

Dubois, Y., Dagan, G., Hupkes, D., & Bruni, E. (2020). Location Attention for Extrapolation to Longer Sequences. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 403–413. https://doi.org/10.18653/v1/2020.acl-main.39

Ebert, T., Belz, J., & Nelles, O. (2014). Interpolation and extrapolation: Comparison of definitions and survey of algorithms for convex and concave hulls. *2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, 310–314. https://doi.org/10.1109/CIDM.2014.7008683

Efheij, H., Albagul, A., & Ammar Albraiki, N. (2019). Comparison of Model Predictive Control and PID Controller in Real Time Process Control System [ISSN: 2573-539X]. *2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, 64–69. https://doi.org/10.1109/STA.2019.8717271

EFNARC. (2005). Specifications and Guidelines for the use of specialist products for Mechanized Tunnelling (TBM) in Soft Ground and Hard Rock. *Recommendation of European Federation of Producers and Contractors of Specialist Products for Structures*.

Efron, B. (2020). Prediction, Estimation, and Attribution [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/insr.12409]. *International Statistical Review*, *88*(S1), S28–S59. https://doi.org/10.1111/insr.12409

Efron, N., & Read, M. (2012). *Analysing International Tunnel Costs* (tech. rep.). Worcester Polytechnic Institute.

Erharter, G. H., Marcher, T., & Reinhold, C. (2019). Comparison of artificial neural networks for TBM data classification [Num Pages: 7]. In *Rock Mechanics for Natural Resources and Infrastructure Development - Full Papers*. CRC Press.

Erharter, G. H., & Marcher, T. (2020). MSAC: Towards data driven system behavior classification for TBM tunneling. *Tunnelling and Underground Space Technology*, *103*, 103466. https://doi.org/10.1016/j.tust.2020.103466

Erharter, G. H., Marcher, T., & Reinhold, C. (2020). Artificial Neural Network Based Online Rockmass Behavior Classification of TBM Data. In A. G. Correia, J. Tinoco, P. Cortez, & L. Lamas (Eds.), *Information Technology in Geo-Engineering* (pp. 178–188). Springer International Publishing. https://doi.org/10.1007/978-3-030-32029-4_16

Fox, H. (2000). *World Bank Urban Transport Strategy Review: Mass Rapid Transit in Developing Countries* (tech. rep.). World Bank. Washington, D.C.

Freedman, D. A. (1999). Ecological Inference and the Ecological Fallacy (N. J. Smelser & P. B. Baltes, Eds.).

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1), 1–22.

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 [Conference Name: IEEE Access]. *IEEE Access*, *8*, 64310–64323. https://doi.org/10.1109/ACCESS.2020.2984515

Garcia, G. R., Michau, G., Einstein, H. H., & Fink, O. (2021). Decision support system for an intelligent operator of utility tunnel boring machines. *Automation in Construction*, *131*, 103880. https://doi.org/10.1016/j.autcon.2021.103880

Garimella, R. V. (2017). *A Simple Introduction to Moving Least Squares and Local Regression Estimation* (tech. rep. LA–UR-17-24975, 1367799). https://doi.org/10.2172/1367799

Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, *31*(14), 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*(1), 3–42. https://doi.org/10.1007/s10994-006-6226-1

Glover, F., & Laguna, M. (1993). Tabu search. In C. Reeves (Ed.), *Modern Heuristic Techniques for Combinatirial Problems*. Blackwell Scientific Publications.

Godard, J. P., & Hugonnard, J. C. (1989). Appraisal of underground urban public transportation projects. *Tunnelling and Underground Space Technology*, *4*(1), 31–41. https://doi.org/10.1016/0886-7798(89)90030-8

Goerzen, C., Kong, Z., & Mettler, B. (2009). A Survey of Motion Planning Algorithms from the Perspective of Autonomous UAV Guidance. *Journal of Intelligent and Robotic Systems*, *57*(1), 65. https://doi.org/10.1007/s10846-009-9383-1

Goh, A. T. C., Zhang, W., Zhang, Y., Xiao, Y., & Xiang, Y. (2018). Determination of earth pressure balance tunnel-related maximum surface settlement: A multivariate adaptive regression splines approach. *Bulletin of Engineering Geology and the Environment*, *77*(2), 489–500. https://doi.org/10.1007/s10064-016-0937-8

Gomes, H. M., Read, J., Bifet, A., Barddal, J. P., & Gama, J. (2019). Machine learning for streaming data: State of the art, challenges, and opportunities. *ACM SIGKDD Explorations Newsletter*, *21*(2), 6–22. https://doi.org/10.1145/3373464.3373470

Gorishniy, Y., Rubachev, I., Khrulkov, V., & Babenko, A. (2021). Revisiting Deep Learning Models for Tabular Data. *Advances in Neural Information Processing Systems*, *34*, 18932–18943. Retrieved October 28, 2022, from https://proceedings.neurips.cc/paper/2021/hash/9d86d83f925f2149e9edb0ac3b49229c-Abstract.html

Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, *27*(3), 659–678. https://doi.org/10.1007/s11222-016-9646-1

Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data? [arXiv:2207.08815 [cs, stat]]. Retrieved October 12, 2022, from http://arxiv.org/abs/2207.08815

Groves, W., & Gini, M. (2015). On Optimizing Airline Ticket Purchase Timing. *ACM Transactions on Intelligent Systems and Technology*, *7*(1), 1–28. https://doi.org/10.1145/2733384

Guglielmetti, V. (Ed.). (2008). *Mechanized tunnelling in urban areas: Design methodology and construction control* [OCLC: ocn144768630]. Taylor & Francis.

Hammerer, N. (2015). *Influence of steering actions by the machine operator on the interpretation of TBM performance data* (tech. rep.). University of Innsbruck. Innsbruck, Austria.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition* (2nd ed.). Springer-Verlag. Retrieved July 21, 2019, from https://www.springer.com/gp/book/9780387848570

Heidemann, J., Silva, F., Intanagonwiwat, C., Govindan, R., Estrin, D., & Ganesan, D. (2001). Building efficient wireless sensor networks with low-level naming. *Proceedings of the eighteenth ACM symposium on Operating systems principles*, 146–159. https://doi.org/10.1145/502034.502049

Hinkley, D. V. (1971). Inference about the Change-Point from Cumulative Sum Tests [Publisher: [Oxford University Press, Biometrika Trust]]. *Biometrika*, *58*(3), 509–523. https://doi.org/10.2307/2334386

Hooker, G. (2004). *Diagnostics and extrapolation in machine learning* (Ph.D.) [ISBN: 9780496044641]. Stanford University. United States – California. Retrieved November 21, 2022, from https://www.proquest.com/docview/305128701/abstract/3879643771E44C45PQ/1

Hooker, G., Mentch, L., & Zhou, S. (2021). Unrestricted permutation forces extrapolation: Variable importance requires at least one more model, or there is no free variable importance. *Statistics and Computing*, *31*(6), 82. https://doi.org/10.1007/s11222-021-10057-z

Hoshino, T., Matsubara, K., Ozawa, Y., & Tanaka, Y. (2020). Development of Learning System for Shield Machine Atitude for Automatic Operation of Shield Machine. Retrieved October 20, 2022, from https://confit.atlas.jp/guide/event/jsce2020/subject/VI-592/detail

Hu, M., Wu, B., Zhou, W., Wu, H., Li, G., Lu, J., Yu, G., & Qin, Y. (2022). Self-driving shield: Intelligent systems, methodologies, and practice. *Automation in Construction*, *139*, 104326. https://doi.org/10.1016/j.autcon.2022.104326

Huayong, Y., Hu, S., Guofang, G., & Guoliang, H. (2009). Electro-hydraulic proportional control of thrust system for shield tunneling machine. *Automation in Construction*, *18*(7), 950–956. https://doi.org/10.1016/j.autcon.2009.04.005

Ishwaran, H., & Kogalur, U. (2022). Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC), R package version 3.1.0.

Ishwaran, H., Tang, F., Lu, M., & Kogalur, U. (2021). randomForestSRC: Multivariate splitting rule vignette. http://randomforestsrc.org/articles/mvsplit.html.

ITA, W. M. T. (2000). *Recommendations and Guidelines for Tunnel Boring Machines (TBMs)* (tech. rep.). International Tunneling Association.

Jakobsen, P. D., Langmaack, L., Dahl, F., & Breivik, T. (2013). Development of the Soft Ground Abrasion Tester (SGAT) to predict TBM tool wear, torque and thrust. *Tunnelling and Underground Space Technology*, *38*, 398–408. https://doi.org/10.1016/j.tust.2013.07.021

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R* [Google-Books-ID: qcI_AAAAQBAJ]. Springer Science & Business Media.

Jankowsky, K., & Schroeders, U. (2022). Validation and generalizability of machine learning prediction models on attrition in longitudinal studies [Publisher: SAGE Publications Ltd]. *International Journal of Behavioral Development*, *46*(2), 169–176. https://doi.org/10.1177/01650254221075034

Jiang, H., Jin, S., & Wang, C. (2010). Parameter-based data aggregation for statistical information extraction in wireless sensor networks. *IEEE Transactions on Vehicular Technology*, *59*(8), 3992–4001.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

JSCE. (2016). *Standard Specifications for Tunneling: Shield Tunnels* (tech. rep.). Japan Society of Civil Engineers.

Kaliampakos, D., Benardos, A., & Mavrikos, A. (2016). A review on the economics of underground space utilization. *Tunnelling and Underground Space Technology*, *55*, 236–244. https://doi.org/10.1016/j.tust.2015.10.022

Kaliampakos, D. (2016). Underground Development: A Springboard to Make City life Better in the 21st Century. *Procedia Engineering*, *165*, 205–213. https://doi.org/10.1016/j.proeng.2016.11.792

Kammerer, L., Kronberger, G., & Kommenda, M. (2020). Data Aggregation for Reducing Training Data in Symbolic Regression. In R. Moreno-Díaz, F. Pichler, & A. Quesada-Arencibia (Eds.), *Computer Aided Systems Theory – EUROCAST 2019* (pp. 378–386). Springer International Publishing. https://doi.org/10.1007/978-3-030-45093-9_46

Kasper, T., & Meschke, G. (2004). A 3D finite element simulation model for TBM tunnelling in soft ground [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nag.395]. *International Journal for Numerical and Analytical Methods in Geomechanics*, *28*(14), 1441–1460. https://doi.org/https://doi.org/10.1002/nag.395

Kasper, T., & Meschke, G. (2006a). On the influence of face pressure, grouting pressure and TBM design in soft ground tunnelling. *Tunnelling and Underground Space Technology*, *21*(2), 160–171. https://doi.org/10.1016/j.tust.2005.06.006

Kasper, T., & Meschke, G. (2006b). A numerical study of the effect of soil and grout material properties and cover depth in shield tunnelling. *Computers and Geotechnics*, *33*(4), 234–247. https://doi.org/10.1016/j.compgeo.2006.04.004

Khedo, K., Doomun, R., & Aucharuz, S. (2010). READA: Redundancy Elimination for Accurate Data Aggregation in Wireless Sensor Networks [Number: 04 Publisher: Scientific Research Publishing]. *Wireless Sensor Network*, *02*(04), 300. https://doi.org/10.4236/wsn.2010.24041

Kim, D., Pham, K., Oh, J.-Y., Lee, S.-J., & Choi, H. (2022). Classification of surface settlement levels induced by TBM driving in urban areas using random forest with

data-driven feature selection. *Automation in Construction*, *135*, 104109. https://doi.org/10.1016/j.autcon.2021.104109

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. MIT Press.

Komiya, K., Soga, K., Akagi, H., Hagiwara, T., & Bolton, M. D. (1999). Finite Element Modelling of Excavation and Advancement Processes of a Shield Tunnelling Machine. *Soils and Foundations*, *39*(3), 37–52. https://doi.org/10.3208/sandf.39.3_37

Kong, X., Ling, X., Tang, L., Tang, W., & Zhang, Y. (2022). Random forest-based predictors for driving forces of earth pressure balance (EPB) shield tunnel boring machine (TBM). *Tunnelling and Underground Space Technology*, *122*, 104373. https://doi.org/10.1016/j.tust.2022.104373

Koyama, Y. (2003). Present status and technology of shield tunneling method in Japan. *Tunnelling and Underground Space Technology*, *18*(2), 145–159. https://doi.org/10.1016/S0886-7798(03)00040-3

Kurosawa, S. (1981). Earth Pressure Balanced Shield Tunneling [Publisher: American Society of Civil Engineers]. *Journal of the Construction Division*, *107*(4), 609–618. https://doi.org/10.1061/JCCEAZ.0000991

Li, K., & Shao, C. (2015). Optimal control for a shield machine subject to multi-point earth pressure balance. *Systems Science & Control Engineering*, *3*(1), 397–403. https://doi.org/10.1080/21642583.2015.1053004

Li, S., Xu, L. D., & Wang, X. (2013). Compressed Sensing Signal and Data Acquisition in Wireless Sensor Networks and Internet of Things [Conference Name: IEEE Transactions on Industrial Informatics]. *IEEE Transactions on Industrial Informatics*, *9*(4), 2177–2186. https://doi.org/10.1109/TII.2012.2189222

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable AI: A Review of Machine Learning Interpretability Methods [Number: 1 Publisher: Multidisciplinary Digital Publishing Institute]. *Entropy*, *23*(1), 18. https://doi.org/10.3390/e23010018

Liu, X., Shao, C., Ma, H., & Liu, R. (2011). Optimal earth pressure balance control for shield tunneling based on LS-SVM and PSO. *Automation in Construction*, *20*(4), 321–327. https://doi.org/10.1016/j.autcon.2010.11.002

Liu, X., Xu, S., & Shao, C. (2020). Optimal control of earth pressure balance of shield tunneling machine based on dual-heuristic dynamic programming [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/oca.2612]. *Optimal Control Applications and Methods*, *41*(5), 1510–1523. https://doi.org/10.1002/oca.2612

Liu, X., & Zhang, K. (2019). Earth pressure balance control of shield tunneling machine based on nonlinear least squares support vector machine model predictive control [Publisher: SAGE Publications Ltd]. *Measurement and Control*, *52*(1-2), 3–10. https://doi.org/10.1177/0020294018811333

Liu, Y., Wang, Y., & Zhang, J. (2012). New Machine Learning Algorithm: Random Forest. In B. Liu, M. Ma, & J. Chang (Eds.), *Information Computing and Applications* (pp. 246–252). Springer. https://doi.org/10.1007/978-3-642-34062-8_32

Loganathan, N., & Poulos, H. G. (1998). Analytical Prediction for Tunneling-Induced Ground Movements in Clays. *Journal of Geotechnical and Geoenvironmental Engineering*, *124*(9), 846–856. https://doi.org/10.1061/(ASCE)1090-0241(1998)124:9(846)

Maaten, L. v. d., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*(Nov), 2579–2605. Retrieved August 23, 2020, from https://jmlr.org/papers/v9/vandermaaten08a.html

Madden, S., Szewczyk, R., Franklin, M., & Culler, D. (2002). Supporting aggregate queries over ad-hoc wireless sensor networks. *Proceedings Fourth IEEE Workshop on Mobile Computing Systems and Applications*, 49–58. https://doi.org/10.1109/MCSA.2002.1017485

Maher, J. I. (2015). *Improved tunneling knowledge through robust machine learning* (PhD Thesis). Colorado School of Mines. Arthur Lakes Library.

Mahmoodzadeh, A., Mohammadi, M., Daraei, A., Farid Hama Ali, H., Kameran Al-Salihi, N., & Mohammed Dler Omer, R. (2020). Forecasting maximum surface settlement caused by urban tunneling. *Automation in Construction*, *120*, 103375. https://doi.org/10.1016/j.autcon.2020.103375

Maidl, B., Herrenknecht, M., Maidl, U., & Wehrmeyer, G. (2013). *Mechanised Shield Tunnelling* [Google-Books-ID: mxsWBRIgHPUC]. John Wiley & Sons
Seems the most relevant reference for TBM control & operation.

Mair, R. J., & Taylor, R. N. (1999). Bored tunnelling in the urban environments. *Fourteenth International Conference on Soil Mechanics and Foundation Engineering. ProceedingsInternational Society for Soil Mechanics and Foundation Engineering*, *4*.

Malley, J. D., Kruppa, J., Dasgupta, A., Malley, K. G., & Ziegler, A. (2012). Probability Machines: Consistent Probability Estimation Using Nonparametric Learning Machines. *Methods of Information in Medicine*, *51*(1), 74–81. https://doi.org/10.3414/ME00-01-0052

Marcher, T., Erharter, G. H., & Winkler, M. (2020). Machine Learning in tunnelling – Capabilities and challenges [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/geot.202000001]. *Geomechanics and Tunnelling*, *13*(2), 191–198. https://doi.org/10.1002/geot.202000001

Martins, F. F., & Miranda, T. F. (2013). Prediction of hard rock TBM penetration rate based on Data Mining techniques [Publisher: Presses des Ponts].

Membah, J., & Asa, E. (2015). Estimating cost for transportation tunnel projects: A systematic literature review. *International Journal of Construction Management*, *15*(3), 196–218. https://doi.org/10.1080/15623599.2015.1067345

Merritt, A. S. (2004). *Conditioning of clay soils for tunnelling machine screw conveyors* (Doctoral dissertation). University of Cambridge.

Merritt, A., Jefferis, S., & Storry, R. (2023). Soil conditioning for EPB tunnelling in coarse grained soils based on laboratory model tests [Pages: 788-795 Publication Title: Geotechnical Aspects of Underground Construction in Soft Ground. 2nd Edition]. https://doi.org/10.1201/9781003355595-104

Meschke, G. (2018). From advance exploration to real time steering of TBMs: A review on pertinent research in the Collaborative Research Center "Interaction Modeling in Mechanized Tunneling". *Underground Space*, *3*(1), 1–20. https://doi.org/10.1016/j.undsp.2018.01.002

Meschke, G., Cao, B. T., Freitag, S., Egorov, A., Saadallah, A., & Morik, K. (2019). Big data and simulation – A new approach for real-time TBM steering [Num Pages: 10]. In *Tunnels and Underground Cities: Engineering and Innovation meet Archaeology, Architecture and Art*. CRC Press.

Mokhtari, S., & Mooney, M. (2019). Feasibility study of EPB shield automation using deep learning. In D. Peila, G. Viggiani, & T. Celestino (Eds.), *Tunnels and Underground Cities: Engineering and Innovation meet Archaeology, Architecture and Art* (pp. 2691–2699). CRC Press. https://doi.org/10.1201/9780429424441-285

Mokhtari, S., & Mooney, M. A. (2020). Predicting EPBM advance rate performance using support vector regression modeling. *Tunnelling and Underground Space Technology*, *104*, 103520. https://doi.org/10.1016/j.tust.2020.103520

Mokhtari, S., Navidi, W., & Mooney, M. (2020). White-box regression (elastic net) modeling of earth pressure balance shield machine advance rate. *Automation in Construction*, *115*, 103208. https://doi.org/10.1016/j.autcon.2020.103208

Mori, L. (2016). *ADVANCING UNDERSTANDING OF THE RELATIONSHIP BETWEEN SOIL CONDITIONING AND EARTH PRESSURE BALANCE TUNNEL BORING MACHINE CHAMBER AND SHIELD ANNULUS BEHAVIOR* (Doctoral dissertation). Colorado School of Mines.

Musicant, D. R., Christensen, J. M., & Olson, J. F. (2007). Supervised Learning by Training on Aggregate Outputs [ISSN: 2374-8486]. *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, 252–261. https://doi.org/10.1109/ICDM.2007.50

Nicodemus, K. K., Malley, J. D., Strobl, C., & Ziegler, A. (2010). The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinformatics*, *11*(1), 110. https://doi.org/10.1186/1471-2105-11-110

Ning, Z., Galisson, L., & Smith, P. (2019). Case Study: Geotechnical Instrumentation and Monitoring of Alaskan Way Viaduct Replacement Project. *GSP 313*, 10.

Ocak, I., & Seker, S. E. (2013). Calculation of surface settlements caused by EPBM tunneling using artificial neural network, SVM, and Gaussian processes. *Environmental Earth Sciences*, *70*(3), 1263–1276. https://doi.org/10.1007/s12665-012-2214-x

O'Shea, M. (2005). Sensing, perceiving, and acting. In M. O'Shea (Ed.), *The Brain: A Very Short Introduction* (p. 0). Oxford University Press. https://doi.org/10.1093/actrade/9780192853929.003.0005

Paden, B., Čáp, M., Yong, S. Z., Yershov, D., & Frazzoli, E. (2016). A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles [Conference Name: IEEE Transactions on Intelligent Vehicles]. *IEEE Transactions on Intelligent Vehicles*, *1*(1), 33–55. https://doi.org/10.1109/TIV.2016.2578706

Paraskevopoulou, C., & Boutsis, G. (2020). Cost Overruns in Tunnelling Projects: Investigating the Impact of Geological and Geotechnical Uncertainty Using Case Studies

[Number: 9 Publisher: Multidisciplinary Digital Publishing Institute]. *Infrastructures*, *5*(9), 73. https://doi.org/10.3390/infrastructures5090073

Paulheim, H. (2017). Knowledge graph refinement: A survey of approaches and evaluation methods [Publisher: IOS Press]. *Semantic Web*, *8*(3), 489–508. https://doi.org/10.3233/SW-160218

Peck, R. B. (1969). Deep Excavation and Tunneling in Soft Ground. State-of-the-Art Report. *Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering, Mexico, 225-325.*

Peila, D. (2014). Soil conditioning for EPB shield tunnelling. *KSCE Journal of Civil Engineering*, *18*(3), 831–836. https://doi.org/10.1007/s12205-014-0023-3

Peila, D., Martinelli, D., Todaro, C., & Luciani, A. (2019). Soil conditioning in EPB shield tunnelling – An overview of laboratory tests [Publisher: Wiley-Blackwell]. *Geomechanik und Tunnelbau*, *12*(5), 491–498. https://doi.org/10.1002/geot.201900021

Pinto, F., & Whittle, A. J. (2014). Ground Movements due to Shallow Tunnels in Soft Ground. I: Analytical Solutions [Publisher: American Society of Civil Engineers]. *Journal of Geotechnical and Geoenvironmental Engineering*, *140*(4), 04013040. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000948

Pinto, F., Zymnis, D. M., & Whittle, A. J. (2014). Ground Movements due to Shallow Tunnels in Soft Ground. II: Analytical Interpretation and Prediction. *Journal of Geotechnical and Geoenvironmental Engineering*, *140*(4), 04013041. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000947

Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/widm.1301]. *WIREs Data Mining and Knowledge Discovery*, *9*(3), e1301. https://doi.org/https://doi.org/10.1002/widm.1301

Qi, Y. (2012). Random Forest for Bioinformatics. In C. Zhang & Y. Ma (Eds.), *Ensemble Machine Learning* (pp. 307–323). Springer US. https://doi.org/10.1007/978-1-4419-9326-7_11

Radovanovic, S., Vukicevic, M., Kovacevic, A., Stiglic, G., & Obradovic, Z. (2015). Domain knowledge Based Hierarchical Feature Selection for 30-Day Hospital Readmission Prediction. In J. H. Holmes, R. Bellazzi, L. Sacchi, & N. Peek (Eds.), *Artificial Intelligence in Medicine* (pp. 96–100). Springer International Publishing. https://doi.org/10.1007/978-3-319-19551-3_11

Raghu, V. K., Ge, X., Chrysanthis, P. K., & Benos, P. V. (2017). Integrated Theory-and Data-Driven Feature Selection in Gene Expression Data Analysis [ISSN: 2375-026X]. *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, 1525–1532. https://doi.org/10.1109/ICDE.2017.223

Ras, G., Xie, N., Gerven, M. v., & Doran, D. (2022). Explainable Deep Learning: A Field Guide for the Uninitiated. *Journal of Artificial Intelligence Research*, *73*, 329–396. https://doi.org/10.1613/jair.1.13200

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models [Conference Name: IEEE Transactions on Systems, Man, and Cybernetics]. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(3), 257–266. https://doi.org/10.1109/TSMC.1983.6313160

Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals [Publisher: [American Sociological Association, Sage Publications, Inc.]]. *American Sociological Review*, *15*(3), 351–357. https://doi.org/10.2307/2087176

Robnik-Šikonja, M., & Kononenko, I. (2003). Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning*, *53*(1), 23–69. https://doi.org/10.1023/A:1025667309714

Roby, J., & Willis, D. (2014). Achieving fast EPB advance in mixed ground: A study of contributing factors [Publisher: Society for Mining, Metallurgy, and Exploration]. *North American Tunneling: 2014 Proceedings*, 182–194.

Rudin, C., & Radin, J. (2019). Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition. *Harvard Data Science Review*, *1*(2). https://doi.org/10.1162/99608f92.5a8a3a3d

Saadallah, A., Priebe, F., & Morik, K. (2019). A Drift-Based Dynamic Ensemble Members Selection Using Clustering for Time Series Forecasting. *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2019, Würzburg, Germany, September 16–20, 2019, Proceedings, Part I*, 678–694. https://doi.org/10.1007/978-3-030-46150-8_40

Saeys, Y., Abeel, T., & Van de Peer, Y. (2008). Robust Feature Selection Using Ensemble Feature Selection Techniques. In W. Daelemans, B. Goethals, & K. Morik (Eds.), *Machine Learning and Knowledge Discovery in Databases* (pp. 313–325). Springer. https://doi.org/10.1007/978-3-540-87481-2_21

Schöning, J., Riechmann, A., & Pfisterer, H.-J. (2022). AI for Closed-Loop Control Systems: New Opportunities for Modeling, Designing, and Tuning Control Systems. *2022 14th International Conference on Machine Learning and Computing (ICMLC)*, 318–323. https://doi.org/10.1145/3529836.3529952

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, *406*, 109–120. https://doi.org/10.1016/j.ecolmodel.2019.06.002

Schwenzer, M., Ay, M., Bergs, T., & Abel, D. (2021). Review on model predictive control: An engineering perspective. *The International Journal of Advanced Manufacturing Technology*, *117*(5), 1327–1349. https://doi.org/10.1007/s00170-021-07682-3

Scutari, M., & Denis, J.-B. (2015). *Bayesian networks: With examples in R*. CRC press.

Scutari, M., & Ness, R. (2020). Bnlearn: Bayesian Network Structure Learning, Parameter Learning and Inference. Retrieved August 31, 2021, from https://CRAN.R-project.org/package=bnlearn

Seo, S. (2006). *A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets* (Doctoral dissertation). University of Pittsburgh.

Sheather, S. (2009). *A Modern Approach to Regression with R* (G. Casella, S. Fienberg, & I. Olkin, Eds.). Springer New York. https://doi.org/10.1007/978-0-387-09608-7

Sheil, B. B., Suryasentana, S. K., & Cheng, W.-C. (2020). Assessment of Anomaly Detection Methods Applied to Microtunneling. *Journal of Geotechnical and Geoenvironmental Engineering*, *146*(9), 04020094. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002326

Sheil, B. B., Suryasentana, S. K., Mooney, M. A., & Zhu, H. (2020). Machine Learning to Inform Tunnelling Operations: Recent Advances and Future Trends. *Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction*, 1–18. https://doi.org/10.1680/jsmic.20.00011

Shi, J., Ortigao, J. a. R., & Bai, J. (1998). Modular Neural Networks for Predicting Settlements during Tunneling [Publisher: American Society of Civil Engineers]. *Journal of Geotechnical and Geoenvironmental Engineering*, *124*(5), 389–395. https://doi.org/10.1061/(ASCE)1090-0241(1998)124:5(389)

Shrivastava, N., Buragohain, C., Agrawal, D., & Suri, S. (2004). Medians and beyond: New aggregation techniques for sensor networks. *Proceedings of the 2nd international conference on Embedded networked sensor systems*, 239–249. https://doi.org/10.1145/1031495.1031524

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, *81*, 84–90. https://doi.org/10.1016/j.inffus.2021.11.011

Sousa, R. L. (2010). *Risk analysis for tunneling projects* (Thesis) [Accepted: 2010-09-02T17:21:51Z]. Massachusetts Institute of Technology. Retrieved August 14, 2020, from https://dspace.mit.edu/handle/1721.1/58282

Sousa, R. L., & Einstein, H. H. (2012). Risk analysis during tunnel construction using Bayesian Networks: Porto Metro case study. *Tunnelling and Underground Space Technology*, *27*(1), 86–100. https://doi.org/10.1016/j.tust.2011.07.003

SOWJ-JV. (2013a). *Jakarta MRT CP104 Geotechnical Interpretative Report* (tech. rep.). Prepared by Mott MacDonald.

SOWJ-JV. (2013b). *Jakarta MRT CP105 Geotechnical Interpretative Report* (tech. rep.). Prepared by Mott MacDonald.

Sramoon, A., Sugimoto, M., & Kayukawa, K. (2002). Theoretical Model of Shield Behavior During Excavation. II: Application. *Journal of Geotechnical and Geoenvironmental Engineering*, *128*(2), 156–165. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:2(156)

Stack, B. (1982). *Handbook of Mining and Tunnelling Machinery*. Wiley.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, *9*(1), 307. https://doi.org/10.1186/1471-2105-9-307

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*(1), 25. https://doi.org/10.1186/1471-2105-8-25

Su, J., Wang, Y., Niu, X., Sha, S., & Yu, J. (2022). Prediction of ground surface settlement by shield tunneling using XGBoost and Bayesian Optimization. *Engineering Applications of Artificial Intelligence*, *114*, 105020. https://doi.org/10.1016/j.engappai.2022.105020

Subramanian, S. V., Jones, K., Kaddour, A., & Krieger, N. (2009). Revisiting Robinson: The perils of individualistic and ecologic fallacy. *International Journal of Epidemiology*, *38*(2), 342–360, author reply 370–373. https://doi.org/10.1093/ije/dyn359

Sugimoto, M., & Sramoon, A. (2002). Theoretical Model of Shield Behavior During Excavation. I: Theory. *Journal of Geotechnical and Geoenvironmental Engineering*, *128*(2), 138–155. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:2(138)

Sugimoto, M., Sramoon, A., Konishi, S., & Sato, Y. (2007). Simulation of Shield Tunneling Behavior along a Curved Alignment in a Multilayered Ground. *Journal of Geotechnical and Geoenvironmental Engineering*, *133*(6), 684–694. https://doi.org/10.1061/(ASCE)1090-0241(2007)133:6(684)

Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: A review [Publisher: World Scientific Publishing Co.]. *International Journal of Pattern Recognition and Artificial Intelligence*, *23*(04), 687–719. https://doi.org/10.1142/S0218001409007326

Suwansawat, S. (2002). *Earth pressure balance (EPB) shield tunneling in Bangkok : Ground response and prediction of surface settlements using artificial neural networks* (Thesis) [Accepted: 2006-03-29T18:25:26Z]. Massachusetts Institute of Technology. Retrieved August 14, 2020, from https://dspace.mit.edu/handle/1721.1/32222

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. https://doi.org/10.1016/j.tust.2005.06.007

Tang, L., & Na, S. (2021). Comparison of machine learning methods for ground settlement prediction with different tunneling datasets. *Journal of Rock Mechanics and Geotechnical Engineering*, *13*(6), 1274–1289. https://doi.org/10.1016/j.jrmge.2021.08.006

Therneau, T. M., & Atkinson, E. J. (1997). *An introduction to recursive partitioning using the RPART routines* (tech. rep.). Technical report Mayo Foundation.

Thewes, M., Budach, C., & Bezuijen, A. (2012). Foam conditioning in EPB tunnelling. *Geotechnical Aspects of Underground Construction in Soft Ground*, *127*.

Thewes, M., & Budach, C. (2010). Soil conditioning with foam during EPB tunnelling / [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/geot.201000023]. *Geomechanics and Tunnelling*, *3*(3), 256–267. https://doi.org/https://doi.org/10.1002/geot.201000023

Todaro, C., Carigi, A., Peila, L., Martinelli, D., & Peila, D. (2021). Soil conditioning tests of clay for EPB tunnelling. *Underground Space*. https://doi.org/10.1016/j.undsp.2021.11.002

Turkandi, T., Sidarto, Agustyanto, D. A., & Purbo Hadiwijoyo, M. M. (1992). Geological Map of Jakarta and Kepulauan Seribu Quadrangles, Jawa. Quadrangle: 1209-4 & 1210-1.

Uematsu, S., Nakanishi, T., Takamizawa, S., Matsueda, K., Nanbu, T., & Fukuhara, S. (1996). Development of an Automatic Direction Control System for TBM. https://doi.org/10.22260/ISARC1996/0097

United Nations, D. o. E. a. S. A., Population Division. (2019). *World Urbanization Prospects 2018: Highlights* (tech. rep. ST/ESA/SER.A/421). United Nations. Retrieved November 14, 2022, from https://population.un.org/wup/Publications/Files/WUP2018-Highlights.pdf

United Nations, D. o. E. a. S. A., Population Division. (2022). *World Population Prospects 2022: Summary of Results* (tech. rep. DESA/POP/2022/TR/NO. 3). United Nations.

Wada, K., Sugiyama, H., Nozawa, K., Honda, M., & Yamamoto, S. (2021). Guidance System for Directional Control in Shield Tunneling Using Machine Learning Techniques [Series Title: Lecture Notes in Civil Engineering]. In E. Toledo Santos & S. Scheer (Eds.), *Proceedings of the 18th International Conference on Computing in Civil and Building Engineering* (pp. 73–88). Springer International Publishing. https://doi.org/10.1007/978-3-030-51295-8_7

Wan, M. S. P., Standing, J. R., Potts, D. M., & Burland, J. B. (2017a). Measured short-term ground surface response to EPBM tunnelling in London Clay. *Géotechnique*, *67*(5), 420–445. https://doi.org/10.1680/jgeot.16.P.099

Wan, M. S. P., Standing, J. R., Potts, D. M., & Burland, J. B. (2017b). Measured short-term subsurface ground displacements from EPBM tunnelling in London Clay [Publisher: ICE Publishing]. *Géotechnique*, *67*(9), 748–779. https://doi.org/10.1680/jgeot.SIP17.P.148

Wan, M. S. P., Standing, J. R., Potts, D. M., & Burland, J. B. (2019). Pore water pressure and total horizontal stress response to EPBM tunnelling in London Clay. *Géotechnique*, *69*(5), 434–457. https://doi.org/10.1680/jgeot.17.P.309

Wang, F., Gou, B., & Qin, Y. (2013). Modeling tunneling-induced ground surface settlement development using a wavelet smooth relevance vector machine. *Computers and Geotechnics*, *54*, 125–132. https://doi.org/10.1016/j.compgeo.2013.07.004

Wang, L., Yang, X., Gong, G., & Du, J. (2018). Pose and trajectory control of shield tunneling machine in complicated stratum. *Automation in Construction*, *93*, 192–199. https://doi.org/10.1016/j.autcon.2018.05.020

Webb, T., Dulberg, Z., Frankland, S., Petrov, A., O'Reilly, R., & Cohen, J. (2020). Learning Representations that Support Extrapolation [ISSN: 2640-3498]. *Proceedings of the 37th International Conference on Machine Learning*, 10136–10146. Retrieved November 21, 2022, from https://proceedings.mlr.press/v119/webb20a.html

Westermann, H., Šavelka, J., r, Walker, V. R., Ashley, K. D., & Benyekhlef, K. (2021). Data-Centric Machine Learning: Improving Model Performance and Understanding Through Dataset Analysis [Publisher: IOS Press]. *Legal Knowledge and Information Systems*, 54–57. https://doi.org/10.3233/FAIA210316

Wright, M. N., & Ziegler, A. (2017). Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R [arXiv: 1508.04409]. *Journal of Statistical Software*, *77*(1). https://doi.org/10.18637/jss.v077.i01

WSDOT. (2010a). *Geotechnical and Environmental Data Report* (tech. rep.). Prepared by Shannon & Wilson, Inc.

WSDOT. (2010b). *Interim Report CT-6: Geologic Characterization* (tech. rep.). Prepared by Shannon & Wilson, Inc.

WSDOT. (2010c). *Revised Geotechnical Baseline Report* (tech. rep.). Prepared by Parsons Brinckerhoff, Shannon & Wilson, Inc.

WSDOT. (2013). SR 99 Tunneling machine animation. Retrieved November 24, 2022, from https://www.youtube.com/watch?v=guWkPRReUaE

Wu, W., Gong, G., Chen, Y., & Zhou, X. (2022). Performance Analysis of Electro-Hydraulic Thrust System of TBM Based on Fuzzy PID Controller [Number: 3 Publisher: Multidisciplinary Digital Publishing Institute]. *Energies*, *15*(3), 959. https://doi.org/10.3390/en15030959

Xie, H., Duan, X., Yang, H., & Liu, Z. (2012). Automatic trajectory tracking control of shield tunneling machine under complex stratum working condition. *Tunnelling and Underground Space Technology*, *32*, 87–97. https://doi.org/10.1016/j.tust.2012.06.002

Xiong, J. L. J., Shen, S. L. K., Batty, R. J., & Ho, J. C. J. (2020). The Pursuit of an Autonomous Tunnel Boring Machine, 139–146.

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. https://doi.org/10.1016/j.tust.2020.103699

Yamamoto, T., Shirasagi, S., Yamamoto, S., Mito, Y., & Aoki, K. (2003). Evaluation of the geological condition ahead of the tunnel face by geostatistical techniques using TBM driving data. *Tunnelling and Underground Space Technology*, *18*(2), 213–221. https://doi.org/10.1016/S0886-7798(03)00030-0

Yang, H., Shi, H., Gong, G., & Hu, G. (2009). Earth pressure balance control for EPB shield. *Science in China Series E: Technological Sciences*, *52*(10), 2840–2848. https://doi.org/10.1007/s11431-009-0245-7

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061

Ye, X., Wang, S., Yang, J., Sheng, D., & Xiao, C. (2017). Soil Conditioning for EPB Shield Tunneling in Argillaceous Siltstone with High Content of Clay Minerals: Case Study. *International Journal of Geomechanics*, *17*(4), 05016002. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000791

Ye, Z., Abouzeid, A. A., & Ai, J. (2009). Optimal Stochastic Policies for Distributed Data Aggregation in Wireless Sensor Networks [Conference Name: IEEE/ACM Transactions on Networking]. *IEEE/ACM Transactions on Networking*, *17*(5), 1494–1507. https://doi.org/10.1109/TNET.2008.2011644

Yu, B., & Kumbier, K. (2020). Veridical data science. *Proceedings of the National Academy of Sciences*, *117*(8), 3920–3929. https://doi.org/10.1073/pnas.1901326117

Yu, H. (2021). *USING TBM-GENERATED DATA FOR PRESSURE MODELING, GROUND CHARACTERIZATION AND OPTIMAL CONTROL IN EARTH PRESSURE BALANCE TUNNELING* (Doctoral dissertation). Colorado School of Mines.

Yue, M., Sun, W., & Hu, P. (2012). Dynamic coordinated control of attitude correction for the shield tunneling based on load observer. *Automation in Construction*, *24*, 24–29. https://doi.org/10.1016/j.autcon.2012.02.017

Zeger, S. L., & Liang, K.-Y. (1992). An overview of methods for the analysis of longitudinal data [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.4780111406]. *Statistics in Medicine*, *11*(14-15), 1825–1839. https://doi.org/10.1002/sim.4780111406

Zhang, H., Nettleton, D., & Zhu, Z. (2017). Regression-Enhanced Random Forests. *Statistics Conference Proceedings, Presentations and Posters*, *9*.

Zhang, P., Chen, R.-P., & Wu, H.-N. (2019). Real-time analysis and regulation of EPB shield steering using Random Forest. *Automation in Construction*, *106*, 102860. https://doi.org/10.1016/j.autcon.2019.102860

Zhang, P., Wu, H.-N., Chen, R.-P., & Chan, T. H. T. (2020). Hybrid meta-heuristic and machine learning algorithms for tunneling-induced settlement prediction: A comparative study. *Tunnelling and Underground Space Technology*, *99*, 103383. https://doi.org/10.1016/j.tust.2020.103383

Zhang, P., Wu, H.-N., Chen, R.-P., Dai, T., Meng, F.-Y., & Wang, H.-B. (2020). A critical evaluation of machine learning and deep learning in shield-ground interaction prediction. *Tunnelling and Underground Space Technology*, *106*, 103593. https://doi.org/10.1016/j.tust.2020.103593

Zhang, Q., Liu, Z., & Tan, J. (2019). Prediction of geological conditions for a tunnel boring machine using big operational data. *Automation in Construction*, *100*, 73–83. https://doi.org/10.1016/j.autcon.2018.12.022

Zhang, T., Ramakrishnan, R., & Livny, M. (1996). BIRCH: An efficient data clustering method for very large databases. *ACM SIGMOD Record*, *25*(2), 103–114. https://doi.org/10.1145/235968.233324

Zhang, W. G., Li, H. R., Wu, C. Z., Li, Y. Q., Liu, Z. Q., & Liu, H. L. (2021). Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. *Underground Space*, *6*(4), 353–363. https://doi.org/10.1016/j.undsp.2019.12.003

Zhang, Y., Wu, Z., Charoenphakdee, N., & Sugiyama, M. (2020). Learning from Aggregate Observations, 13.

Zhao, J., Shi, M., Hu, G., Song, X., Zhang, C., Tao, D., & Wu, W. (2019). A Data-Driven Framework for Tunnel Geological-Type Prediction Based on TBM Operating Data [Conference Name: IEEE Access]. *IEEE Access*, *7*, 66703–66713. https://doi.org/10.1109/ACCESS.2019.2917756

Zhou, J., Qiu, Y., Zhu, S., Armaghani, D. J., Li, C., Nguyen, H., & Yagiz, S. (2021). Optimization of support vector machine through the use of metaheuristic algorithms in

forecasting TBM advance rate. *Engineering Applications of Artificial Intelligence*, *97*, 104015. https://doi.org/10.1016/j.engappai.2020.104015