We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,600
Open access books available

138,000
International authors and editors

175M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Sources Identification of Water Inrush in Coal Mines Using Technique of Multiple Non-Linear Machine Learning Modelling

Yao Shan

Abstract

Water inrush is a major threat to the working safety for coal mines in the Northern China coal district. The inrush pattern, threaten level, and also the geochemical characteristics varies according to the different of water sources. Therefore, identifying the water source correctly is an important task to predict and control the water inrush accidents. In this chapter, the algorithms and attempts to identify the water inrush sources, especially in the Northern China coal mine district, are reviewed. The geochemical and machine learning algorithms are two main methods to identify the water inrush sources. Four main steps need to apply, namely data processing, feature selection, model training, and evaluation, in the process of machine learning (ML) modelling. According to a calculation instance, most of the major ions, and some trace elements, such as Ti, Sr, and Zn, were identified to be important in light of geochemical analysis and machine learning modelling. The ML algorithms, such as random forest (RF), support vector machine (SVM), Logistica regression (LR) perform well in the source identification of coal mine water inrush.

Keywords: water inrush, source identification, coal mines, non-linear machine learning, groundwater

1. Introduction

Water inrush is one of severe hazards to coal mines in China. According to statistical material, more than 25 billion tons of coal resources are at the risk of water inrushes in China. From 2000 to 2015, 1162 water inrush accidents were reported, causing 4676 deaths. The number of accidents and deaths took 3.3% and 7.8% of all accidents in coal mines. In spite of the low proportion, major accidents often took place, leading to severe property and live loss.

Northern China district is an important coal base area, reserves of which takes nearly 40% of all country. Therefore, the prevention of water inrush accident is a key issue to the mining safety. The main threats of water inrush to the working face can be grouped into mainly four types, namely surface water, coal roof aquifer water, coal floor aquifer water, and goaf water. The coal roof water is usually relative to coal seam sandstone aquifers, sometime associate with quaternary aquifers. Goaf water formed when the working face closed and ground water filling up this

space. The coal floor water is usually relative to limestone aquifers in the Ordovician system and the Taiyuan Formation in the Carboniferous system.

The different types of water inrush threats show various foreshadow, bursting behaviour, and hazard rating, and corresponding treating technology is essential. Therefore, the technique to predict and evaluate the accident potential, forecast the accident occurrence, and identify water inrush sources, is a key step to prevent the accidents or disasters, and protect the working safety and human health and lives.

In this chapter, the main techniques that used to identify water inrush sources and its application, mainly focusing on the Northern China district, are illustrated.

2. Methods and it's applications for source identification

A basic strategy to identify source of water inrush is based on the geochemical characteristics. Some researchers have compared concentrations of major ions, including K^+, Na^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, CO_3^{2-}, HCO_3^-, and also total dissolved solid, between different aquifers to determine water sources.

In different aquifers, the water composition is a response of its original characteristics and water-rock interaction process. In the Northern China, two main groups of aquifers are coal bearing strata aquifers and limestone aquifers. The geochemists used to find key ions in water, sometimes using geochemical figures, to determine the water sources.

While the geochemical strategy is based on some unique ions and parameters in a lower dimension, another strategy, namely the machine learning (ML) algorithms, is based on multivariate analysis, including some specific methods, and provide more quantitate and reliable results.

2.1 Geochemical methods

The geochemical method is a popular technique in the water inrush identification, for mainly two reasons. First, some coal mines, especially for the large companies, have their own laboratory to test water geochemistry. Therefore, it is easy to obtain data. Second, the experienced technicians are familiar with the water geochemical data, especially for the major ions and important parameters. Researchers usually begin their study from the normal water geochemistry, to investigate water characteristics in every aquifer, set up identification model to distinguish water type from others, and find out the water-rock interaction mechanism for the water composition.

An easy-to-handle method to identify water source is to analyse the major ion characteristics. Cheng et al. [1] analysed water geochemistry in quaternary aquifers, magmatic aquifers, limestone aquifers in the Huaibei coal mine district, Anhui province. The data was grouped into different chemical types, which can be used as database for the water source identification. Chen and Gui [2] discussed water geochemistry in Wanbei coal mine district in Anhui province. Zhang and Cao [3] analysed ground water in Hancheng coal mine district in Shannxi province, founding that the potential water burst point was related with the limestone aquifer. Dai et al. [4] discussed water characteristics in Xiangshan coal mine in Shannxi province. The data was grouped using SPSS to set up a database for further coal mine monitoring and forecasting. The author's group have collected and analysed more than 30 water samples in the Lu'an coal mine district in Shanxi province, the pattern of water flow underground and water characteristics in every ground aquifer were summarised, some important ions, trace elements, and parameters, were identified and used to distinguish water sources from others.

A geochemical chart, the piper diagram, is usually used to analyse and group water samples into different groups by drawing the data as points in two triangle and a diamond figure. Zhang and Cao [3], Dai et al. [4] have applied this technique to identify the water sources. Author's group have collected samples in 2019, **Table 1** shows part of the data, and **Figure 1** shows the water geochemistry in a piper diagram.

As **Figure 1** showing, water in coal bearing seam shows similar characteristics, Na^+ and K^+ take more than 80% and up to more than 95% of all the cations. TDS of the most water sample were less than 1000 mg. The limestone water shows a spanning pattern. TDS of the limestone aquifer water also showed a much wider range, from less than 500 mg to higher than 3000 mg. In the limestone aquifer, water volume is larger, and water-rock interaction is stronger than that in the coal bearing seam, which maybe the reason to the water characteristics in the limestone aquifers.

The source identification using basic geochemical technique is a qualitative, or semi-quantitative method, which may mainly depend on researchers' experiences. If distinguished differences between aquifers are observed in low dimensions, the basic geochemical technique is useful and easy to use. However, while the differences reveal in a higher dimension, i.e. the difference of ions' composition, this method system may lead to a confusing result.

Not only the major ions, but also trace element concentrations and isotope values can be used to distinguish one source from others. Some researchers have used the trace elements to distinguish water samples from others or set up discriminant models. Feng and Han [5] analysed concentration and occurrence of trace elements and modelled its formation using PHREEQC. Chen et al. [6] collected 24 samples from the quaternary aquifers, coal seam sandstone aquifers, and limestone aquifers in Wanbei coal mine district in Anhui province and tested 24 types of trace elements, including Be, B, Sc, V, Cr, etc. The samples and trace elements were clustered. Then eight trace elements, including Be, Zn, Ga, Sr, U, Zr, Cs, Ba, were found to be key parameters to set up discriminant model. The key trace elements were used to train Bayes discriminant analytical model with a good performance.

Isotopes are also used in the water inrush in the coal mines. The most popular isotopes are δD and $\delta^{34}S$ of water. In recent years, the studies are applied in Wanbei coal mine district [7–9], and Fushun coal mine district [10], etc.

No.	$K^+ + Na^+$	Ca^{2+}	Mg^{2+}	Fe^{3+}	Fe^{2+}	NH_4^+	Total
1	363.95	0.84	0.23	0.03	0.09	0.00	365.14
2	17.07	138.56	38.96	0	0	0	194.59
3	12.86	61.84	13.63	0	0	0	88.33
4	41.58	69.88	32.14	0	0	0.07	143.67
5	12.01	65.87	32.14	0	0	0.2	110.22
6	240.88	0.08	0.02	0.15	0	0.44	241.57
7	315.05	12.44	4.14	0	0	1.51	333.14
8	88.16	356.63	38.96	0.11	0	0.38	484.24
9	351.19	1.07	0.54	0.07	0	0.47	353.34
10	334.4	0.21	0.21	0	0	0.41	335.23

Table 1.
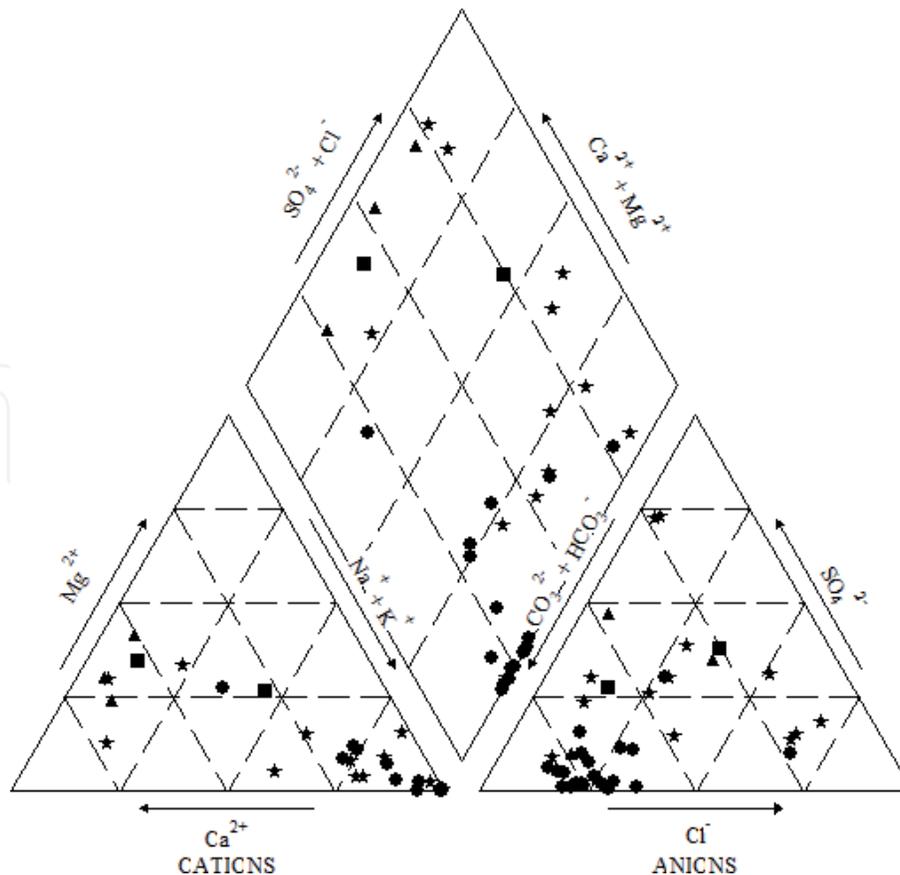*Major ion data in ground aquifer (mg/L).*

Figure 1.
Piper drawing of the ground water (In the figure, the squares stand for surface water, the triangles stand for quaternary water, the circles stand for coal bearing seam water, and the stars stand for limestone aquifer water).

In the author's research in Lu'an coal mine district in Shanxi province, the major ions and trace element were treated together, then SO_4^{2-}, Ti, Sr, Mg, K + Na, Zn, and Cl^- were chosen to be typical ions or elements to train models.

Furthermore, the water form in a scale of whole water unit, therefore the analysis should be carried out in a scale of whole water unit, but not a single point. In the Northern China area, several ground units can be divided, the water-rock interactions among which show similar pattern in different coal mine district. Therefore, the analysis of coal mine district scale and comparison between different coal mine district is an important task to summarise the common mechanism of the water-rock interaction and distinction models.

2.2 Machine learning methods

The geochemical method is effective only if the water samples can be grouped and divided very clearly by one or very few parameters. In most scenario, the ion-distinguishing method is confusing and lack of accuracy. The difference of water samples is embedded in a high dimension, i.e. the combination of major ions, trace elements and other parameters. It is hard to find the dividing mode just by observation or simple drawing. Benefiting from the developing of data science and technology, the environmental and geological issues, including the ground water can be described, and divided by ML methods.

The ML algorithm can be simply divided into supervised, unsupervised, and semi-supervised, depending on how the target variables are labelled. For some environmental and geological problems, the target variables cannot be labelled, then the unsupervised ML algorithm, such as principal components analysis (PCA)

are applied. For example, Shan et al. [11] applied the PCA method to analyse the occurrence and leaching mechanism in coal and host rock, Pumure et al. [12] found out successfully of the occurrence of As and Se in coal host rock. Self-Organising Maps (SOM) is a kind of unsupervised artificial neural network (ANN) used in a large data amount scenario [13–14]. The PCA algorithm is only used for the water inrush if the target variables cannot be labelled [15–17].

While the researchers carrying out their studies, discriminant models should be trained. The training data is obtained from the samples collected from every aquifer. In this step, the data is usually marked clearly. Therefore, the target variables can be obtained for most research cases, and the supervised ML algorithm can be used, which shows high precise and accuracy than the unsupervised ML algorithm. There are several algorithms are suitable for the model training, such as artificial neural network (ANN), support vector machine (SVM), discriminant analysis (DA), decision tree (DT), random forest (RF), boosting, and regression, etc.

In the Northern China, supervised ML algorithm has been used in several coal mine districts. **Table 2** shows part of the research cases in the Northern China area in recent years. It can be concluded from the table that DT criterions are most implemented, some other methods, such as SVM, and ANN, are also used.

2.2.1 Supervised machine learning algorithm

Up to present, the DA is a most popular method analysis to identify sources of water inrush in the Northern China district. Two criterions are usually used, namely Fisher criterion and Bayes criterion. In the framework of Fisher-criterion based DA algorithm, high dimensional data is projected to a one-dimension space, then a discriminant criterion is obtained to achieve the maximum variance between two groups and the minimum in-group variance. Because this method is used to handle a two-group problem, many rounds of calculation are needed for a multiple-group problem. The Bayes-criterion base DA method calculate the posterior probabilities of the sample in each group, then the sample can be classified into the group with the highest posterior probability. Comparing with the Fisher criterion, the Bayes criterion is more frequently used.

The DA is a kind of linear algorithm. Along with the development of ML technology, non-linear modelling is widely used in researches, including the geological, environmental, and engineering area. In order to deal with problems of surface water and ground water, the SVM method is applied to predict water quality and water level [23, 24], ANN and DT are used to predict the $[NO_3^-]$ of ground water [25], set up the water quality monitoring system [26]. Boosting tree is also used to classify distributed water and ground water.

However, the non-linear ML method is relatively less applied to deal with the water inrush problems in coal mines, though higher accuracy maybe achieved

Coal mine districts	Years	Algorithm	Sources
Anhui province	2014	Fishes discriminant	[18]
Henan Province	2015	Feature ions- discriminant	[19]
Shanxi Province	2015	Support vector machine	[20]
Anhui Province	2016	Bayes Discriminant	[6, 7, 21]
Henan province	2019	PCA – grey relationship analysis	[22]

Table 2.
Some implementation of water inrush source identification in the northern China in recent years.

compared to the linear algorithm. According to literature research, the ANN [27] and SVM [20] have been implement in this area. The ANN is a very popular technique in many areas, including figure and voice identification, driverless driving, etc. However, the ANN usually needs large amount of data to train model to control its over-fitting problem. On the other hand, the data of the environmental and geological area, including the water inrush analysis are usually structured data, and limited to a small data quantity. As a result, the problem of over-fitting problem is hard to control, which means low accuracy of prediction is prospected while using the ANN model to check using the testing data, though a high accuracy may be obtained while testing the model using the training data. The algorithm of SVM perform better to control the over-fitting problem. Other than SVM, the DT, DT, boosting tree, Bayes network (BN) also have good prospect, in consideration of the characteristics of the coal mine ground water, i.e. structured small data quantity.

2.2.2 Data selection and feature engineering

The tested data of ground water is material of model training. However, the data preparing is essential to ensure or enhance the model quality. The data preparing work mainly includes data selection and feature engineering.

In a wide sense, the data selection includes data cleaning, which means treatment of unit and missing data. Then the data should be selected to determine those used in the model training step. The data selection is applied in two stages, before model training and after model training. Before the model training, the suitable data for the model training means to make sure all the data is labelled correct. Uncorrected marked data leads to wrong model definitely, regardless of the quality of models. After the model training, the training data should be checked again. The data have to be checked very carefully to find wrong classified data. While it is determined to wrong pre-labelled, then the data should be deleted, and new model need to be trained.

The other important work before the model training is feature engineering. The basic mechanism to process feature engineering is to achieve a best performance of the model. **Figure 2** shows the idea of feature engineering. Number of features, or parameters, means the model complexity. More features in the model lead to a higher complexity of the model. As **Figure 2** showing, the prediction performance is related to model complexity. A very simple model lead to very bad model
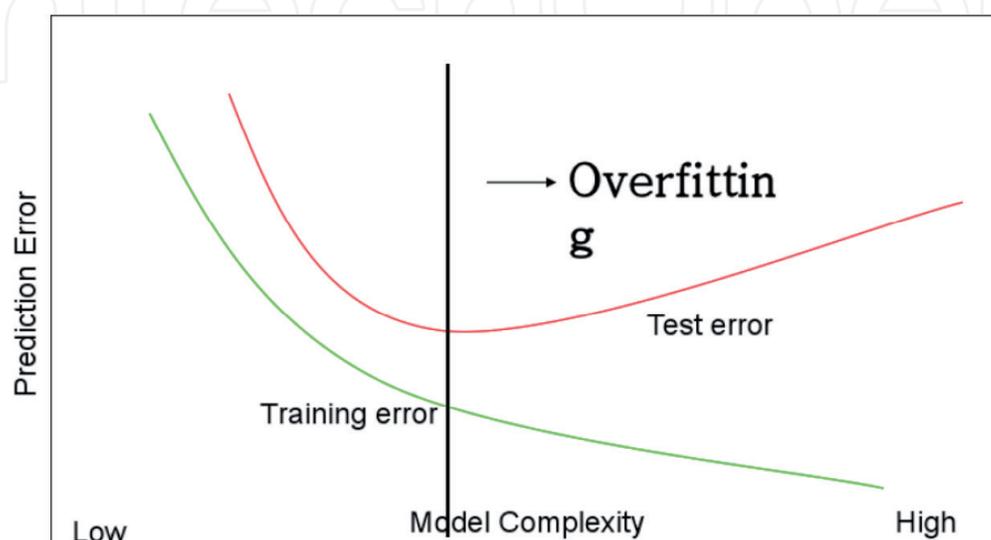

Figure 2.
Correlation of the model complexity and prediction error.

performance, that's why the non-linear models are used. Along with the increasing of model complexity, the prediction error of the training samples becomes lower steadily. On the other hand, the prediction error of the testing samples gets lower at first, then higher again. That suggested over-fitting problem in the ML model. Therefore, the feature has to be processed if a good performing model is acquired.

The feature engineering includes feature fusion and feature selection. A common feature fusion method is PCA. The PCA can reduce dimension of data, then the features in a lower space could stand for most data information. As combination of the original features, the new feature cannot reflect the data characteristics directly. While the researchers want to analyse the importance of the parameters in the original data, the feature selection technique should be used.

Popular feature methods include RF, and Lasso regression, etc. The RF based feature selection undergoes the following steps.

1. The data set X contain N samples, draw samples randomly from the data set X using the bootstrap resampling method. The resampling is carried out k times, to construct k regression tree. In this process, the probability of no drawing of each sample is $p = (1–1/N)^N$. The p tends to 0.37 while the N increasing to infinity. That means that about 37% of the samples in the data set X are not drawn, these data are not used in the DT training, calling out-of-bag (OOB) data. These OOB data is used to test the regression trees.

2. For k bootstrap samples, k unpruned regression trees are created respectively. In the training process, for each node, m attributes are randomly selected from the total M attributes as internal nodes. Then, an optimal attribute is selected from m attributes as a split variable to make the branches grow, according to the minimum Gini index principle.

3. The k decision tress comprises a random forest, the model quality could be evaluate using two indices: large mean square error of OOB (MSE_{OOB}) and low coefficients of determination (R^2_{RF}).

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

$$R^2_{RF} = 1 - \frac{MSE_{OOB}}{\hat{\sigma}^2_y} \tag{2}$$

Where n is the total number of the samples, $\hat{y}_i$ is the predicted output obtained by the generated RFR regression model, y_i is the observed output value, and the $\hat{\sigma}^2_y$ is the predicted variance of the OOB output.

4. The RF regression model provides two methods to determine the importance degree of each variable index: mean decrease in Gini index and mean decrease in accuracy. In a regression model, the mean decrease in Gini is usually used, and the mean decrease in accuracy is more applied for the classification problem. The water inrush source identification is a kind of classification problem, therefore the mean decrease in accuracy is selected.

While carrying out the inrush source identification, the attributes could be used in the model includes major ions, trace elements, important parameters, and isotopes, etc. In which, the data of major ions and important parameters are easier

to obtain. On the other hand, adding of trace elements and isotopes into the models may enhance the model performance, for these parameters carries a lot of information of the water samples. In consideration of easy using, only major ions and important parameters is used, while considering for the model accuracy, more parameters could be added. Therefore, it is a balance need to consider while building models.

In our previous study in the Lu'an coal mine district in Shanxi Province, all the prescribed parameters have been tested. In the first step, the feature selection was applied on the major ions and important parameters. **Figure 3** shows the calculation result while using the algorithm of RF. As **Figure 3** showing, key attributes to the mode are SO_4^{2-}, C (stands for CO_3^{2-} and HCO_3^-), $K^+ + Na^+$, Mg^{2+}, Cl^-, and Ca^{2+}, in a
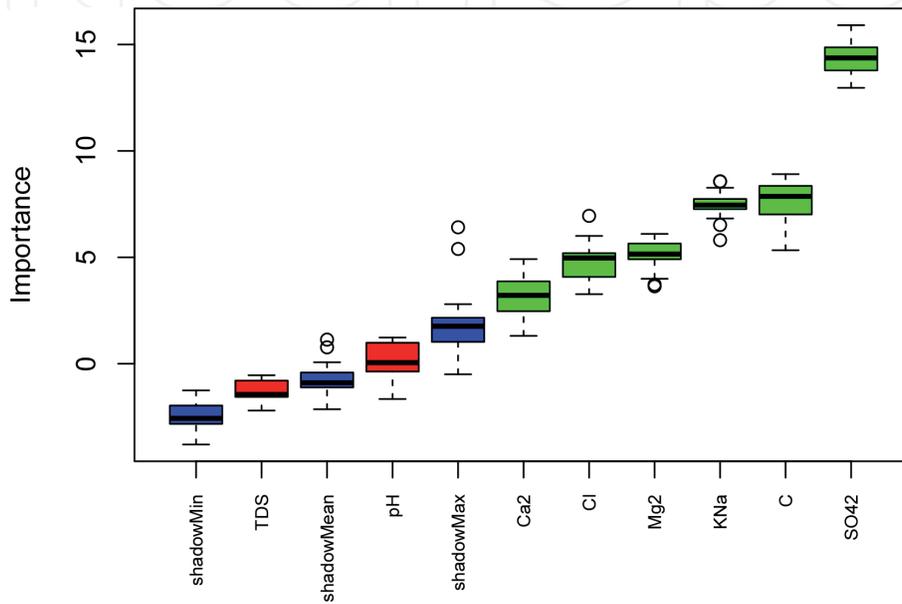

Figure 3.
Feature selection result of major ions and important paraments using RF algorithm.
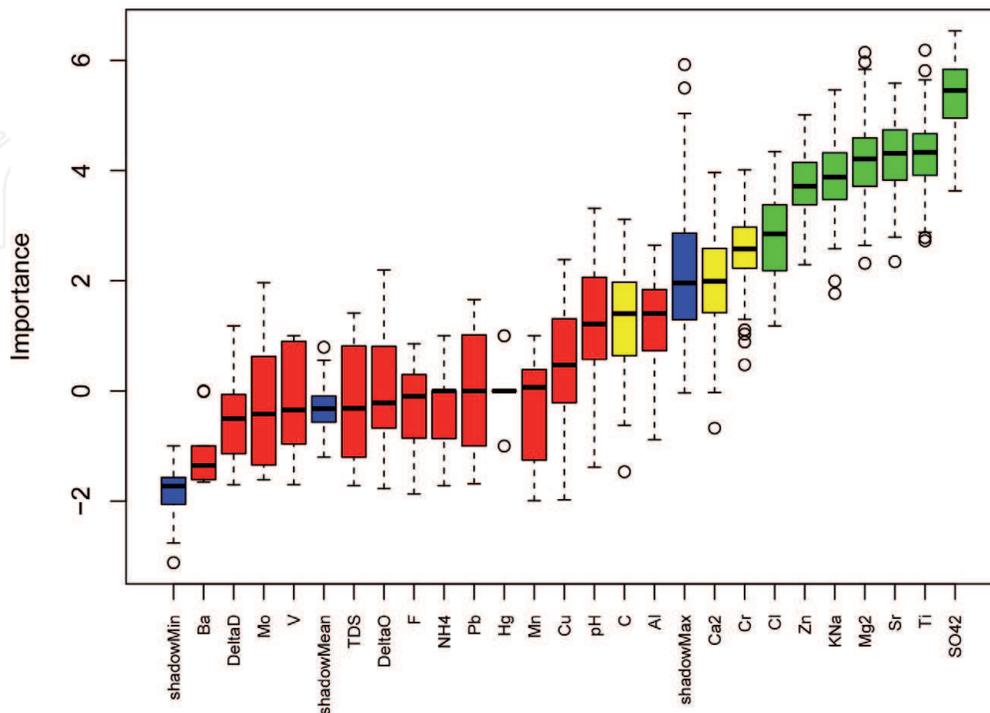

Figure 4.
Feature selection result of all data using RF algorithm.

descending order. It can be concluded that all the major parameters contribute to the explanation of water source identification.

For the trace elements and isotopes may carry information of the water samples, all the parameters were calculated following the same process of feature selection illustrated in the previous section, which is shown in **Figure 4**. As the figure showing, the key attributes for the model turned out to be SO_4^{2-}, Ti, Sr, Mg^{2+}, K^++Na^+, Zn, and Cl^-. Comparing with the first feature select, the major ions still play important role in the water inrush identification, that's because of the significant water-rock interaction in the ground aquifer. However, some trace elements are also important in the distinguishing water source from others, including Ti, Sr, and Zn. These trace elements show low concentrations in coal bearing seam water, and higher concentrations in limestone aquifer water.

We have also applied the Lasso regression to determine important attributes for the model training. The RF perform better, so only the RF result is illustrated in this chapter.

2.2.3 Model selection

In the previous description, the data and attributes have been selected to train models, and the model frameworks need to select and evaluate. According to the literature review, the non-linear ML modelling is seldom applied to solve the problem of water inrush in coal mines in the Northern China area. However, by consideration the mechanism of ML algorithm, and relative studies in the environmental and geological area, some non-linear ML method were applied, evaluated and compared.

Li et al. [20] applied the SVM algorithm to determine the source of water inrush. The SVM algorithm project data from a lower dimensional space to a higher dimensional space by applying the kernel functions. Then the data that cannot divided in a lower dimensional space can be separated. Several types of kernel can be evaluated, such as radial basis and poly-nominal functions, etc.

Li et al. [20] compared the performance of geochemical and machine leaning methods on the water inrush issues based on the Lu'an coal mine district. It was found that, the accuracy by using the SVM method has achieved 100%, while only 42% and 48% have obtained while using one (SO_4^{2-}) or two (SO_4^{2-} and Ca^{2+}) parameters as the key attributes, respectively. In this research, the radial basis and poly-nominal function were not compared. However, a mixing kernel function was proposed.

The ANN has also used to identify the source of coal mine water inrush [27]. For example, in the research of Li et al. [27], an improved genetic algorithm combining PCA and back propagation ANN were applied. The research had achieved a simple, reasonable, and effective distinguishing result.

In a previous study carried out by the authors' group, 42 samples were collected from the Lu'an coal mine district, Shanxi province, which belong to six water types, namely sandstone aquifer water, Ordovician limestone water, Taiyuan formation limestone water, goaf water, spring water, and fault-oxidising zone water, respectively. The parameters tested included K^+, Ca^{2+}, Mg^{2+}, SO_4^{2-}, pH, Fe, NO_2^-, I, Cl^- and temperature. The data was normalised before the model training. Some non-linear ML was applied. The research has focused on the following issues: first, the key attributes should be selected by the ML method, which has been described in the previous section, second, the data need to be cleaned and selected, third, a wide scope of parameters should be considered, including major ions, important traditional parameters, trace elements, and isotopes, and the fourth, the ML framework need to be evaluated and compared.

We have applied several algorithms; the identification result is shown in **Table 3**. Considering the easy using of the model, only major ions, and key parameters are used

Attribute scope	ML algorithm	Corrected identification
Major ions and parameters	RR	28/31
Major ions and parameters	RF	29/31
Major ions and parameters	SVM (radial basis)	29/31
Major ions and parameters	SVM (poly-nominal)	30/31
Major ions and parameters	LR	29/31
Major ions and parameters	DA	29/31
All data	RR	23/25
All data	RF	25/25
All data	SVM (radial basis)	24/25
All data	SVM (poly-nominal)	25/25
All data	LR	25/25
All data	DA	25/25

Table 3.
Identification result of water resource using the non-linear ML algorithm.

in the first-round model training. In a traditional process, the data should be divided into two groups with a ratio of 7:3. The first part is used to train models, and the second part is used to evaluate the model performance. Because the amount of data is limit, the models' test was carried out by using the method of re-calculating the sample data using models. Other than the ridge regression (RR), other algorithm showed better performance. The DA was based on the Bayes criterion, showing an acceptable result. Comparing the two SVM method, poly-nominal kernel got a better result than that of the radial kernel SVM.

The wrong identified data is important to the modelling. Four main reason may lead to the result: First, the water sample was wrong labelled; second, the water is a mix of multi water sources; third, the attributes selected is not suitable; and fourth, the model need to be improved or more suitable model framework is needed.

For the first reason, the data should be analysed very carefully. If not representative, the samples should be deleted. For the second reason, the water mix should be calculated using other method, such as the nonlinear programming. In order to improve the model performance, the scope of the attributes enlarged to include trace elements and isotopes. For the trace elements of some samples was not tested, the overall sample data was less than the first-round training. According to the results, better performance was observed when using the trace elements. It suggested that some trace elements can be used as the key attributes in the water source identification. And the wrong-identified samples were probably because of poorer performance of the first-round modelling, rather than the wrong-labelling of the samples. Comparing the ML framework, the RF, SVM, LR, and the Bayes-based discriminant analysis all showed high accuracy. The accuracy of ridge regression (RR) is less than others. Comparing the two kernels used for the SVM, the poly nominal based SVM showed a better performance.

3. Conclusions

In this chapter, the main methods to identify water inrush sources in the coal mines, especially in the Northern China area, were reviewed.

The basic idea of the geochemical method is to find characteristic ions, which shows distinct behaviour in one group of water samples from others, by geochemical analysis and drawing.

In the process of non-linear machine learning, four main steps need to apply, they are: data processing, feature selection, model training, and evaluation. According to the studies of literature review and the author's previous study, some key attributes/ parameters were selected using the ML algorithm, then the ML algorithm was compared. It was found that, most of the major ions, and some trace elements, such as Ti, Sr, and Zn, were important in the model training. Then ML algorithms, RF, SVM, LR perform well in the source identification of coal mine water inrush.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author details

Yao Shan
School of Emergency Technology and Management, North China Institute of Science and Technology, Yanjiao, 101601, China

*Address all correspondence to: 9106350@qq.com

IntechOpen

References

[1] Cheng C, Ge X, Wang, D. Hydrogeochemical characteristics of Baishan coal mining area and the application to the source discrimination of water invasion in the mine. Journal of Huaian Mining Institute. 1994; 14(1): 25-32. (In Chinese)

[2] Chen L, Gui H, Hu Y, Xu G, Yuan W. Study of hadrochemical characteristics in Karst water of coal seam floor in the Mining District of North Anhui. Coal Geology & Exploration. 2003; 31(2): 27-31. (In Chinese)

[3] Zhang L, Cao H. Distinguishing the sources of water inrush in Sangshuping coal mine by hadrochemical characteristics. Coal Geology & Exploration. 2013; 41(4): 42-45. (In Chinese)

[4] Dai G, Xue X, Niu C. Discrimination of water inrush source in xiangshan coal mine based on chemical characteristics of groundwater. Journal of Xi'An University of Science and Technology. 2017; 37(2): 213-218. (In Chinese)

[5] Feng Q, Han B. Occurrence of trace elements in reservoir, oilfield water, and scaling of Wumishan Formation, Renqiu, China. Journal of China University of Mining & Technology. 2001; 30(2): 122-125. (In Chinese)

[6] Chen L, Gui H, Yin X, Qian J, Pan L. The standard type trace elements and the discriminant model of water bursting source in the Linhuan coal district. Hydrogeology & Engineering Geology. 2010; 37(3): 18-22. (In Chinese)

[7] Gui H, Chen S. sotopic geochemical characteristics of groundwater and its geological significance in Sunan mining area. Earth Science Frontiers. 2016; 23(3): 133-139. (In Chinese)

[8] Chen L, Gui H, Yin X. Composing characteristic of hydrogen and oxygen stable isotopes and tracing of hydrological cycle. Journal of China Coal Society. 2008; 33(10): 1107-1111. (In Chinese)

[9] Chen L, Gui H, Yin X, Qian J. Characteristic 0of 18 O and D and Current Field in Deep Groundwater. Journal of China University of Mining & Technology. 2008; 37(6): 854-859. (In Chinese)

[10] Hu W, Ma Z, Cao H, Liu F, Li T, Dou H. Application of isotope and hydrogeochemical methods in distinguishing mine bursting water source. Journal of Earth Sciences and Environment. 2010; 32(3): 268-271. (In Chinese)

[11] Shan Y, Wang W, Qin Y, Gao L. Multivariate analysis of trace elements leaching from coal and host rock. Groundw Sustain Dev. 2019; 8: 402-12.

[12] Pumure I, Renton JJ, Smart RB. The interstitial location of selenium and arsenic in rocks associated with coal mining using ultrasound extractions and principal component analysis (PCA). J Hazard Mater. 2011;198:151-8.

[13] Lacassie JP, Roser B, Ruiz Del Solar J, et al. Discovering geochemical patterns using self- organizing neural networks: a new perspective for sedimentary provenance analysis. Sedimentary Geology. 2004; 165(1-2): 175-191.

[14] Nakagawa K, Yu Z-Q, Berndtsson R, et al. Temporal characteristics of groundwater chemistry affected by the 2016 Kumamoto earthquake using self-organizing maps. Journal of Hydrology. 2020; 582: 1-14.

[15] Valder J F, Long A J, Davis A D, et al. Multivariate statistical approach to estimate mixing proportions for unknown end members. Journal of Hydrology. 2012; 460-461: 65-76.

[16] Mencić A, Folch A, Mas-Pla J. Identifying key parameters to differentiate groundwater flow systems using multifactorial analysis. Journal of Hydrology. 2012; 472-473: 301-313.

[17] Ju Q, Hu Y, Zhang S. Mine water inrush source identification method based on principal component analysis and Bayesian discriminant. Coal Engineering. 2018; 50(12): 90-94. (In Chinese)

[18] Yang X, Wang L, Zhang Z. The Application of Fisher Approach to Discrimination of Mine Water-inrush Source. Science & Technology Information. 2014; 11: 97-98. (In Chinese)

[19] Zhu Q, Li X. Water source discrimination of inrush water in mine based on water chemistry characteristics analysis. Journal of North China Institute of Science and Technology. 2015; 12(3): 29-33. (In Chinese)

[20] Li J, Xu Y, et al. Research on method of automatic recognition of water sources based on Support vector machine. Journal of North China Institute of Science and Technology. 2015; 12(2): 25-29. (In Chinese)

[21] Chen M, Wu Y, Gao D, Chang M. Shallow Groundwater Hydrogeochemical Evolution Process and Controlling Factors in Plain Zone of Guanghan City. Journal of Jilin University (Earth Science Edition). 2016; 46(3): 831-843. (In Chinese)

[22] Yang Z, Ding F, Huang P. Research on PCA-GRA Discrimination Model of Water Inrush Source in Mines. Coal Technology. 2019; 38(12): 84-87. (In Chinese)

[23] Singh K P, Basant N, Gupta S. Support vector machines in water quality management. Analytica Chimica Acta. 2011; 703(2): 152-162.

[24] Kisi O, Shiri J, Karimi S, et al. A survey of water level fluctuation predicting in Urmia Lake using support vector machine with firefly algorithm. Applied Mathematics and Computation. 2015; 270: 731-743.

[25] Wagh V, Panaskar D, Muley A, et al. Neural network modelling for nitrate concentration in groundwater of Kadava River basin, Nashik, Maharashtra, India. Groundwater for Sustainable Development. 2018; 7: 436-445.

[26] O'Connor E, Smeaton A F, O'Connor N E, et al. A Neural Network Approach to Smarter Sensor Networks for Water Quality Monitoring. Sensors. 2012; 12(4): 4605-4632.

[27] Li Y, Niu G, Liu H. Application of improved GA-BP neural network on identification of water inrush source in mine. Journal of Safety Science and Technology. 2016; 12(7): 77-81. (In Chinese)