Research Article

Space-Time Distribution Laws of Tunnel Excavation Damaged Zones (EDZs) in Deep Mines and EDZ Prediction Modeling by Random Forest Regression

Qiang Xie and Kang Peng

1School of Civil Engineering, Central South University, Changsha 410075, China
2Hunan City University, Yiyang 413000, China
3State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing 400044, China
4College of Resources and Environmental Science, Chongqing University, Chongqing 400044, China

Correspondence should be addressed to Kang Peng; pengkang@cqu.edu.cn

Received 19 December 2018; Revised 23 April 2019; Accepted 7 May 2019; Published 29 May 2019

The formation process of EDZs (excavation damaged zones) in the roadways of deep underground mines is complex, and this process is affected by blasting disturbances, engineering excavation unloading, and adjustment of field stress. The range of an excavation damaged zone (EDZ) changes as the time and space change. These changes bring more difficulties in analyzing the stability of the surrounding rock in deep engineering and determining a reasonable support scheme. In a layered rock mass, the distribution of EDZs is more difficult to identify. In this study, an ultrasonic velocity detector in the surrounding rock was used to monitor the range of EDZs in a deep roadway which was buried in a layered rock mass with a dip angle of 20–30°. The space-time distribution laws of the range of EDZs during the excavation process of the roadway were analyzed. The monitoring results showed that the formation of an EDZ can be divided into the following stages: (1) the EDZ forms immediately after the roadway excavation, which accounts for approximately 82%–95% of all EDZs. The main factors that affect the EDZ are the blasting load, the excavation unloading, and the stress adjustment; (2) as the roadway excavation continues, the range of the EDZs increases because of the blasting excavation and stress adjustment; (3) the later excavation zone has a comparatively larger EDZ value; and (4) an asymmetric supporting technology is necessary to ensure the stability of roadways buried in layered rocks. Additionally, the predictive capability of random forest modeling is evaluated for estimating the EDZ. The root-mean-square error (RMSE) and mean absolute error (MAE) are used as reliable indicators to validate the model. The results indicate that the random forest model has good prediction capability (RMSE = 0.1613 and MAE = 0.1402).

1. Introduction

The geo-stress level has a close relationship with the fracture of surrounding rock and the onset of rock dynamic disasters in underground engineering [1–8]. In general, the geo-stress consists of two aspects: the vertical stress and horizontal stress. The vertical stress is rooted in the gravity of the overlying strata and increases by 27 MPa per kilometer with buried depth. The horizontal stress is influenced by the geological structure and has an irregular increasing trend with buried depth [9–13]. Consequently, the excavation of deep roadways can disturb the geo-stress field and lead to a redistribution of the initial stress. As a result, the radial stress decreases, and the tangential stress increases [13]. The deep rocks distributed in the area of redistribution stress are called engineering surrounding rocks. Many underground mines have stepped into the deep mining ranks. The mining depth is approximately 1000–2000 m, and the redistribution stress level is high. The deep buried rocks are subjected to high static stress and excavation disturbances, such as blasting dynamic loading and excavation unloading, which can induce weakening in the engineering rocks [14, 15], as shown in Figure 1.

Under the coupled action of high stress, blasting vibration, and unloading disturbance, the cracks in surrounding
rocks are induced to expand and the strength is weakened forming an unstable area [1, 16–18]. The unstable area is also called the excavation damage zone (EDZ) [14, 15]. The blasting excavation that influences the surrounding rocks can be better understood if the size of the EDZ is measured accurately. In addition, details about the size of the EDZ can provide a theoretical basis for supporting parameter design for underground roadways.

Theoretical analysis, numerical simulations, and field tests are commonly used to predict the size of an EDZ in underground engineering. Theoretical analyses are based mainly on different constitutive models and failure criteria, such as Mohr–Coulomb and Hoek–Brown failure criteria [19], which do not consider the effect of intermediate principal stress on EDZs. The Drucker–Prager criterion and unified strength theory have been used to solve the elastic-plastic problem of surrounding rocks in underground roadways [20, 21]. In general, the results of theoretical analyses have large gaps compared with results from field tests. Numerical simulation and field testing are two effective methods in the stability analysis of rock engineering [22–25] that have been widely applied to investigate the distribution ranges of EDZs. The size of an EDZ is a function of time and moisture and has a significant effect on its development, which is affected by rock mass structure, initial stress, and tunnel radius [26–29]. Moreover, considering that it is time-consuming to conduct numerous tests to characterize the EDZ, supervised machine learning algorithms based on random forest [29–32] models are proposed in this study, and the EDZ is predicted as the output variable.

As discussed above, many studies on the size of EDZs have been carried out. However, the time and space distribution regulation of EDZs has rarely been the focus. Thus the roadways buried at 550 m in an underground mine were chosen to study the time-space distribution laws of EDZs. The results of this study showed that the proposed analytical solutions were able to predict damage regions. Moreover, these findings can provide measures to control damage zones and can also provide theoretical guidance for underground tunnel excavation and support. Lastly, random forest [33] models are adopted to obtain the nonlinear relationship between the EDZ and various predictors. Then the developed model is assessed with cross validation, and the performance of the proposed model is compared with multiple linear regression [34].

2. Field Testing Scheme

The section shape of the roadway being tested is a three-centered arch type. The surrounding rock lithology of the roadway is violet red slate, which obviously has an inclined bedding structure. The uniaxial compressive strength of an intact slate specimen that is 50 mm in diameter and 100 mm in height is 107 MPa, which categorizes very strong rock [29]. The dip angle of the inclined stratum is approximately 20–30°. From a geo-stress survey, the value of the vertical stress is much smaller than the horizontal stress; the direction of the maximum principle stress, $\sigma_1$, is horizontal and the direction of the minimum principle stress, $\sigma_3$, is vertical. Under the special buried environment of the roadway, roof caving and floor heaving are two main rock failure types. Thus, it is necessary to carry out studies on the regular distribution of EDZs and the supporting design, as shown in Figure 2.

At present, many methods have been used to determine the size of EDZs in the field [29–32]. In these studies, a single-hole wave velocity monitoring method was applied. The primary device was a BA-II ultrasonic crack detector, whose main components were a host computer, a water plugged up block, a water pipe, an ultrasonic source probe, and receiving probes. A simple illustration of the principle of the single-hole wave velocity monitoring technique is shown in Figure 3. Seven boreholes were chosen for measuring the size of an EDZ on a section of the roadway. The testing borehole 6 was used to monitor the wave velocity. The main scheme of the technique is outlined in Figure 3.

The source and receiving probes were inserted into a borehole, and then water was injected into the borehole, and
the borehole was filled until the testing was finished. The source probe was triggered by ultrasonic waves, which were received by the receiving probe. The difference between the generation time of the source probe and the arrival time of the receiving probe was the travel time, $\Delta t$, of the ultrasonic wave, and the distance from the source probe to the receiving probe was the travel distance, $L$, of the ultrasonic wave; the travel distance of the BA-II ultrasonic crack detector was 140 mm. The formulation of the wave velocity in the surrounding rocks at each monitored point of the borehole can be expressed as

$$V_p = \frac{L}{(\Delta t)}$$  \hspace{1cm} (1)

The ultrasonic velocity decreases with the development of fractures in the surrounding rocks. That is, the velocity was high if the surrounding rocks were intact; otherwise, the velocity was low. Hence, the thicknesses of the EDZ around a roadway can be determined by analyzing the sudden change point in the wave velocity, as shown in Figure 3 (the black line is near borehole 6).

### 3. Measurement Results and Application

Three typical sections of the roadway are chosen in this study. The distribution of the chosen sections is shown in Figure 4. The total length of the drilling roadway is 40 m, and the chosen sections are located at 10 m, 20 m, and 30 m of the roadway. The excavation direction of the roadway is from Section 1 to Section 2. The excavation speed of the roadway is approximately 2.5 m per day. EDZs in every section were detected for 20 days. The interval of detection was 4 days.

The detailed relationship between the wave velocity and the length of the testing holes is shown in Figures 5–7. There is a sudden change point in the wave velocity for each borehole, that is, the size of the EDZ is clear and quantitative. According to the $V_p - L$ curves, the ultrasonic velocity increased with the increase in borehole depth. First, the wave velocity had lower values because of the large number of cracks induced in the rocks near the roadway boundary. The wave velocity in the surrounding rocks near the boundary of the roadway was approximately 1500–2000 m/s. In contrast, the wave velocity reached a high and stable value when the monitoring points were located considerably far from the roadway boundary, which was approximately 3500–3800 m/s. The location of the sudden velocity change point is the extent of the EDZ.

The size of the EDZ as observed from different holes at the same section is different. The EDZ at the top left corner (borehole 7) had the largest value, and those at the roof and top right corner (borehole 1 and borehole 2) also had greater values, which are the main reasons for roof caving. The size of the EDZ at the lower left corner (borehole 4) was also over 110 mm and was much larger than that of other boreholes, such as borehole 3, borehole 5, and borehole 6, which induced the floor heave. An asymmetric distribution regulation of EDZs in the layered rocks was found, as shown in Figures 5(d), 6(d), and 7(d). The dip angle of the rock layer has a large influence on the EDZs in an underground space. After excavation of the roadway, the surrounding rocks were impacted by blasting dynamic loads and the unloading effect; thus, the initial EDZ formed. Then the concentration of $\sigma _1$ occurred on the top and bottom floors of the roadway, and the EDZ size increased slightly. Because of the small value of $\sigma _3$, the EDZ size near the walls of the roadway also had small values. The EDZ size for different sections had similar developmental progress. In addition, if the initial size of the EDZ was small, the wave velocity of rocks located in the range of the EDZ was comparatively high.

A clear time distribution for EDZs was observed, as shown in Figure 8. The observation of the EDZs for each borehole lasted for 20 days. For a single testing borehole, the EDZ size increased with the continuous basting excavation of the roadway. However, the increase in EDZ extent was limited to approximately 5–18% of the initial size of the EDZ. The increment of the EDZ increased with an increase in the initial EDZ size.
In general, the size of the EDZs in different sections, in order from the largest to the smallest, is as follows: Section 3, Section 2, and Section 1, as shown in Figure 9. The order of the EDZ size is opposite of the excavation order of the sections. The EDZ of the later section is affected by the stress concentration and excavation blasting of the previous section.

Bolts with a 2 m length were used to support the roadway. However, roof caving and floor heaving often occurred in the roadway. Thus, based on the distribution regular of the EDZ, the supporting technology was modified as follows: the length of the bolts was changed from 2.0 m to 2.5 m at the top left corner and the lower left corner of the roadway and the length of bolts was kept at 1.5 m at the other positions. The number at the top left corner and the lower left corner of the roadway was changed from 1 to 3, as shown in Figure 10. After these changes, the occurrence of roof caving and floor heaving was reduced.

4. EDZ Prediction by the Random Forest Model

4.1. Data Set and Descriptive Analysis. In this study, the developed random forest method was used to obtain the predictive results. A group of 63 experimental results were utilized to evaluate the feasibility of the proposed model. Three relative parameters were selected as input variables to establish the EDZ prediction model, namely, \( V \) (m/s), the angle between borehole and bedding (\( \theta \)), and distance between section and working face (\( D \), m). The relationship between the EDZ and other input variables is demonstrated in the correlation matrix plot in Figure 11, which shows the pairwise relationship between parameters along with corresponding correlation coefficients for each parameter. It can also be concluded that several parameters have relatively good/meaningful correlations with each other. Particularly, it can be seen that the parameter \( A \) is highly correlated with the EDZ.

4.2. Random Forest Regression. Random forest regression is presently one of the most successful prediction methods. It is a tree-based and nonparametric ensemble technique that was proposed by Breiman [33, 35]. As a viable method for classification and regression, random forest can be applied to many engineering problems. Particularly, random forests have successfully been used in the fields of geotechnical and mining engineering, e.g., the prediction of rockburst in hard rock, pillar stability, tunnel-induced ground settlements, daily air surface temperature, and shear strength of rockfill materials [36–42]. Compared with traditional statistical methods utilizing a single standard regression tree, the RF uses a group of decision trees, such as a forest, to predict the results. In this study, the main aim is to predict the depth of an EDZ. As a nonparametric regression approach, RF is made up of a series of \( W \) trees \( \{\xi_1(X), \xi_2(X), \ldots, \xi_W(X)\} \), where \( X = [x_1, x_2, \ldots, x_D] \) is a \( D \)-dimensional input parameter that forms a forest. The ensemble generates \( P \) outputs related to each tree, \( Y_p, (p = 1, 2, \ldots, P) \). The final output is acquired through averaging all tree predictions. The training procedure is employed as follows [33, 35–42]:

(a) From the dataset, a bootstrap sample was drawn with a random sample.

(b) The bootstrap is used to evolve a tree with the following modifications: the best split within a randomly selected subset of \( m_{try} \) descriptors is chosen at each node. In the random forest algorithm, \( m_{try} \) plays the role of an essential tuning parameter. The tree grows to the maximum size and does not get pruned back.

(c) Step (b) is repeated until the number of trees (ntree) grows to a user defined number.

As for regression, random forest establishes an amount of \( K \) regression trees and averages the predicted results. The ultimate predicted values are achieved by the aggregation of all the individual trees that constitute the forest [35]. After growing to \( W \) trees \( \{\xi(x)\} \), the random forest regression predictor is calculated using the following equation:

\[
\varphi(x) = \frac{\sum_{i=1}^{W} \xi(x)}{W}.
\]

To construct each regression tree, a new training set (bootstrap sample) is chosen from the original training set. Furthermore, in order to assess the prediction accuracy of random forests, about one-third of samples do not participate in establishing models and these data are also called out of bag (OOB) samples which can be used for evaluating the model error. In many studies, the OOB error is employed as an effective tool to evaluate the generalization error.
When independent test data are used, this built-in validation feature is able to improve the tree’s generalization capability in the RF.

4.3. Performance Metric. Two common verification statistics [34–43], root-mean-square error (RMSE) and mean absolute error (MAE), are regularly employed to evaluate the performance of RF regression models. These metrics can be obtained by the following equations:

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_{i}^\text{obs} - y_{i}^\text{pred})^2}{n}},
\]

\[
\text{MAE} = \frac{\sum_{i=1}^{n}|(1 - y_{i}^\text{pred})/y_{i}^\text{obs}|}{n},
\]
where $n$ is the total number of measured data points and $y_{i,\text{obs}}$ and $y_{i,\text{pred}}$ are the observed and predicted values of the $i$th observation, respectively.

4.4. Results of Hyperparameter Tuning. In the process of the regression of the random forests, an approximate function is finally determined to predict the value of the output. First, the original datasets of EDZs are assigned randomly to either the training set or the test set. Among these, 47 sets of the available data are chosen as the training dataset. The remaining data are considered as the testing data. The test set is used for evaluating the accuracy of the function and estimating the performance of the regression model.

This section uses the abovementioned methodology for EDZ prediction. Input parameters ($A$, $V$, and $D$) in the random forest models are the factors influencing an evaluation target (EDZ). To conduct the experiments, the 63 groups of data are randomly split into two subsets with a stratified sampling method that divides the datasets into training (47) and test (16) datasets. This method also creates...
the training set by selecting from these subsets with the same outcome distribution as the whole dataset. Then, the training set is trained to establish a regression model that is used for evaluating the parameters for the random forest regression methods. To determine reasonable and optimized hyperparameters for the random forests, a fivefold CV procedure [20, 36–42] was implemented to determine the optimal parameter values based on the original training dataset. In addition, the sqrt(x) function was utilized to determine each node of the random tree where x is the number of the input variables. In this study, we will stick to tuning two parameters, namely, the number of trees (ntree) and the number of variables randomly sampled as candidates at each split (mtry). These parameters have the following effect on our random forest model.

In the application of random forest experiments, the caret package [35] in R was used to build a regression model. To investigate the behavior of the random forest hyperparameters (ntree and mtry), the model was tuned using a relatively fine search grid with CV methods, and tuning parameters were set as follows: mtry = c(1 : 3), ntree = seq(50, 1000, by = 50). RMSE was used to select the
optimal model using the smallest value. Figure 12 shows the cross-validated RMSE profile for the random forest. It can be seen that the RMSE value of the random forest model is more sensitive to mtry than to ntree. The final values used for the random forest model with a 5-fold CV procedure (Figure 12) were ntree = 650 and mtry = 2. The RMSE and MAE of the random forest model for 47 sets of training data were found to be 0.2499 and 0.2054, respectively.

4.5. Results of Independent Test Set. To validate the predictive models based on the predicted and measured (real) values, 16 testing samples were validated by the optimized random forest model. The results are presented in Figure 13 and show actual vs. predicted EDZ by linear regression and random forest algorithm using the test data. Figure 13 suggests that all of the predicted values by random forest lie within ±15% error off the line of perfect agreement. The RMSE and MAE between the observed and predicted values
of the random forest model are found to be 0.1613 and 0.1402 for test data, respectively. Additionally, the RMSE and MAE of the predicted values found using the conventional multiple linear regression method [34] were found to be 0.2075 and 0.1806 for test data, respectively. A comparison of RMSE and MAE values indicates improved performance by the random forest modeling approach compared to conventional multiple linear regression.

4.6. Variable Importance Analysis. The variable importance can be estimated easily by using varImp() for random forest models in the caret package [35]. Among all the three variables, it is noteworthy that $A$ is the most influential of the parameters among the indicators for prediction of EDZs, which is also consistent with the correlation matrix for the variables that showed the highest coefficients for these variables. The indicator $V$ is slightly sensitive. The factor $D$ is
not sensitive, and this means that this factor does not contribute to the aforementioned prediction modeling, as depicted in Figure 14. Note that the importance score for the random forest model relied on the dataset used in this work. In the future, a larger dataset is desirable to obtain more reliable importance scores for the influence of variables.

5. Limitations

The RF (random forest) approach for the size of EDZs, however, has some limitations that need to be improved in future research. First, the limitation of RF approach is that the datasets are relatively small, only a few dozen cases involved in RF prediction model. A larger dataset related to EDZs is able to raise the model’s performance and can improve the interpretability of variable importance scores. In this study, the variable importance score of D is not sensitive, which is controversial to measured EDZ; a larger dataset is expected to make further explanation. Second, a further research about more influencing variables involving time after excavation needs to be explored and more hyperparameters need to be tuned to ensure the RF algorithm more feasible. Last, other advanced supervised machine learning approaches that have been proved to show excellent predictive performance on modeling complicated nonlinear engineering problems, such as support vector machine [44] and stochastic gradient boosting approach [45], have not been investigated and compared on the size of EDZs.

6. Conclusions

In this paper, according to field testing, the EDZ size of a roadway buried in a layered rock mass was analyzed. The main conclusions are as follows:
The excavation damaged zone of an underground roadway is affected by several factors. First, the excavation blasting dynamic load and the unloading effect are the main reasons that the initial EDZ is induced. Subsequent to this, the stress concentration can enlarge the size of the EDZ and the increment of the EDZ limit.

(2) There is an asymmetric distribution regulation of EDZs in the layer rocks, and the size of the EDZ increases by approximately 5%–18% with the excavation progress of roadway. In addition, the later excavation zone has a comparatively larger value.

(3) It is necessary to design the supporting parameters according to the distribution of the EDZ, that is, an asymmetric supporting technology is necessary to maintain the stability of roadways buried in layer rocks.

(4) Compared with conventional multiple linear regression, it can be concluded that random forest regression algorithms perform well in predicting the EDZ in terms of its explanatory value and error. Random forest algorithms can effectively identify useful input parameters which affect EDZs.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to acknowledge the financial support of National Natural Science Foundation of China (no. 51774058) and supported by the Chongqing Basic Research and Frontier Exploration Project (cstc2018jcyjA3320 and cstc2016jcyjA1861), which are gratefully acknowledged.

References


References


