HHS Public Access

Author manuscript

Min Metall Explor. Author manuscript; available in PMC 2024 February 12.

Published in final edited form as:

Min Metall Explor. 2023 ; 40(4): 1399–1407. doi:10.1007/s42461-023-00800-7.

Identifying the Location and Size of an Underground Mine Fire with Simulated Ventilation Data and Random Forest Model

Yuting Xue[1], **Davood Bahrami**[1], **Lihong Zhou**[1]

[1]CDC NIOSH, Pittsburgh, PA, USA

Abstract

Underground mine fires are a threat to the safety and health of mine workers. The timely determination of the location and size of an underground fire is of great importance in developing firefighting strategies and reducing the risk of any injuries. Machine learning was used in this paper to develop a predictive model for fire location and fire size in an underground mine. The ventilation data were obtained by simulating different mine fire scenarios with MFire. The ventilation data of all airways were used as features to predict the fire location. Based on the feature importance, five airways were selected to monitor, and the airflow data of the selected airways were used to predict the fire location and fire size. An accuracy score of 0.920 was obtained for the prediction of fire location. In addition, in-depth analyses were conducted to characterize the wrong predictions with the purpose of improving the performance of the random forest model. The results show that the occurrence of fire at closely connected airways at some locations can generate misleading ventilation data for each other and the model performance can be further improved to 0.962 by grouping them. Fire size is another factor affecting the model performance and the model accuracy increases with increasing fire size. The result from this study can help mine safety personnel make informed decisions during a mine fire emergency.

Keywords

Underground mine fire; Machine learning; Airflow; Random forest; Fire location; Fire size

1 Introduction

The number of fire-related injuries in mining has been decreasing in the past decade; however, mine fire remains a safety and health threat to underground miners. The occurrence of fire in underground mines can significantly affect the ventilation network in the form of increasing airway resistance, changing airflow quantity, changing airflow direction, and unexpectedly contaminating fresh air escape routes. An underground mine fire can also produce a large amount of toxic gases and smoke and release significant levels of heat [4, 5].

✉Yuting Xue, qcj1@cdc.gov.

Conflict of Interest The authors declare no competing interests.

Disclaimer The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention. Mention of any company or product does not constitute endorsement by NIOSH.

Prompt determination of an underground fire location and its condition is of great importance for developing firefighting strategies in underground mines. The occurrence of an unknown fire can be detected by the presence of smoke and gases such as carbon monoxide (CO). Currently, the fire location determination is based on the visual reports of smoke or measured hazardous gases. The increasing availability of monitoring data from the Atmospheric Monitoring System (AMS) sensors makes it possible to automatically determine the potential fire source airways(s) [1, 13, 19]. The AMS uses environmental and air-quality sensors installed in underground mines to monitor atmospheric parameters. These parameters have been used to detect fires in underground mines. However, the AMS, especially CO sensors, is normally installed in the entries where fires likely take place, for example, in belt entries. The sensors cannot cover the whole mine. If a fire occurs in a location that is not covered, it is very likely that the fire cannot be detected promptly. This disadvantage makes way for machine learning.

Machine learning (ML) is a powerful tool for solving complex problems. Using computational methods, ML models “learn” information directly from data without relying on a predetermined equation as a model. They adaptively improve their performance with increasing training data. ML has been used to address various mining problems, including equipment mechanical failure prediction, production optimization, and ore body delineation [14]. Principal component analysis and artificial neural network were used to predict methane emissions in longwall mines in the USA [8]. The rough set theory was used to select the important parameters from 12 monitored ventilation parameters and the selected parameters were used to evaluate the risk of a ventilation network with a support vector machine (Cheng & Yang, 2012). Instead of manually adjusting regulators, Kashnikov and Levin [9] optimized the location and setting of ventilation regulators for desired flow in a mine ventilation network with ML. Recently, an SVM regression–based ML model was developed to predict the propagation of a gas explosion in underground coal mines with simulation data [11]. In addition, there are increasing applications of ML and artificial intelligence for fire detection in tunnels [7, 16, 17]. However, there is currently no available study of ML for the identification of fire location and fire size in the ventilation network of an underground mine.

An ML-based model, if properly trained, can be used to identify an underground mine fire. It requires training a ML model with ventilation data and using the trained ML model to predict mine fires with real-time monitored data. Since ML learns from experience, a good ML model needs to be trained with the ventilation data under a broad range of fire scenarios, like different fire locations and fire sizes. However, it is impractical to collect the real ventilation data for a specific ventilation network under different fire scenarios. An alternative approach is to use the simulated ventilation data. An attempt has been made to train a ML model with the simulated ventilation data under different fire scenarios to monitor the occurrence of underground mine fires [2]. Different ML algorithms commonly used by other researchers were employed to determine the fire location with simulated ventilation data and their performance was evaluated based on model accuracy and the time of training and prediction.

In this study, a detailed analysis of using ML to predict fire location and fire size in underground mines with simulated airflow data was conducted. First, a database with fire location and fire size, as output/target, and airflow data, as input/feature, was generated by simulating different fire scenarios with a calibrated ventilation network. The training dataset with the airflow data of all airways was then used to train a Random Forest classifier to classify the fire location. Based on the feature importance, the five most important airways were selected to monitor, and the classifier was further trained with the airflow data of the selected airways. In addition, the training dataset with the airflow data of the selected airways was used to train a Random Forest regressor for the fire size prediction. After training the models, detailed analyses were conducted on the wrong predictions to investigate the reasons with the purpose of improving model accuracy.

2 Database Generation

The same dataset as in the previous study was used in this study, and it is briefly described here [2]. The ventilation network of the Safety Research Coal Mine (SRCM) at Pittsburgh Mining Research Division was used. The layout of the facility is shown in Fig. 1. A total of 166 airway segments exist in this network. A simplified ventilation network setup in Ventsim is shown in Fig. 2. The ventilation model is based on a well-calibrated network from the previous work of NIOSH researchers [19].

This model is manually converted to a model input file that can be used by MFire fire simulation software within MATLAB [20]. Within the ventilation network, the airway number was used to label each airway so that the fire source location and the airflow data in the dataset could be easily identified. Forty-two airways with very low airflow, shown in light blue in Fig. 2, were excluded from the dataset. These airways are typically the ones that are closed off. The focus of the model is on the active airways with numerically significant airflows, larger than 0.05 m/s. After excluding the 42 inactive airways, the remaining 123 active airways were considered the potential fire source in this study.

The ventilation model input was fed to an automated macro to build a large set of fire scenarios of different fire sizes placed in each airway of the network. Twenty fire sizes were randomly selected from a range of 50 and 147,860 Btu/min for each airway. A total of 200 fire scenarios were generated. After running the fire simulations using MFire software, the results of each scenario in terms of airflow in every airway, fire location, and fire size were recorded in a tabular form with each scenario data being placed in one row. Each row of data represents one case, and they form the database. The database contains 24,600 cases with fire location and fire size as targets and with airflow data of 123 airways as features.

The database was divided into two datasets with the ratio of 75%−25%, 75% for training and 25% for testing. The split of the training and testing dataset was conducted at the beginning and was kept for all the following learning, testing and analyses.

3 Results and Discussions

3.1 Fire Location Prediction with All Airflow Data

Python was used for ML, and the Random Forest (RF) model from the Scikit-Learn package was used for the classification of fire location and for the prediction of fire size [15]. Due to fast processing speed and good accuracy, the RF model has been widely used in ML projects for mine safety and health [3, 6, 10, 12, 18]. As an ensemble approach, it has been identified as one of the best models in the previous study [2].

The fire location prediction is a classification problem, and the RF classifier needs to classify the 123 potential fire locations based on the airflow data of the 123 airways. The classifier was trained with the training dataset, representing 75% of all data, and was evaluated with the testing dataset. The accuracy score is used as the metrics to evaluate the model performance, and it is defined as the fraction of accurate predictions among the testing data. The result shows that there are 129 wrong predictions with 6150 cases in the testing dataset, leading to a model accuracy of 0.979.

Even with the airflow of all the airways being monitored, the fire location cannot be predicted with 100% accuracy, and there were still 129 wrong predictions. Why? A histogram plot of the wrong prediction is shown in Fig. 3. The airways with a large number of wrong predictions were analyzed. By comparing Figs. 2 and 3, it was found that the airways with a larger number of wrong predictions are connected somewhere within the ventilation network. Airways 18 and 165 with the largest and third largest number of wrong predictions are the same airway, as shown in Fig. 2. The influence of fire within these two airways can be misleading, which makes it difficult to differentiate them. This is the same situation for airways 94 and 95, which are further connected with airway 97. The third group of airways, including airways 147, 104, 7, 148, and 149, also has a higher wrong prediction rate than others. All these airways are connected and the air flows from one airway to another. The occurrence of fire in either airway may have a similar influence on the airflow distribution, potentially making it difficult to differentiate from the change in airflow.

3.2 Fire Location Prediction with the Airflow Data of Selected Airways

3.2.1 Feature selection and model training—The fire location was predicted with the airflow data of all airways in the previous section. It indicates that all the airways need to be monitored to get the airflow data in an operating mine. However, it is impractical, especially when there is a large ventilation network with hundreds of airways. A practical approach is to select a few “important” airways and use their airflow data to predict the fire location with the loss of certain model accuracy. The airflow data in each airway was considered one feature for ML. By selecting the import features, the represented airway can be selected as important airways to monitor. The selection of important features was accomplished by ranking feature importance, a parameter determined when training of the RF model with the airflow data of all airways in the previous section. Feature importance represents the usefulness of features in classifying the fire location. A high feature importance value can be obtained if one feature is frequently used in the RF model to classify the fire locations, and as a result, using a few important features will only lead to minor loss of accuracy. But how many important features should be selected? Different numbers of important features were used to predict the fire location. After a few trials, it was found that the model accuracy can be above 0.900 when five most important airways were selected. The drop of one more airway will lead to an accuracy less than 0.900. The selected airways, including 61, 34, 12, 101, and 128, are marked on the simplified ventilation network in Fig. 2. It is noticeable that the five selected airways do not concentrate in specific areas but rather are distributed evenly within the network.

During the training of the classifier with the airflow data of the five selected airways, the method of grid search with cross-validation was used to tune the model hyper-parameters. The process is shown in Fig. 4. There are various hyper-parameters for a RF classifier, such as the number of decision tree, the criterion to split the data, and the minimum amount of data for a split. These hyper-parameters control the training process and affect the performance. The grid search cross-validation method allows the search of the best hyper-parameters by exploring different combinations of the hyper-parameters. The cross-validation method involves randomly splitting the training dataset into K groups (5, by default). For each iteration, one group of data is used as the validation dataset, while the remaining groups are taken as the training dataset. The classifier with one group of hyper-parameters is fitted with the training data and is evaluated with the testing data to obtain an evaluation score. After K iterations, the average evaluation score is used to quantify the performance of the hyper-parameter combination. After comparing the average evaluation score for all the hyper-parameter combinations, the best hyper-parameters can be selected and the RF classifier with the best performance can be used as the trained classifier.

3.2.2 Results with selected airways—The loss of model accuracy was expected after dropping 118 features. The results show that there were 495 wrong predictions of the 6150 cases within the testing dataset. An accuracy score of 0.920 was obtained with the airflow data of only 5 selected airways.

In order to understand the reasons for wrong predictions and to further improve the model accuracy, a detailed analysis of the wrong predictions was conducted. A comparison of the airflow at the selected airways between the right and wrong predictions was made. The potential logic behind the comparison is that the influence of fire should be reflected in the airflow data of the selected airways; otherwise, it cannot be identified from the selected airflow data. The airflow data in the testing dataset was separated into two groups, right and wrong prediction. A statistical analysis of these two groups of data was carried out, and the results are presented with the bar chart with error bar in Fig. 5. The error bars represent the standard deviation. It can be found that the mean values of the airflows are very close, while the right predictions have higher variation than the wrong predictions. Higher variation of the airflow data in the group of right predictions indicates that the airflow of the selected airways is more significantly affected by the occurrence of fire, while the wrong predictions involve fire with negligible influence on the airflow of the selected airways. It indicates that only when the airflow is affected to a certain extent that the classifier can correctly identify the location of the fires. Three factors can potentially affect this process. The first one is the number of selected airways. As the airflow of the selected airways needs to be affected, more monitoring stations increase the possibility of change in airflow being monitored. The second factor is the relative distance of the fire to the selected airways within the ventilation network. The relative distance is not the physical distance of the fire to the selected airways. It can be affected by the airflow sequence. If the air of the fire location is flowing from or to the selected airways, the airflow of the selected airways can be easily affected. The third factor is the fire size, which potentially affects the change in airflow and will be discussed in the next section.

Figure 5 also shows that the airflow within the selected airways is relatively small. A possible reason is that the airways with small airflow are more sensitive to the change in airflow than the airways with large airflow. However, the potential problem is that they may be too sensitive to the change in airflow for real airflow data. If there is airflow fluctuation during normal mining operation or sensor error, the change in airflow can lead to wrong prediction and even false alarm.

In addition, a histogram plot of the 495 wrong predictions is presented in Fig. 6. For comparison purposes, the wrong predictions with all airflow data shown in Fig. 3 are also included. Figure 6 shows that, after dropping 118 features, there are increases in the count of wrong fire location prediction everywhere within the ventilation network. However, significant increases concentrate at certain locations and there is only a minor increase at other locations.

Three observations can be made from Fig. 6 after the dropping of most airway data. The first observation is that there are no (or a small number of) wrong predictions with either all or selected airflow data. This is the dominant case, which leads to the high accuracy of the ML model. The second and third situations all involve a significant increase in the number of wrong predictions after dropping most airway data. Due to the difference in performance with all airflow data, they were separated into different situations. The second situation is that there are no (or small number of) wrong predictions with all the airflow data and the drop of most airways leads to a significant increase in the number of wrong predictions. The examples include airways 8, 73, and 123. For the third situation, the wrong prediction rate was already higher than others with all the airflow data and the dropping of most airway data made the situation worse. The examples include airways 18, 94, 95, 147, and 165.

A detailed analysis focusing on the airways with a wrong prediction number larger than 10 was conducted. These airways, except airways 123 and 150, were separated into three groups based on their location within the ventilation network. The first group includes airways 18 and 165; the second group includes airways 73, 86, 94, 95, and 97; and the third group includes airways 7, 8, 15, 104, 147, 148, and 149. These three groups of airways are marked in Fig. 2. It can be found that these three groups cover the connected airways with a higher wrong prediction rate that was discussed with Fig. 3. Due to the large number of airways as potential fire locations, it is difficult to visualize the results with a confusion matrix. For these three groups, the airway ID, total count in testing dataset, number of wrong predictions, predicted airway ID, and the count for each predicted airway ID are summarized in Table 1.

Table 1 shows that most predicted fire locations of the wrong predictions are within the same group. As shown in Fig. 2, the airways within each group are connected and the air flows from one airway to another. Maybe due to the specific location of these three groups of airways within the ventilation network, the fire within each group of airways has a similar influence on the change of airflow in the monitored airways. The result is that the ML model cannot accurately differentiate between the airways of each group. Since the airways of each group are connected and are physically close to one another, it is reasonable to group them when predicting the fire location. The airways presented in Table 1 were merged into three airway IDs. Airways 10, 18, and 165 are represented as airway 168; airways 7, 8, 15, 104, 142, 147, 148, and 149 are merged to be airway 167 and airways 73, 85, 86, 94, 95, and 97 are represented as airway 166. These three special airways represent three areas. The model was evaluated again after merging these groups of airways. The number of wrong predictions dropped to 231 from 495, leading to a model accuracy of 0.962. A histogram plot of the wrong predictions is presented in Fig. 7. It can be found that the large numbers of wrong predictions are eliminated. For the wrong predictions summarized in Table 1, there are only two left, one for airway 166 (group 3) and one for airway 168 (group 1). It should be noted that the failure of differentiating the connected airways potentially occurs at other locations. However, due to the small number of wrong predictions, attention was not paid to analyze and group them.

3.3 Fire Size Prediction with the Airflow Data of Selected Airways

The fire size prediction is a regression problem and a RF regressor was used. The regressor was trained with the airflow data of the five selected airways discussed in the previous section. The prediction results are shown in Fig. 8, and an R^2 value of 0.977 was obtained. It indicates a good agreement between the predicted and actual fire size. As shown in Fig. 8, most of the predicted fire sizes are close to the actual values, and the data points are located along the diagonal between these two triangles. However, there are some predictions with larger error, as marked by the two triangles.

An analysis was then conducted for the fire size predictions with large error. The split of the training and testing dataset was conducted at the beginning and was kept for all the training and testing. The testing dataset for fire location prediction and fire size prediction is the same one. An analysis of the wrong fire location prediction was conducted with the fire size prediction. All the wrong fire location predictions were collected, and their airflow data were used to predict the fire size with the trained RF regressor. The predicted fire sizes with wrong fire location prediction are marked with red in Fig. 9, and the blue ones represent the right fire location prediction. It should be noted that there are 6150 data points in total and 495 data points with wrong fire location prediction. There are much more right fire location predictions than the wrong ones. Figure 9 shows that, within the triangle areas, there are more wrong fire location predictions than the right predictions, especially for the ones with a large error in fire size prediction. It indicates that the performance of the RF regressor is not good when the fire location prediction with the RF classifier is wrong, where both models use the same airflow data for training and testing.

At the same time, Fig. 9 shows that there are much more wrong fire location predictions when the fire size is smaller because the red data points are denser with smaller fire size. A histogram plot of the wrong fire location predictions with actual fire size is presented in Fig. 10. It shows that the wrong predictions drop significantly with the increase in fire size. The wrong fire location predictions with actual fire size less than 10,000 Btu/min represent 32.3% of all the wrong predictions. The potential reason for the influence of fire size on fire location prediction is that a small fire has limited influence on the airflow of the ventilation network. The negligible change in the airflow of the selected airways is difficult to use to predict the fire location correctly. This further coincides with the results in Fig. 5, where the wrong predictions of fire location have much lower variation than the right predictions. When the influence on airflow of the selected airways is limited, the variation of the airflow is reduced. The model accuracy can be further increased if only the cases with actual fire size larger than 10,000 Btu/min are used. The confidence in the model prediction increases with increasing fire size.

3.4 Limitations

The trained machine learning models for the prediction of fire location and fire size with ventilation data in underground mines show promising results. However, there are some limitations with this method. First of all, ML learns from “experience” indicating that the trained model will not work well if there are some changes making the experience invalid. In this study, the ML model learns the experience from the airflow data of a specific ventilation network. When there are large changes in the ventilation network resulting from the change in mining operations, the airflow data used to train the ML model will not fit the new ventilation network, making the trained ML model not applicable for the new network. This also applied to the situation with two different ventilation networks where the ML model trained with a ventilation network could not be used for the other one. To overcome these problems, the ML model needs to be retrained after the change or update of the ventilation network. Second, the data used for training and testing are from simulations that are conducted under some simplifications. There are significant fluctuations in the real airflow data, potentially affecting the ML model performance. Finally, the fire size evolves with time and is not a fixed value. At the early stage when the fire is small, the performance of the ML model may not be good, and the predicted results can be misleading. It is only when the fire develops to a certain size that we can obtain confident predictions.

4 Conclusions

The application of machine learning to identify the location of an underground mine fire with simulated airflow data demonstrates some promising results. An accuracy score of 0.979 was obtained with all the airflow data, and it dropped to 0.920 with the airflow data of five selected airways. The statistical analysis of the airflow in selected airways between right and wrong predictions shows that the wrong predictions involve airways with much lower airflow variation than the right predictions. It indicates that, in order to be predicted correctly, the fire needs to affect the airflow of the selected airways to a certain extent. In addition, the histogram plot of the wrong predictions with the airflow data of all and selected airways shows that the airways with a large number of wrong predictions were concentrated at specific locations. They are normally closely connected airways and the air flows from one airway to another. The occurrence of fire in either airway potentially has a similar influence on the airflow of the selected airways, making it difficult to differentiate between them. When three groups of closely connected airways with large numbers of wrong predictions for fire location were identified and merged, the accuracy score increased from 0.920 to 0.962.

ML was also used to predict the fire size with the airflow data of the five selected airways. A R^2 value of 0.977 was obtained for the relationship between the actual and predicted fire size, indicating a good agreement between the predicted and actual fire size. However, there are a small number of predictions with large error. The comparison with the fire location prediction shows that the fire size predictions with large error normally have wrong predictions for fire location as well. The analysis also shows that fire size is one critical factor affecting the fire location prediction. There is a significant reduction in the number of wrong fire location predictions with the increase in fire size. The study demonstrates that a training ML model can be used to provide additional information to help mine safety personnel make informed decisions during a mine fire emergency.

References

1. Bahrami D, Yuan L, Zhou L, Rowland JH (2019) New mine fire location model utilizing real-time gas monitoring data. Proceedings of the 17th North American Mine Ventilation Symposium, Montreal, QC, Canada
2. Bahrami D, Zhou L, Xue Y, Yuan L (2021) Application of machine learning to determine underground hazard location. Proceedings of the 18h North American Mine Ventilation Symposium, Rapid City, SD, USA
3. Chen R, Zhang P, Wu H, Wang Z, Zhong Z (2019) Prediction of shield tunneling-induced ground settlement using machine learning techniques. Front Struct Civ Eng 13(6):1363–1378
4. Gillies S, Wu H (2004) Case studies from simulating mine fires in coal mines and their effects on mine ventilation systems. Fifth Australasian Coal Operators Conference, Wollongong, NSW, Australia
5. Hansen R, Ingason H (2013) Heat release rate measurement of burning mining vehicles in an underground mine. Fire Saf J 61:12–25
6. Hasan AN, Twala B (2015) Improving single classifiers prediction accuracy for underground water pump station in a gold mine using ensemble techniques. Proceedings of IEEE EUROCON 2015 - International Conference on Computer as a Tool (EUROCON), Salamanca, Spain
7. Hong Y, Kang J, Fu C (2022) Rapid prediction of mine tunnel fire smoke movement with machine learning and supercomputing techniques. Fire Saf J 127:103492
8. Karacan CÖ (2008) Modeling and prediction of ventilation methane emissions of U.S. longwall mines using supervised artificial neural networks. Int J Coal Geol 73(34):371–387
9. Kashnikov AV, Levin L (2017) Applying machine learning techniques to mine ventilation control systems. XX IEEE International Conference on Soft Computing and Measurements (SCM), St. Petersburg, Russia
10. Lin Y, Zhou K, Li J (2018) Prediction of slope stability using four supervised learning methods. IEEE Access 6:31169–31179
11. Liu L, Liu J, Zhou Q (2021) An SVR-based machine learning model depicting the propagation of gas explosion disaster hazard. Arab J Geosci 46(10):10205–10216
12. Qi C, Fourie A, Du X, Tang X (2018) Prediction of open stope hangingwall stability using random forests. Nat Hazards 92:1179–1197

13. Smith A, Litton C (2015) The use of atmospheric monitoring system for fire detection in underground coal mines. Proceedings of the 15th North American Mine Ventilation Symposium, Blacksburg, VA, USA
14. Stewart P (2016) Machine learning enters mines. Mining Magazine. https://www.miningmagazine.com/innovation/opinion/1263928/machine-learning-enters-mines
15. VanderPlas J (2016) Python data science handbook: essential tools for working with data. O'Reilly
16. Wu X, Zhang X, Huang X, Xiao F, Usmani A (2021) A real-time forecast of tunnel fire based on numerical database and artificial intelligence. Build Simul 1–14, Tsinghua University Press
17. Wu X, Park Y, Li A, Huang X, Xiao F, Usmani A (2021) Smart detection of fire source in tunnel based on the numerical database and artificial intelligence. Fire Technol 57:657–682
18. Zhou J, Li X, Mitri HS (2015) Comparative performance of six supervised learning methods for the development of models of hard rock pillar stability prediction. Nat Hazards 79(1):291–316
19. Zhou L, Yuan L, Bahrami D, Thomas RA, Cole G, Rowland J (2019) Study on integration of real-time atmospheric monitoring system monitoring data and MFIRE simulation. Proceedings of the 17th North American Mine Ventilation Symposium, Montreal, QC, Canada
20. Zhou L, Smith AC, Yuan L (2016) New improvements to MFIRE to enhance fire modeling capabilities. Min Eng 68(6):45 [PubMed: 27375301]

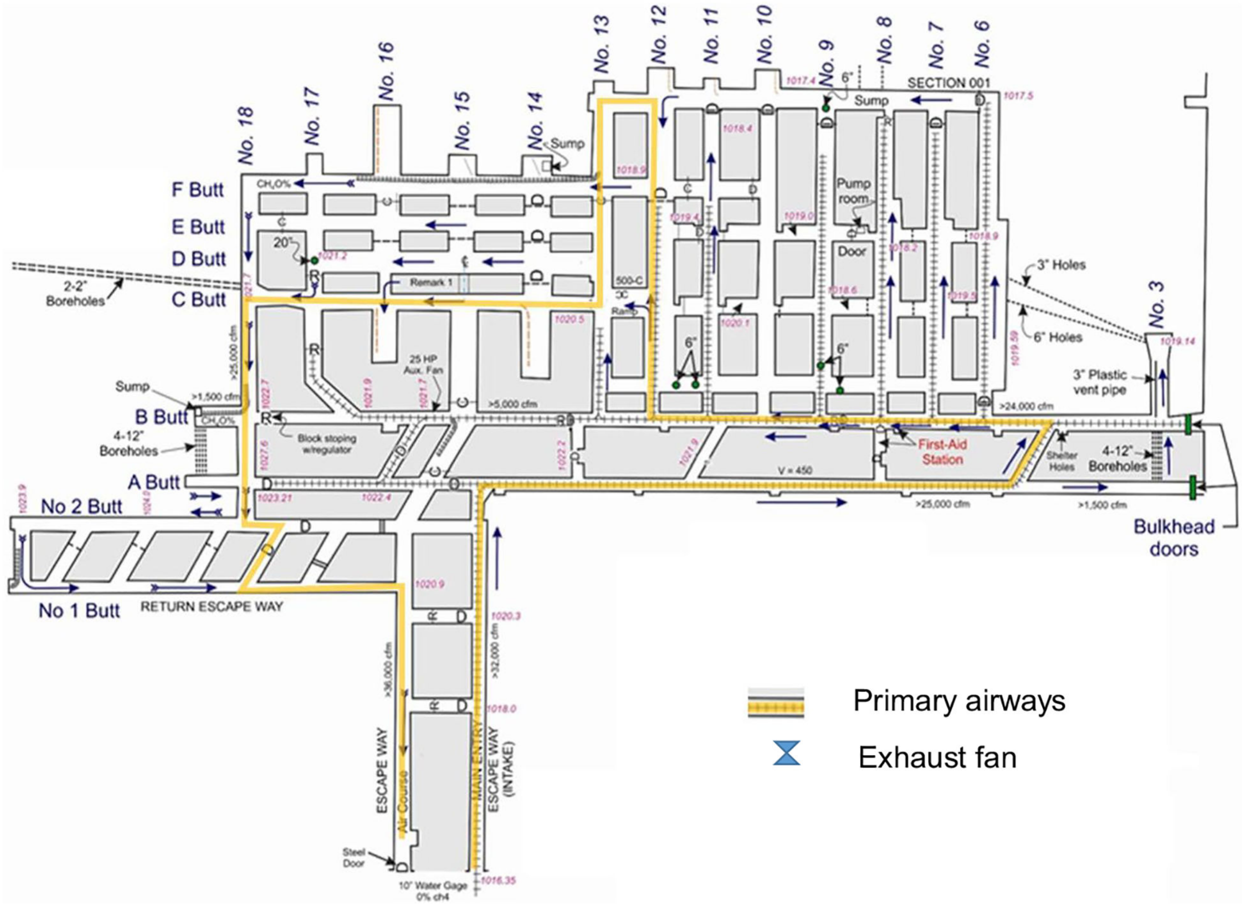

Fig. 1.

Layout of SRCM at Pittsburgh Mining Research Division

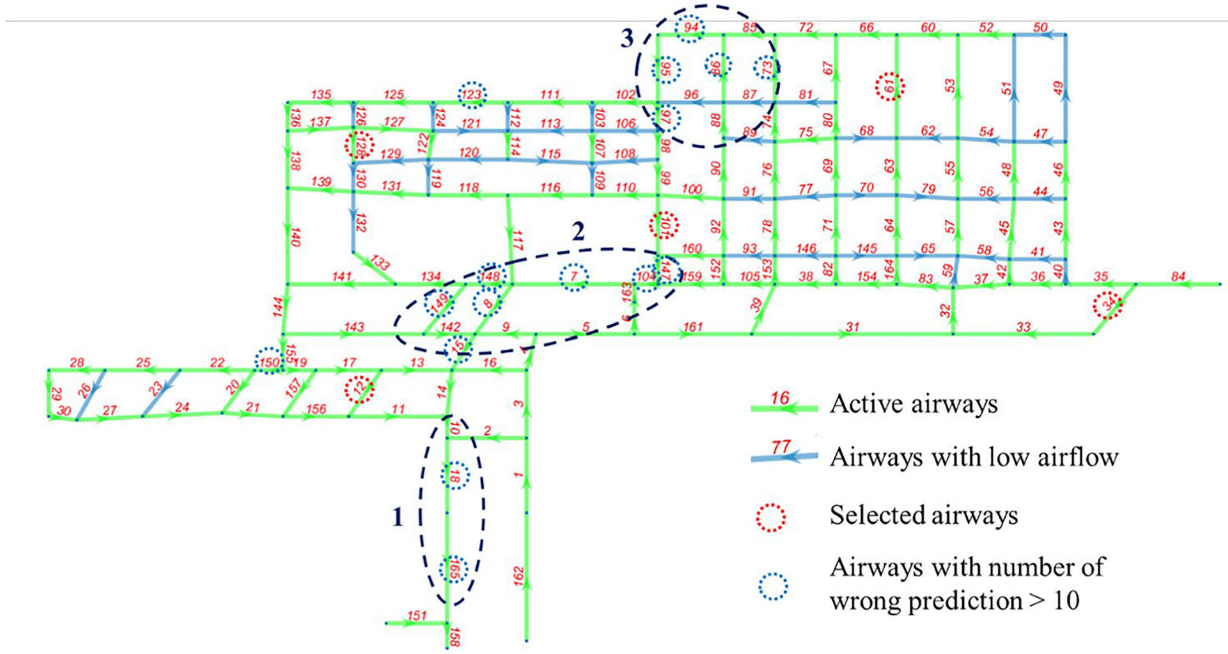

Fig. 2.
Simplified ventilation network for SRCM

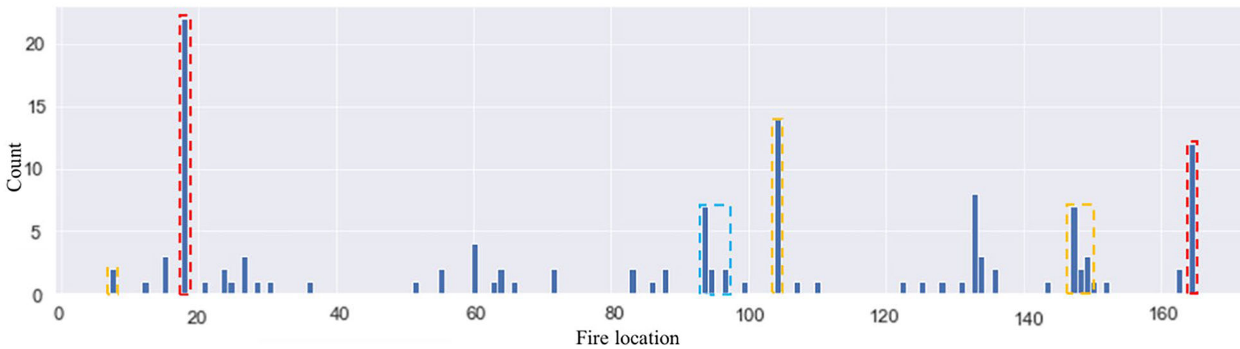

Fig. 3.

Histogram plot of the wrong prediction of the RF classifier trained with all airflow data

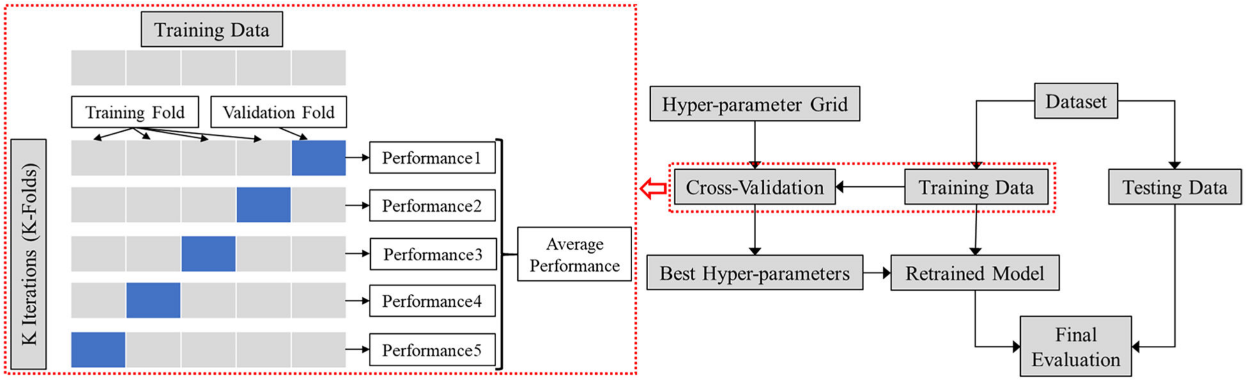

Fig. 4.
Grid search with cross-validation to tune hyper-parameters

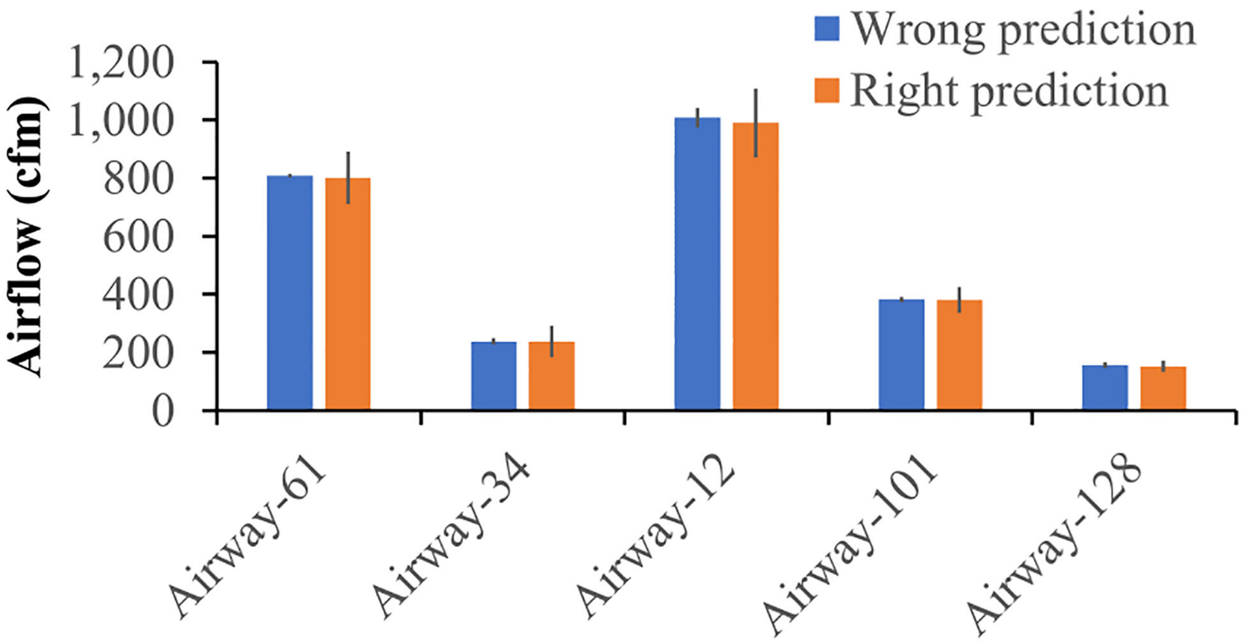

Fig. 5.

Error bar plot of the airflows within the selected airways for the right and wrong predictions

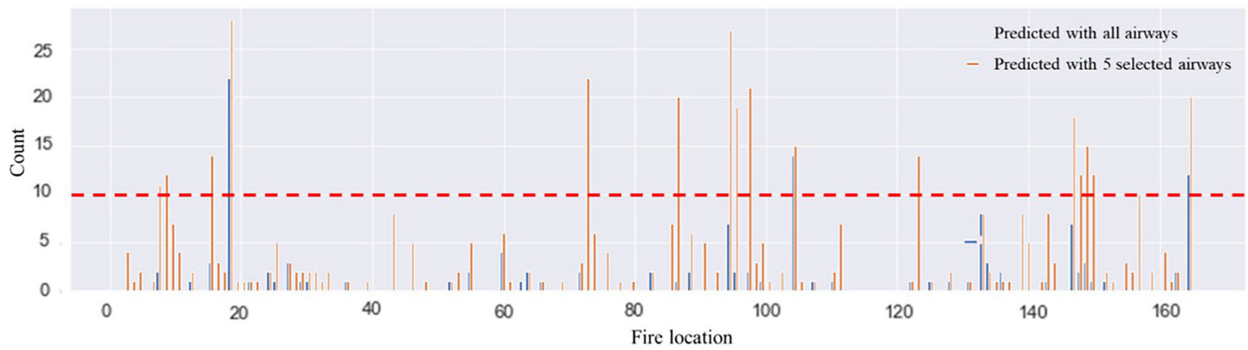

Fig. 6.

Comparison of the wrong fire location prediction between the models trained with the airflow data of all airways and the 5 selected airways

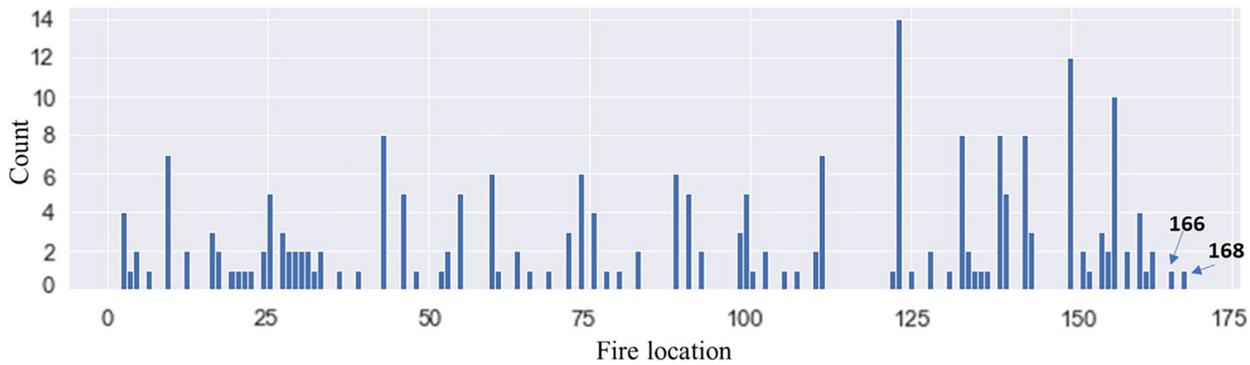

Fig. 7.
Histogram plot of the wrong prediction with selected airways after grouping

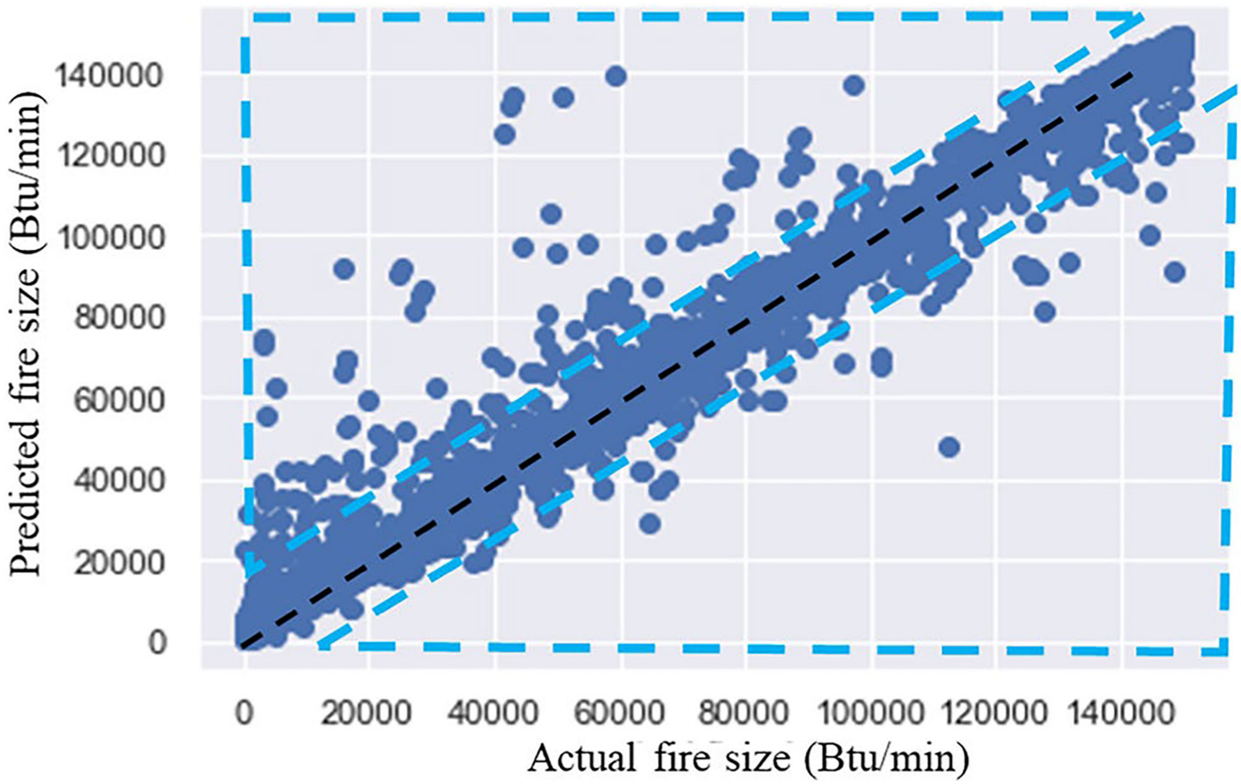

Fig. 8.
Fire size prediction results with the selected airways

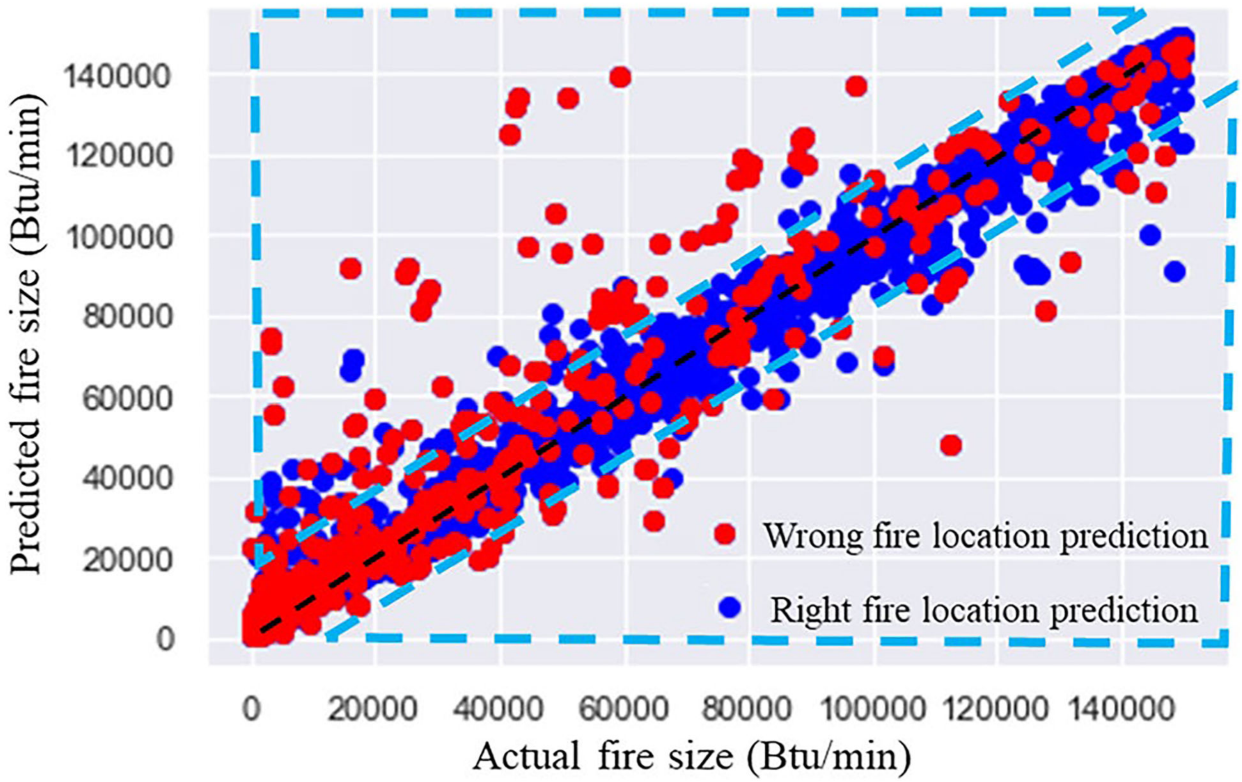

Fig. 9.

Fire size prediction results marked with right and wrong fire location prediction

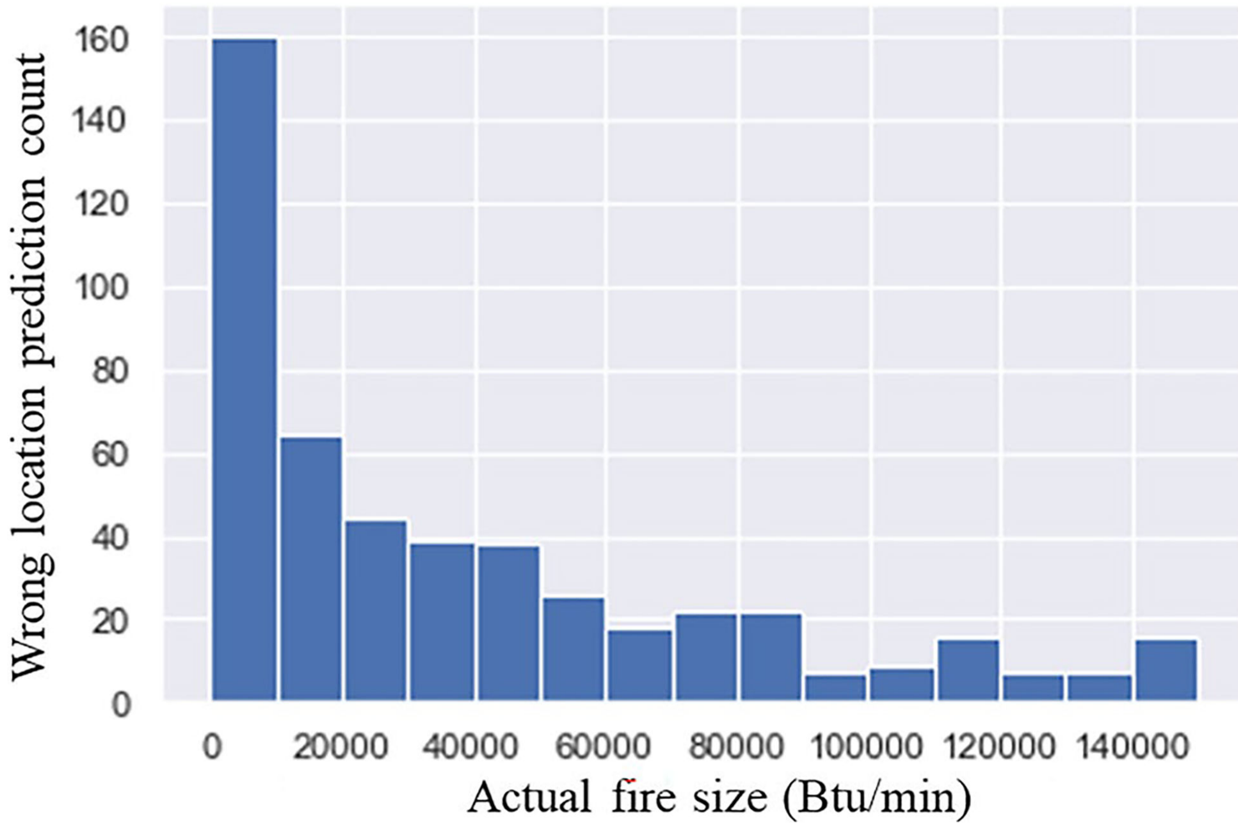

Fig. 10.
Histogram plot of the wrong fire location prediction with different fire sizes (unit: Btu/min)

Table 1

Summary of the wrong predictions for the three groups

Group ID	Airway ID	Total count in testing dataset	Count of wrong predictions	Predicted airway ID	Count
1	**18**	54	28	**165**	25
				10	3
	165	49	20	**18**	12
				10	7
2	**94**	49	27	**95**	18
				97	3
				85	6
	73	49	22	**86**	21
	97	43	21	**94**	7
				95	9
				85	5
	86	53	20	**73**	19
	95	47	19	**94**	13
				97	6
3	**147**	53	18	**104**	14
				7	3
	104	45	15	**147**	12
				7	3
	149	48	15	**148**	8
				8	5
				147	1
				104	1
	15	50	14	**142**	14
	8	45	12	**149**	9
				148	3
	148	43	12	**149**	8
				104	1
				7	2
				8	1
	7	45	11	**104**	8
				147	3