Height prediction of water-flowing fracture zone with a genetic-algorithm support-vector-machine method

Enke Hou[1] · Qiang Wen[1] · Zhenni Ye[1] · Wei Chen[1] · Jiangbo Wei[1]

Abstract Prediction of the height of a water-flowing fracture zone (WFFZ) is the foundation for evaluating water bursting conditions on roof coal. By taking the Binchang mining area as the study area and conducting an in-depth study of the influence of coal seam thickness, burial depth, working face length, and roof category on the height of a WFFZ, we proposed that the proportion of hard rock in different roof ranges should be used to characterise the influence of roof category on WFFZ height. Based on data of WFFZ height and its influence index obtained from field observations, a prediction model is established for WFFZ height using a combination of a genetic algorithm and a support-vector machine. The reliability and superiority of the prediction model were verified by a comparative study and an engineering application. The results show that the main factors affecting WFFZ height in the study area are coal seam thickness, burial depth, working face length, and roof category. Compared with multiple-linear-regression and back-propagation neural-network approaches, the height-prediction model of the WFFZ based on a genetic-algorithm support-vector-machine method has higher training and prediction accuracy and is more suitable for WFFZ prediction in the mining area.

1 Introduction

When mining a coal seam, the roof rock will move, forming a fracture zone (Gao 1996). Fractures in this zone may provide a channel for accumulated water in the overlying aquifer, not only destroying groundwater resources in the aquifer but also causing groundwater to flow into the mine along the water-flowing fracture zone (WFFZ), threatening mine safety and production (Wu et al. 2000, 2016; Li and Li 2012; Garritty 1983; Zhang and Yang 2018; Peng and Zhang 2007). Therefore, coal scientists should pay attention not only to the prediction of

water abundance in roof aquifers but also to the prediction of WFFZ height.

To date, many important results have been obtained through similar-material simulations (Lin et al. 2010; Gao and Wu 2011; Zhao et al. 2011; Sui et al. 2015; Dong et al. 2015; Zhang et al. 2018, 2017), numerical simulation experiments (Liu et al. 2018a; Wu et al. 2014; Liu et al. 2018b; Zhang et al. 2018; Meng et al. 2018; Zhu et al. 2018; Yin et al. 2016; Wu et al. 2015), theoretical analysis (Xu and Sun 2008; Shi et al. 2012; Liu et al. 2018b; Qiao et al. 2017; Yin et al. 2016), physical exploration (Yang et al. 2018), and drilling tests (Luan et al. 2010; Sui et al. 2015; Dong et al. 2015; Zhang and Peng 2005; Yin et al. 2016) for the study of WFFZ height in the roof after coal mining. However, these methods mainly study height prediction of the WFFZ under specific geological and mining conditions, and it is difficult to conduct such studies under complex and variable geological and mining conditions (Wang 2006).

✉ Enke Hou
houek@xust.edu.cn

[1] Xi'an University of Science and Technology, Xi'an 710054, China

In recent years, many scholars have proposed multi-factor-analysis methods for WFFZ height prediction under complex and variable geological conditions. In terms of the main factors that affect WFFZ height, Liu first considered two control factors: the roof category and the coal seam thickness (Liu 1995). Afterwards, some scholars found that the height of the roof WFFZ in coal mining is also controlled by the length of the coal face, the coal seam pitch, and the burial depth of the coal seam (Hu et al. 2012; Liu 2010; Du and Weng 1997). Using multi-factor analysis, early scholars described the relationship between WFFZ height and the main control factors by using linear models, such as multiple linear regression (MLR) and the entropy method (Hu et al. 2012; Wang et al. 2018; Xu and Sui 2013). With the development of system theory and computer technology, a nonlinear method using back-propagation neural-networks (BPNNs) was applied (Chen et al. 2005). Height prediction of WFFZs has been extensively studied using multi-factor analysis. The main factors affecting WFFZ height have been revealed, and various advanced methods have been applied for its prediction. Among these factors, the roof category is divided into three types: weak, medium-hard, and hard. In fact, the influence of roof category on WFFZ height is relatively complicated. Thus, it is necessary to characterise roof classification quantitively to predict WFFZ height accurately, and determining a reasonable index to characterise roof classification is essential. In addition, many nonlinear methods, such as BPNNs, overcome the limitations of linear analysis, but the use of these methods requires large amounts of observational data. Therefore, a new method is urgently needed to establish a prediction model for WFFZ height based on small amounts of observational data.

In this study, we used single-factor analysis to study the influence of coal seam thickness, burial depth, working face length, and roof category on WFFZ height and created an index to characterise the roof category. Using measured values of WFFZ height and its influence indexes, we established a prediction model for this height based on a genetic algorithm (GA) and a support-vector machine (SVM) to obtain better predictions when using small amounts of observational data and to provide the theoretical basis for the prediction and prevention of roof-gushing (inrush) water disasters in coal mining in the future.

2 Descriptions of study area

The Binchang mining area covers an area of about 670 km², located in the middle of the Huanglong Jurassic Coalfield in Shaanxi Province, and is one of the most important coal-industry bases in China (Fig. 1). The landform consists of the Loess gully and hilly areas, and the
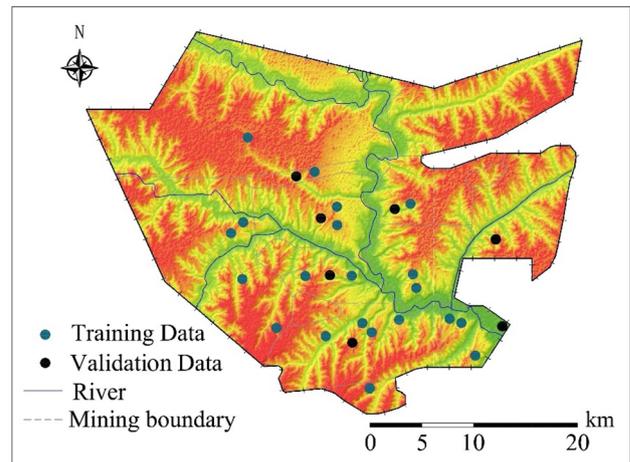

Fig. 1 Location of the study area and distribution of the WFFZ height observation points

terrain is high in the northwest and low in the southeast. The Yan'an Group of Jurassic strata is a coal-bearing stratum with a total of eight coal seams, among which coal seam number 4 is the current main coal seam with an average coal thickness of about 10 m.

Jing River, Hei River, and Hongya River constitute the main river system in the study area. The Jing River flows through the study area from northwest to southeast. The Hei River flows into the study area from the west and merges into the Jing River. The Hongya River flows into the study area from the east and merges into the Jing River. The groundwater in the study area can be divided into loose-pore unconfined aquifers and bedrock pore-fracture confined aquifers. The bedrock-fracture aquifer of the Cretaceous Luohe Group is the most important and is the main source of domestic and industrial water in the study area, as well as an inrush source that affects the safety of mine production.

3 Data preparation

3.1 In situ observation of WFFZ height

In coal mining, the overburden rock collapses under gravity, causing the stress of surrounding rock to be redistributed, resulting in tensile failure above the mined-out area and shear failure above the coal wall, forming a WFFZ. Therefore, the rock-mass fractures in a WFFZ are more developed than those in normal rock mass. When an exploration hole is drilled into a WFFZ, the consumption of flushing fluid in the borehole will increase markedly, and the water level will continue to decrease. Therefore, WFFZ height can be determined from the consumption of flushing fluid and the variation characteristics of the water level in

the borehole. A total of 28 boreholes were arranged in the study area to observe WFFZ height (Fig. 1). The observational results in Table 1 show that WFFZ height in the study area is mainly between 100 and 240 m.

3.2 Factors of WFFZ height

It is necessary to analyse the factors affecting WFFZ height before predicting it. Based on previous research results and practical experience, the height of the WFFZ roof after coal mining is mainly controlled by coal seam thickness, working face length, coal seam burial depth, and roof category (Table 1). The data show that the stratum in the study area is inclined to the north, the pitch changes little, and the influence of pitch on WFFZ height is small. The

influence of other factors on WFFZ height is showed as follows.

3.2.1 Coal seam thickness

After coal mining, the influence range of roof displacement and failure is directly controlled by coal seam thickness. With similar conditions of working face length, representative data are selected from Table 1 to analyse the relationship between WFFZ height and coal seam thickness (Fig. 2a). The results show that the height of the WFFZ roof increases significantly with increasing coal seam thickness; therefore, coal seam thickness is an important factor affecting WFFZ height. In addition, with similar conditions of coal seam thickness, when the length of the

Table 1 Measured height of the water-flowing fracture zone and its main control factors

No. of boreholes	Working face length (m)	Burial depth (m)	Coal seam thickness (m)	Hard-rock proportion in different ranges of roof thickness			Height of water-flowing fracture zone (m)
				0–7 M	7–10 M	10–15 M	
ZK1	180.0	550.00	3.50	0.25	0.93	0.38	130.00
ZK2	180.0	550.00	3.50	0.35	0.80	0.32	134.00
ZK3	240.0	546.96	3.82	0.56	0.64	0.32	180.60
ZK4	240.0	550.00	3.85	0.48	0.41	0.61	145.00
ZK5	93.4	332.00	8.90	0.64	0.47	0.28	149.48
ZK6	93.4	333.50	9.00	0.51	0.29	0.39	124.50
ZK7	116.0	650.00	9.10	0.09	0.00	0.83	121.03
ZK8	103.9	520.00	9.10	0.26	0.00	0.75	134.98
ZK9	116.0	650.00	9.20	0.34	0.25	0.81	138.38
ZK10	93.4	325.00	9.80	0.30	0.27	0.47	120.00
ZK11	116.0	650.00	9.80	0.52	0.57	0.34	175.80
ZK12	93.4	325.00	9.90	0.56	0.45	0.14	149.00
ZK13	116.0	650.00	9.99	0.38	0.23	0.62	165.83
ZK14	130.0	980.00	10.00	0.08	0.12	1.00	130.00
ZK15	180.0	475.95	11.22	0.52	0.72	0.24	211.09
ZK16	180.0	475.95	11.22	0.46	0.61	0.42	192.00
ZK17	200.0	650.00	12.00	0.56	0.36	0.13	238.67
ZK18	180.0	480.00	12.34	0.53	0.34	0.38	187.00
ZK19	180.0	488.24	12.56	0.39	0.21	0.52	179.50
ZK20	175.0	610.0	12.92	0.34	0.21	0.60	189.52
ZK21	175.0	610.0	13.00	0.42	0.21	0.65	225.43
ZK22	116.0	550.0	3.83	0.53	0.59	0.31	121.00
ZK23	116.0	549.0	3.81	0.49	0.43	0.62	110.00
ZK24	180.0	629.0	9.20	0.34	0.25	0.81	152.60
ZK25	180.0	637.0	9.80	0.52	0.57	0.34	183.20
ZK26	180.0	650.0	9.80	0.47	0.58	0.41	176.30
ZK27	120.0	490.0	12.31	0.51	0.42	0.37	172.30
ZK28	130.0	500.0	12.50	0.42	0.19	0.53	164.20

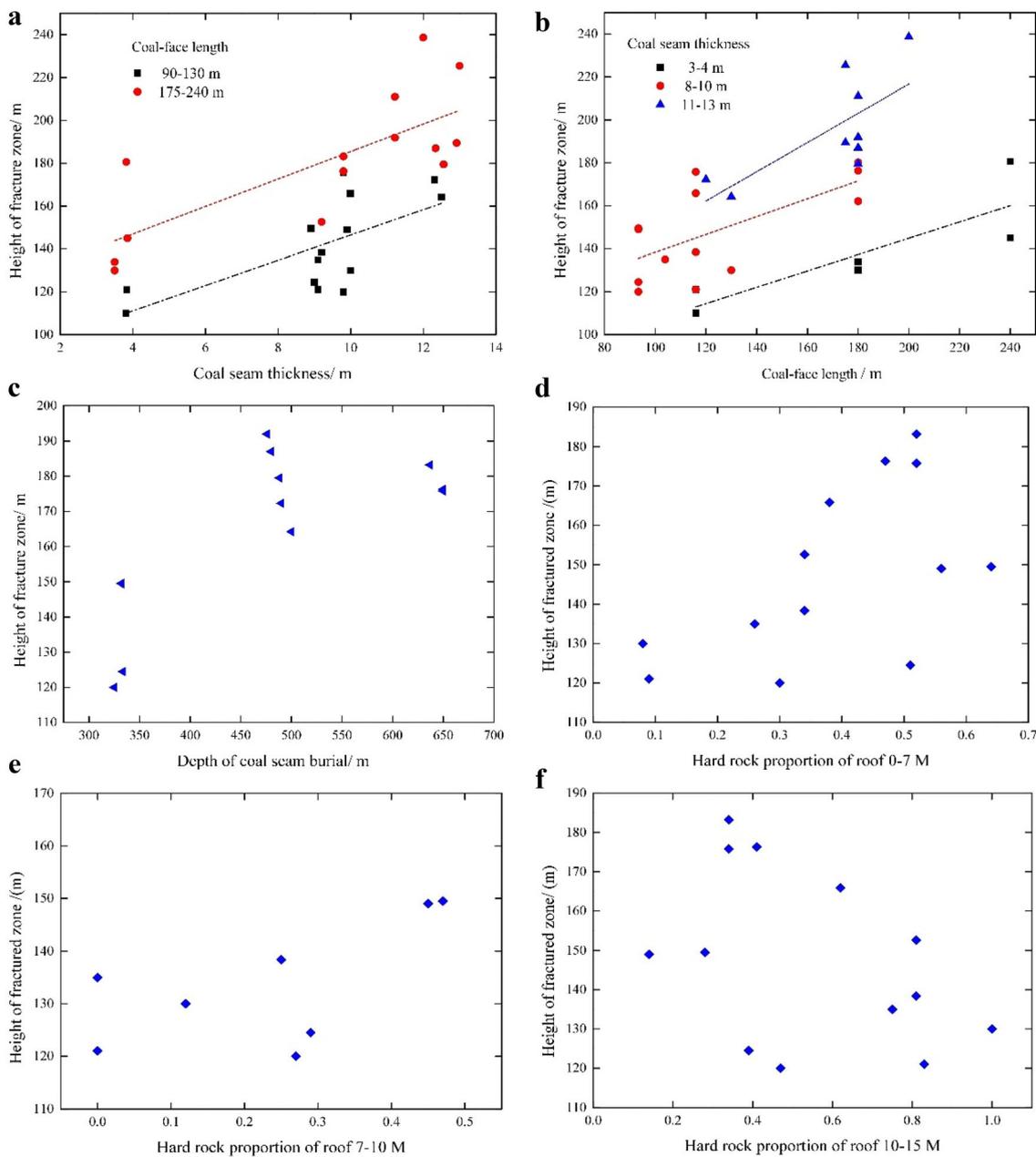M denotes coal seam thickness

Fig. 2 Correlation between WFFZ height and coal seam thickness, working face length, coal seam burial depth and hard rock proportion of roof

working face is large, the height of the water diversion fracture zone is also large.

3.2.2 Working face length

Working face length is an important index of the displacement and failure degree of roofs in coal mining. When the working face length is small, a balanced arch structure is formed in the periphery of the roof-failure area, which leads to a low WFFZ height. Before full mining, WFFZ height increases with increasing working face length, but

after full mining, this increase ceases. According to the observational data obtained in the study area (Fig. 2b), under similar conditions of coal seam thickness, roof category, and other factors, WFFZ height increases with increasing working face length.

3.2.3 Coal seam burial depth

Coal seam burial depth controls the original stress of the surrounding rock. Deeper burial coal seam entails higher original stress on the surrounding rock. With coal mining,

the overburden rock collapses under gravity, causing the stress of surrounding rock to be redistributed, resulting in tensile failure above the mined-out area and shear failure above the coal wall, forming the WFFZ. Accordingly, with the increasing of coal seam burial depth, the unloading and fracture range of the surrounding rock increases gradually after mining. The observational data (Fig. 2c) show that WFFZ height increases markedly when the coal seam burial depth increases from 300 to 500 m, and increases slowly when the burial depth exceeds 500 m.

3.2.4 Roof classification

There is a close relationship between WFFZ height and roof classification. Brittle rocks (conglomerate, sandstone) usually have high compressive strength, but it is difficult to close cracks after failure. The compressive strength of plastic rock (siltstone, mudstone) is low, but the cracks caused by failure are easy to close. The roof of the coal seam often comprises a combination of brittle-plastic rock layers, and different lithologic combinations have different effects on WFFZ height after mining. Some scholars have used the proportion of hard rock, which is the proportion of hard rock (sandstone and conglomerate) to WFFZ height, to characterise the properties of lithologic combinations of the roof (Hu et al. 2012). However, recent results show that lithologic combinations of different positions of the coal seam roof have different effects on WFFZ height after mining (Xu et al. 2009, 2012). In the area close to the coal seam, a hard rock layer will form connected fissures and control the synchronous fracture of soft rock for a given range, conducting the development of a WFFZ. Away from the coal seam, either the hard-rock layer is less affected by mining, which inhibits WFFZ development. Therefore, the proportion of hard rock in the WFFZ is used to characterise the properties of the lithologic combination of the roof, which obscures the effect of different layers of hard rock on the overburden failure process and may lead to large deviations prediction results. In our study, the roof is divided into three parts: A, B, and C (Fig. 3). The observational data (Fig. 2d) show that WFFZ height increases significantly with increasing hard-rock proportion when the distance from the coal roof to its upper strata is 0–7 times the coal seam thickness. With increasing hard-rock proportion in roof B, WFFZ height increases slowly (Fig. 2e). There is a negative correlation between hard-rock proportion in roof C and WFFZ height (Fig. 2f). To describe the effect of the lithologic combination of the roof at different positions in the overburden failure process, the hard-rock proportions of roofs A, B, and C are proposed to jointly characterise the influence of roof category on WFFZ height.

4 Methodology

The SVM model is a prediction model based on small-sample learning, which can be divided into a support-vector classifier and a support-vector regression (SVR) (Vapnik 1995; Fang et al. 2018). The main idea of SVR is to find an optimal surface based on the Vapnik–Chervonenkis generalization theory to minimise the structural risk of the model (Vapnik 1998).

GA is a kind of simulated evolutionary algorithm proposed by Professor Holland of Michigan University in 1969 and summarized by DeJong, Goldberg, and others (Holland 1975; DeJong 1975; Goldberg 1989). The algorithm comes from Darwin's theory of evolution, Weizmann's theory of species selection, and Mendel's theory of population genetics. GA is a type of self-organizing and adaptive artificial-intelligence technology that simulates the process and mechanism of biological evolution in nature to solve extreme-value problems.

SVM parameters greatly impact the prediction ability of the SVM model, therefore, reasonable selection of parameters is important to establish an SVM model. GA is an effective method to select parameters with characteristics of global optimization and computational stability (Lin et al. 2013; Lewis and Randall 2017; Ouyang et al. 2017). Therefore, GA is used to select SVM parameters, avoiding the subjectivity of artificially selecting parameters and improving the prediction ability of SVM. Compared with other methods, SVM optimised by GA (GA-SVM) has the advantages of being only weakly dependent on training samples and having a strong predictive ability.

5 Results and verification

Using the observed WFFZ height and coal seam thickness, working face length, coal seam burial depth and hard rock proportion of roof for the ZK1-ZK21 borehole as training samples, a prediction model for WFFZ height was established by GA-SVM. The ZK22-ZK28 borehole data were applied to the prediction model as verification samples. Through comparative analysis and an engineering application, the superiority and reliability of the model are verified.

5.1 Application of the GA-SVM method

When the GA-SVM method is used, the reciprocal of the sum-squared error of the verification samples is defined as a fitness function. When the population evolves to 80 generations, the sum-squared error of the verification samples and the fitness function (Figs. 4) reach an optimal
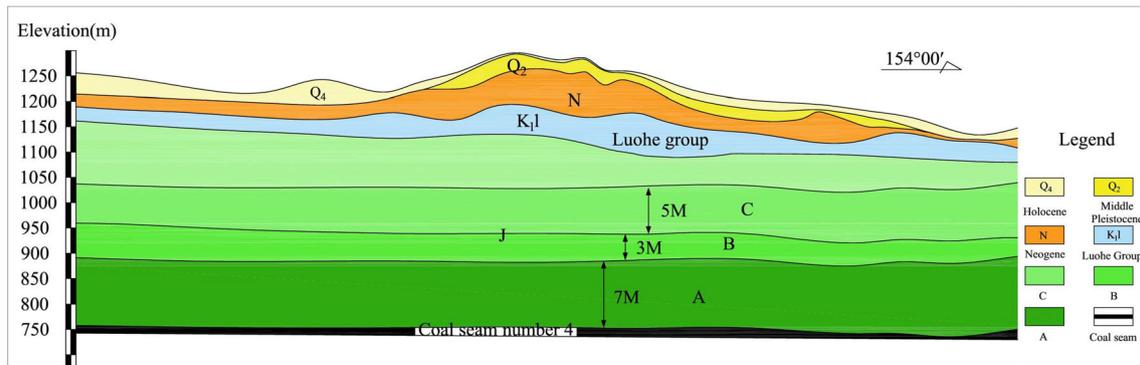
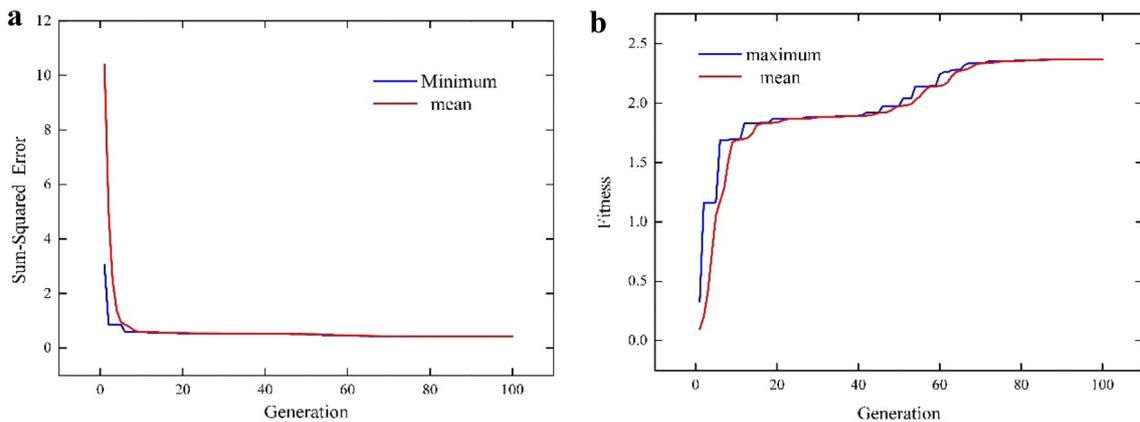**Fig. 3** Geological cross-section of study

Fig. 4 Sum-squared error and fitness-function variation

condition, and the optimum output parameters are $C = 8.001$ and $g = 0.0884$.

The optimum parameters C and g are assigned to the SVM parameters, and the SVM model is established by the training samples:

$$y(x) = \sum_{i=1}^{19} W_i \cdot ^{-0.0884\|X_i - X\|^2} - 0.0946$$

where X_i and W_i are the support vectors and their coefficients, respectively (Table 2), X is the normalised vector of the prediction-sample factors, and $\|X_i - X\|$ is the two-norm distance between the support vector and the normalised vector of the prediction-sample factors.

5.2 Verification

To verify the reliability and superiority of the GA-SVM prediction model, the same training samples were used to establish an MLR and a BPNN prediction model. Then, all sample data were substituted into the three models to obtain predictions of WFFZ height (Fig. 5a). Using to the predicted and observed heights, the residual values of the WFFZ heights were obtained (Fig. 5b).

The determination coefficient R^2 largely reflects the effect of the prediction model. The equation used to calculate R^2 is given below (Gujarati 2003):

$$R^2 = 1 - \frac{\sum (Y_i - \hat{Y})^2 / (n - p - 1)}{\sum (Y_i - \bar{Y})^2 / (n - 1)}$$

where Y_i, $\bar{Y}$, and $\hat{Y}$ are the observed, mean, and predicted values of WFFZ height, respectively; n is the sample size; and p is the number of variables.

The calculation results show that the determination coefficients R^2 of the GA-SVM-based prediction model for the training and verification samples are 0.925 and 0.974, respectively. The corresponding determination coefficients R^2 of the MLR-based prediction model are 0.859 and 0.970, respectively, and those of the BPNN-based prediction model are 0.986 and 0.581, correspondingly. The results show that the SVM- and MLR-based prediction models closely fit the training and verification samples, making the models suitable for WFFZ-height prediction in the study area (Fig. 5a). The determination coefficient of the BPNN-based prediction model for the training samples is much larger than that for the verification sample, and

Table 2 Support vectors and their coefficients

Coefficients, W_i		Support vectors, X_i	Coal seam thickness (m)	Burial depth (m)	Working face length (m)	Hard rock propagation in ranges of roof different heights		
						0–7 M	7–10 M	10–15 M
W_1	4.89	X_1	0.18	− 0.31	− 1.00	− 0.39	1.00	− 0.43
W_2	− 8.00	X_2	0.18	− 0.31	− 1.00	− 0.04	0.72	− 0.56
W_3	4.62	X_3	1.00	− 0.32	− 0.93	0.71	0.38	− 0.56
W_4	− 3.30	X_4	1.00	− 0.31	− 0.93	0.43	− 0.12	0.10
W_5	6.93	X_5	− 1.00	− 0.98	0.14	1.00	0.01	− 0.66
W_6	− 8.00	X_6	− 1.00	− 0.97	0.16	0.54	− 0.38	− 0.40
W_7	− 3.95	X_7	− 0.69	− 0.01	0.18	− 0.96	− 1.00	0.61
W_8	8.00	X_8	− 0.86	− 0.40	0.18	− 0.36	− 1.00	0.43
W_9	− 4.73	X_9	− 0.69	− 0.01	0.20	− 0.07	− 0.46	0.56
W_{10}	− 5.32	X_{10}	− 0.69	− 0.01	0.33	0.57	0.23	− 0.52
W_{11}	− 1.77	X_{11}	− 1.00	− 1.00	0.35	0.71	− 0.03	− 0.98
W_{12}	8.00	X_{12}	− 0.69	− 0.01	0.37	0.07	− 0.51	0.13
W_{13}	− 2.25	X_{13}	− 0.50	1.00	0.37	− 1.00	− 0.74	1.00
W_{14}	5.86	X_{14}	0.18	− 0.54	0.63	0.57	0.55	− 0.75
W_{15}	4.83	X_{15}	0.45	− 0.01	0.79	0.71	− 0.23	− 1.00
W_{16}	− 8.00	X_{16}	0.18	− 0.53	0.86	0.61	− 0.27	− 0.43
W_{17}	− 8.00	X_{17}	0.18	− 0.50	0.91	0.11	− 0.55	− 0.10
W_{18}	2.19	X_{18}	0.11	− 0.13	0.98	− 0.07	− 0.55	0.08
W_{19}	8.00	X_{19}	0.11	− 0.13	1.00	0.21	− 0.55	0.20

M denotes coal seam thickness

over-fitting of the training data occurs without considering the effect of the main factors on WFFZ height. Therefore, compared with GA-SVM and MLR, BPNN performs poorly in WFFZ-height prediction in the study area.

The residual plot (Fig. 5b) shows that the residual value of the GA-SVM prediction model is between −28.69 and 15.57 m in all samples, and the sample size with a residual value between −5 and 5 m is 22, accounting for 78.57% of the total samples. The residual value of the MLR prediction model is between −30.66 and 14.79 m, and the sample size with a residual value between −5 and 5 m is 14, accounting for 50% of the total samples. The results above show that compared with the MLR model, the GA-SVM model can better eliminate the influence of outlier points on the model and can establish a model that most samples obey, achieving accurate fitting of most of the sample data. Therefore, the GA-SVM model has more advantages in the prediction of WFFZ height.

To further verify the reliability of the model, the GA-SVM-based model was used in the first mining area of the Wenjiapo coal mine (Fig. 6).

The Wenjiapo coal mine is located to the east of the study area, with an area of about 79.69 km². Coal seam No. 4 is the current main coal seam, with a thickness of 1.08–15.00 m and a burial depth of 300–500 m. The main aquifer is the bedrock-fracture aquifer of the Luohe Group. The thickness of the combined aquifuge between coal seam No. 4 and the aquifer in the first mining area of the Wenjiapo coal mine is between 180 and 235 m (Fig. 7). From the northwest to the southeast of the first mining area, the thickness of the combined aquifuge first increases and then decreases.

Using drilling data from the first mining area of the mine, thematic maps of coal seam thickness, burial depth, and hard-rock proportions in different layers were drawn (Fig. 8). A thematic map of coal seam thickness (Fig. 8a) shows that the thickness of coal seam No. 4 in the first mining area of the Wenjiapo coal mine is between 3 and 15 m, with an average thickness of 11.3 m, low thickness in the east and high thickness in the west, reaching a maximum in the southwest of the first mining area. The thematic map of burial depth (Fig. 8b) shows that the burial depth of the coal seam is between 250 and 550 m, with an average burial depth of 395 m. From the northwest to the southeast, the burial depth of the coal seam is deep-shallow-deep. The thematic maps of hard-rock proportion, when the distance from the coal roof to its upper strata is 0–7 times, 7–10 times, and 10–15 times the coal seam
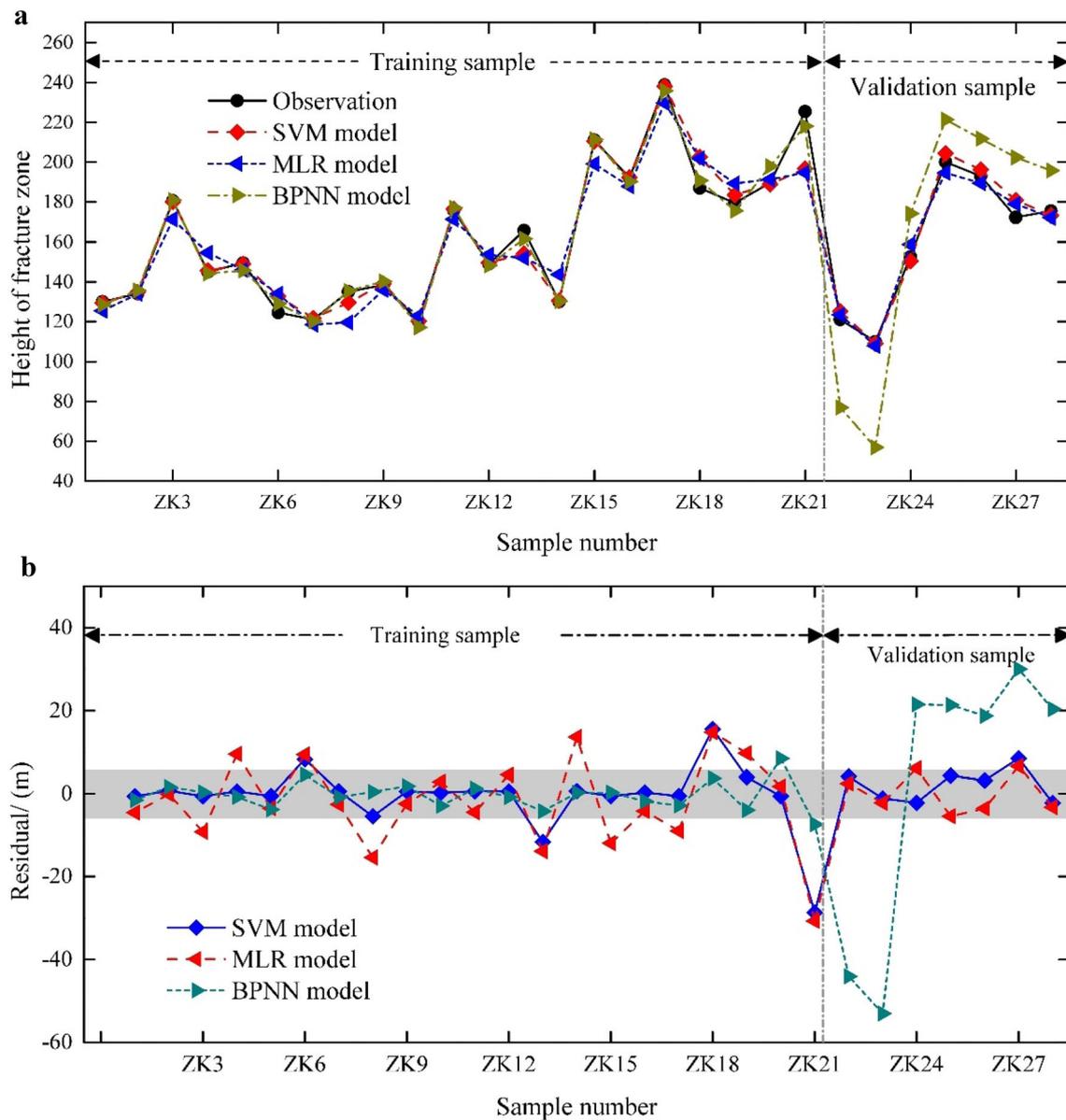
Fig. 5 Prediction and residual values of WFFZ height

thickness, show that the proportion of hard rock in the different roof-thickness ranges is basically between 0.3 and 0.9. In addition, the mining data show that the length of the coal face in the first mining area is 240 m.

The thematic maps of coal seam thickness, coal seam-burial depth, and hard-rock proportion for roofs A, B, and C in the first mining area were applied to the established GA-SVM prediction model. Using the predicted height of the WFFZ and the thickness of the combined aquifuge in the first mining area, a connectivity map of the roof aquifer after mining was drawn (Fig. 9).

The connectivity map of the roof aquifer after mining (Fig. 9) shows that of the three coal faces mined in the first mining area, coal faces 4101 and 4102 have not been

connected to the roof aquifer of the Luohe Group, while the WFFZ formed after mining coal face 4103 has been connected.

According to the measured working face mine-water inflow (Fig. 10), since mining of coal face 4103 began in August 2018, the mine-water inflow from coal faces 4101 and 4102 is about 10 m^3/h and 90 m^3/h, respectively, while the mine-water inflow from coal face 4103 is 180 m^3/h, the latter continuing to exhibit an increasing trend. This indicates that after the mining of coal faces 4101 and 4102, the roof aquifer of the Luohe Group is not connected to the goaf, and the mine-water inflow is relatively small. However, the aquifer of the Luohe Group was connected to the goaf after the mining of coal face 4103, which led to a
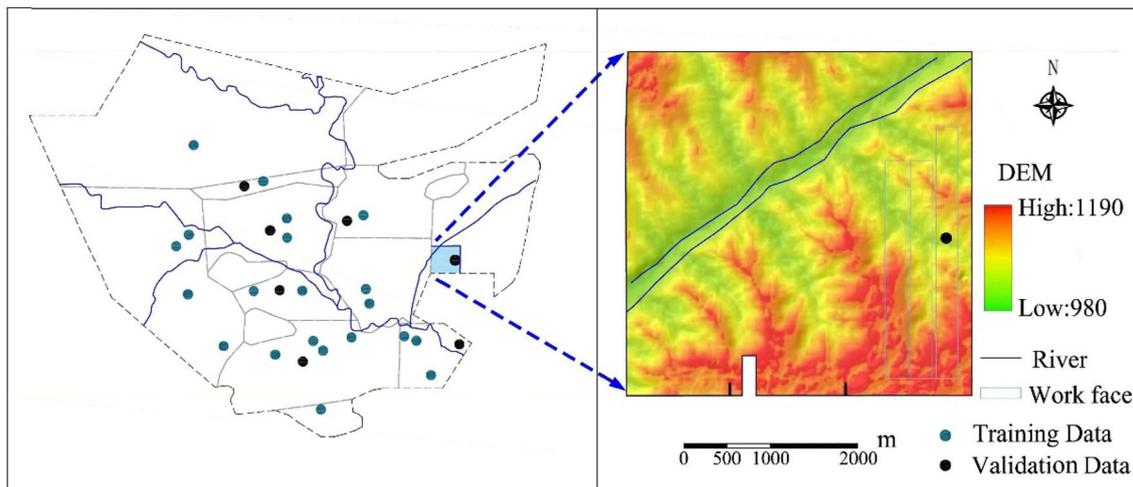
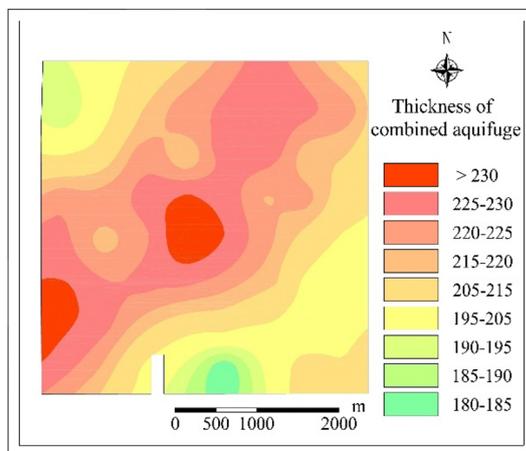**Fig. 6** Location of the first mining area in the Wenjiapo coal mine

Fig. 7 Thickness of the combined aquifuge

significant increase of mine-water inflow. This result is basically consistent with the results of the prediction model and further shows that the GA-SVM-based prediction model is reliable.

After completing the WFFZ-height prediction in the study area, some additional information about the influence of each factor on WFFZ height is obtained. Specifically, Fig. 11 shows 6 variables sorted according to average decreasing accuracy. The average decrease in accuracy is a measure of the contribution of each variable in the SVM model. The more the accuracy of the SVM model decreases due to the exclusion of variables, the more important is the variable; therefore, the greater the average decrease in accuracy, the more important the variable. According to this index, the most important variable is coal seam thickness, followed by working face length and the ratio coefficient of hard rock proportion in the range of 0–7 times the thickness of the coal seam roof.

6 Discussion

In the prediction of roof-gushing (inrush) conditions, the study of WFFZ height is important. Research shows that in coal mining, coal seam thickness, burial depth, working face length, and roof category are contributory factors in roof-WFFZ height. A coal seam roof often comprises a combination of brittle-plastic rock layers, and the hard rock in different roof ranges has different effects on WFFZ height after mining. Therefore, roof category cannot be divided simply into three types: weak, medium-hard, and hard. The hard rock proportions in different roof ranges can be used to characterise the influence of roof category on WFFZ height. This is because the effect of hard rock in different roof ranges on WFFZ height differs. Among the factors of WFFZ height, the most important variable is coal seam thickness, followed by working face length and the ratio coefficient of hard rock proportion in the range of 0–7 times the thickness of the coal seam roof.

In addition, the results show that compared with the MLR and BPNN methods, the GA-SVM-based prediction model has a high fitting accuracy for the training and verification samples and can be used to determine WFFZ height in the study area. This is because, compared with other methods, GA-SVM may meet the accuracy requirements when few training samples are provided. Besides, the method may also effectively reduce the impact of outlier values in the model.

7 Conclusion

Using the Binchang mining area as a study area, we used mathematical statistics to analyse the factors that influence WFFZ height during the mining of coal seams. GA-SVM
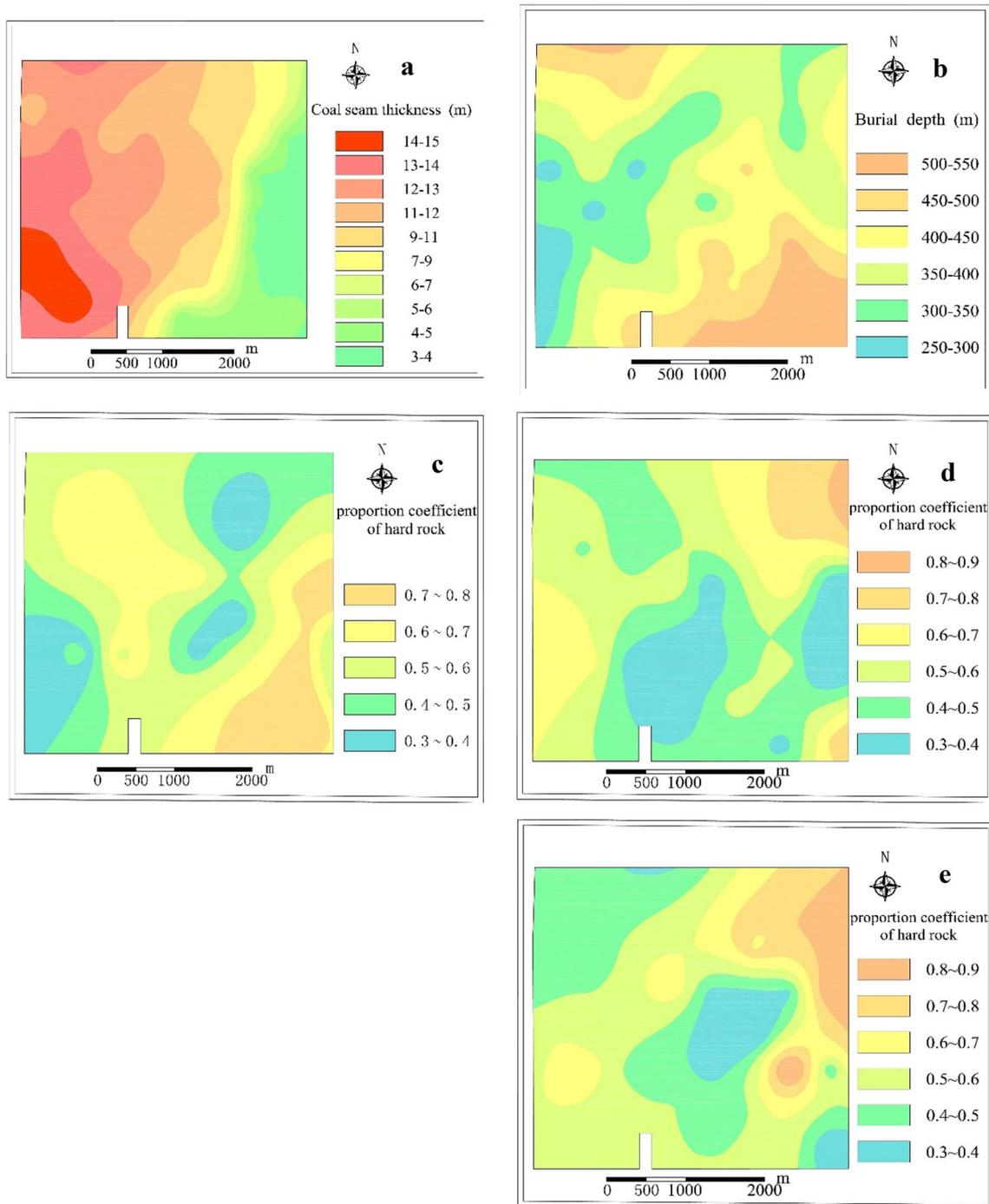
Fig. 8 Thematic maps of main control factors of the water-flowing fracture zone height

was used to establish a prediction model for WFFZ-roof height. The results of this analysis reveal a number of findings outlined below.

(1) The factors that influence WFFZ height in the study area are coal-seam-mining thickness, burial depth, working face length, and roof category. There is a negative correlation between WFFZ height and the

ratio coefficient of hard rock proportions in the range of 10–15 times the thickness of the coal seam roof. There is a positive correlation between coal seam thickness, working face length, coal seam burial depth, hard rock proportion in the range of 10–15 times the thickness of the coal seam roof and WFFZ height.
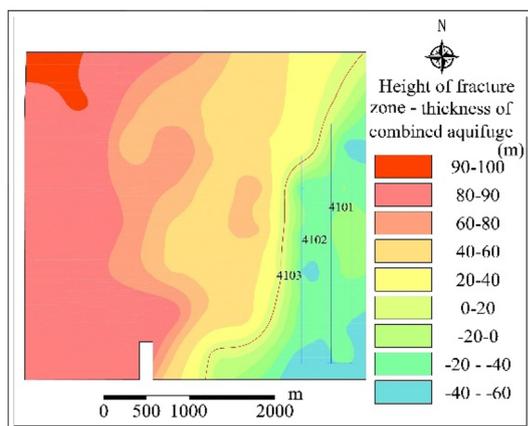
Fig. 9 Connectivity of roof aquifer after mining
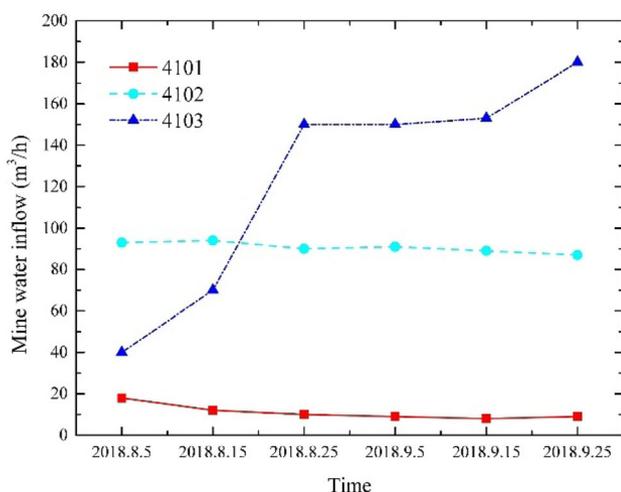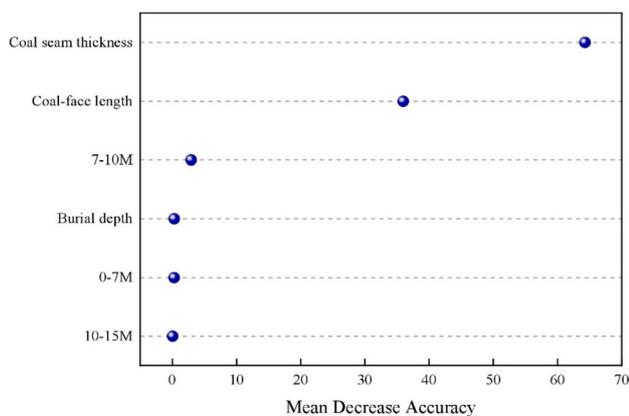

Fig. 10 Working face mine-water inflow curve

Fig. 11 Mean decrease in accuracy

(2) Among the variables, the most important is coal seam thickness, followed by working face length, and the ratio coefficient of hard rock proportion in

the range of 0–7 times the thickness of the coal seam roof.

(3) Compared with MLR and BPNN, the GA-SVM model is more accurate for the verification samples and is more suitable for the prediction of WFFZ-roof height in coal mining.

This research not only proposes a new index to characterise the effect of coal seam-roof category on WFFZ height but also provides a new method for the prediction of WFFZ height in coal mining for the future; this is of great significance for the prediction of roof-gushing (inrush) conditions in coal mining.

References

Chen PP, Liu HQ, Zhu ZX, Yan Y (2005) Height forecast of water conducted zone with top coal caving based on artificial neural network. J China Coal Soc 30(4):438–442

DeJong KA (1975) The analysis of the behavior of a class of genetic adaptive systems. University of Michigan, Ann Arbor

Dong QH, Liu F, Zhang Q (2015) Deformation of soil and rock transition belt caused by the mining damage. Eng Geol Soc Territ 6:813–817

Du SG, Weng XH (1997) Relationship between dip angle of coal seam and height of three zones. J Eng Geol 5(3):211–217

Fang HT, Jhong BC, Tan YC, Ke KY, Chuang MH (2018) A two-stage approach integrating SOM- and MOGA-SVM-based algorithms to forecast spatial–temporal groundwater level with meteorological factors. Water Resour Manag 33:797–818

Gao Y (1996) Four-zone" model of rock mass movement and back analysis of dynamic displacement. J China Coal Soc 21(1):51–56

Gao XC, Wu YP (2011) Experimental study on dynamic distribution law of the spread cracks overlying stratum with rich water in special hard thick seam. Saf Coal Mines 42(3):16–18

Garritty P (1983) Water flow into undersea mine workings. Int J Min Eng 1(3):237–251

Goldberg DE (1989) Genetic algorithms in search, optimization and ma-chine learning. Addison-Wesley Longman Press, Boston

Gujarati D (2003) Basic econometrics, 4th edn. McGraw-Hill, New York

Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor

Hu XJ, Li WP, Cao DT, Liu MC (2012) Index of multiple factors and expected height of fully mechanized water flowing fractured zone. J China Coal Soc 37(4):613–620

Lewis A, Randall M (2017) Solving multi-objective water management problems using evolutionary computation. J Environ Manag 204:179–188

Li JS, Li ZY (2012) Height analysis of water flowing fractured zone of a coal mine mining under large-scale water mass. Saf Coal Mines 43(12):190–192

Lin HF, Li SG, Cheng LH (2010) Model experiment of evolution pattern of mining-induced fissure in overlying strata. J Xi'an Univ Sci Technol 30(5):507–512

Lin GF, Jhong BC, Chang CC (2013) Development of an effective data-driven model for hourly typhoon rainfall forecasting. J Hydrol 495:52–63

Liu TQ (1995) Influence of mining activities on mine rock mass and control engineering. J China Coal Soc 1:1–5

Liu Y (2010) Research on relationship between different width of working faces and height of water-flowing fractured zone. Saf Coal Mines 41(4):13–17

Liu S, Li W, Wang Q (2018a) Height of the water-flowing fractured zone of the Jurassic coal seam in northwestern China. Mine Water Environ 37(2):312–321

Liu WT, Pang LF, Liu YB, Du YH (2018b) Characteristics analysis of roof overburden fracture in thick coal seam in deep mining and engineering application of super high water material in backfill mining. Geotech Geol Eng 37:1–10

Luan YZ, Li JT, Ban XH, Sang CY, Zhang CQ, Ma DP (2010) Observational research on the height of water flowing fractured zone in repeated mining of short distance coal seams. J Min Saf Eng 27(1):139–142

Meng XX, Liu WT, Mu DR (2018) Influence analysis of mining's effect on failure characteristics of a coal seam floor with faults: a numerical simulation case study in the Zhaolou coal mine. Mine Water Environ 37:754–762

Ouyang Q, Lu WX, Hou ZY, Zhang Y, Li S, Luo JN (2017) Chance-constrained multi-objective optimization of groundwater remediation design at DNAPLs-contaminated sites using a multi-algorithm genetically adaptive. J Contam Hydrol 200:15–23

Peng S, Zhang J (2007) Engineering geology for underground rocks pp 221–260

Qiao W, Li WP, Li T, Chang JY, Wang QQ (2017) Effects of coal mining on shallow water resources in semiarid regions: a case study in the Shennan mining area, Shaanxi, China. Mine Water Environ 36(1):104–113

Shi LQ, Xin HQ, Zhai PH, Li SC, Liu TB, Yan Y, Wei WX (2012) Calculating the height of water flowing fracture zone in deep mining. J China Univ Min Technol 41(1):37–41

Sui WH, Hang Y, Ma LX, Wu ZY, Zhou YJ, Long GQ, Wei LB (2015) Interactions of overburden failure zones due to multiple-seam mining using longwall caving. Bull Eng Geol Environ 74(3):1019–1035

Vapnik V (1995) The nature of statistical learning theory. Springer, New York

Vapnik V (1998) Statistical learning theory. Wiley, New York

Wang SM (2006) A brief review of the methods determining the height of permeable fracture zone. Hydrogeol Eng Geol 33(5):126–128

Wang Z, Wang C, Wang Z (2018) The hazard analysis of water inrush of mining of thick coal seam under reservoir based on entropy weight evaluation method. Geotech Geol Eng 36(5):3019–3028

Wu Q, Huang XL, Dong DL, Yin ZR, Li JM, Hong YQ, Zhang HJ (2000) Three maps-two predictions" method to evaluate water bursting conditions on roof coal. J China Coal Soc 25(1):60–65

Wu Q, Fan Z, Zhang Z, Zhou W (2014) Evaluation and zoning of groundwater hazards in Pingshuo No. 1 underground coal mine, Shanxi province, China. Hydrogeol J 22(7):1693–1705

Wu Q, Liu YZ, Luo LH, Liu SQ, Sun WJ, Zeng YF (2015) Quantitative evaluation and prediction of water inrush vulnerability from aquifers overlying coal seams in Donghuantuo Coal Mine, China. Environ Earth Sci 74(2):1429–1437

Wu Q, Xu K, Zhang W (2016) Further research on "three maps-two predictions" method for prediction on coal seam roof water bursting risk. J China Coal Soc 41(6):1341–1347

Xu ZM, Sui WH (2013) Statistical prediction of overburden failure due to coal mining under sea area. In: Huang Y, Wu F, Shi Z, Ye B (eds) New frontiers in engineering geology and the environment. Springer, Berlin, pp 255–258

Xu ZM, Sun YJ (2008) Height prediction of the water conducted zone for mining under reservoir. China Min Mag 17(3):96–99

Xu JL, Wang XZ, Liu WT, Wang ZG (2009) Influence of primary key stratum location on the height of water flowing fracture zone. Chin J Rock Mech Eng 28(2):381–385

Xu JL, Zhu WB, Wang XZ (2012) New method to predict the height of fractured water-conducting zone by location of key strata. J China Coal Soc 37(5):762–769

Yang DM, Guo WB, Tan Y (2018) Application of magnetotelluric method to the detection of overburden failure height in shallow seam mining. Arab J Geosci 11:350

Yin HY, Lefticariu L, Wei JC, Zhu L, Guo JB, Li ZJ, Guan YZ (2016) A multi-method approach for estimating the failure depth of coal seam floor in a longwall coal mine in China. Geotech Geol Eng 34(5):1267–1281

Zhang JC, Peng SP (2005) Water inrush and environmental impact of shallow seam mining. Environ Geol 48(8):1068–1076

Zhang J, Yang T (2018) Study of a roof water inrush prediction model in shallow seam mining based on an analytic hierarchy process using a grey relational analysis method. Arab J Geosci 11(7):153

Zhang SC, Guo WJ, Li YY, Sun WB, Yin DW (2017) Experimental simulation of fault water inrush channel evolution in a coal mine floor. Mine Water Environ 36(3):443–451

Zhang YB, Xu YX, Wang K, Chen P, Wang X, Zheng Q, Zhang ZX (2018) The fracturing characteristics of rock mass of coal mining and its effect on overlying unconsolidated aquifer in Shanxi, China. Arab J Geosci 11(21):666

Zhao DS, Chen F, Wang ZC (2011) The experimental study on the failure of soft overburden in mining of thick coal seam. J Guangxi Univ (Nat Sci Ed) 36(1):177–181

Zhu G, Wu X, Yu S, Qian C, Dong YY, Zhang CJ, Wu C (2018) Surface water control for mining thick, relatively shallow coal seams in the Loess area of western China. Mine Water Environ 37(3):442–455