Mine Water and the Environment (2022) 41:225–236
https://doi.org/10.1007/s10230-021-00805-y

TECHNICAL ARTICLE

Predicting the Height of the Water-conducting Fractured Zone Based on a Multiple Regression Model and Information Entropy in the Northern Ordos Basin, China

Peng Yang[1] · Weifeng Yang[1] · Yuxu Nie[1] · Fahad Saleem[1] · Fei Lu[2] · Rongkun Ma[1] · Ruipeng Li[1]

Received: 4 February 2021 / Accepted: 28 July 2021 / Published online: 10 August 2021

Abstract
Predicting the water-conducting fractured zone (WCFZ) height is the key to roof water disaster prevention and environmental protection. Field WCFZ data were collected from 24 working faces of Shendong, Shenfu, Yushen, and Yuheng mining areas in the northern Ordos Basin. A fitting formula was obtained by a multiple regression model and the weight of each factor was calculated by information entropy theory, allowing a prediction of the WCFZ height in the area. The formula was improved by applying a safety factor of 1.33, which met the 95.8% reliability requirements of the samples. Finally, the 31,114 working face of the Jinjie coal mine was selected as a case study. The spatial characteristics of the WCFZ in the study area were analyzed using a geographic information system (GIS) and the WCFZ height was measured using borehole TV. The predicted value was 49.6 m and the field measured value exceeded 42.4 m, making the model-predicted value much closer that predictions calculated with other empirical formulas. The results indicate that this method quantitatively predicted the WCFZ height in the study area and enabled the WCFZ height in the study area to be visualized with GIS technology. It is an effective tool for analyzing and predicting the height of the WCFZ and for preventing mine water inrush.

Keywords Mine inrush safety · Jinjie coal mine · Nonlinear multiple regression · Geographic information system · Field measurements

Introduction

The Ordos Basin is the most productive of China's coal resource areas. The Jurassic Yan'an formation is the primary coal-bearing stratum in the basin and its roof generally contains Jurassic Zhiluo formation, Cretaceous, and Quaternary Salawusu formations. Therefore, coal mining in the basin faces different degrees of roof water disaster risks (Gui and Lin 2016; Hu and Zhao 2021). The northern Ordos Basin is an arid and semi-arid region with a fragile ecological environment (Li 2018; Sun et al. 2017; Yang et al. 2020a, 2021).

According to the deformation and failure characteristics of the overlying strata during coal mining, the overlying strata are distributed in three distinct zones from bottom to top, named the caved, fractured, and continuous deformation zones (Liu 1981; Peng et al. 2015). The concept is illustrated in Fig. 1. The height of the water-conducting fractured zone (WCFZ) is the sum of the heights of the fractured and caved zones. When the fractures connect with the upper aquifer, a mine water inrush channel is formed, which can easily cause a water inrush disaster and worsen the local ecological environment. Therefore, predicting the height of the WCFZ is the key to roof water disaster prevention and ecological environment protection.

Various research methods have been used to study the height of the WCFZ, such as empirical equations, similar material model tests, numerical simulation, theoretical analysis, and field measurements. Based on these methods, researchers have studied the height of the WCFZ in Jurassic coal seam mining in the Ordos Basin, in western China. In the 1980s, Liu (1981) established some empirical equations based on measured data to derive the height of the WCFZ

✉ Weifeng Yang
yangwf888@163.com

Peng Yang
yangp@cumt.edu.cn

[1] School of Resources and Geosciences, China University of Mining and Technology, Xuzhou 221116, Jiangsu, China

[2] Shendong Coal Grp Co Ltd. Jinjie Coal Mine, Yulin 719315, Shaanxi, China

using the mining height and overburden rock hardness as variables. In the "Codes for Setting Coal Pillar and Mining Pressed Coal Under Buildings, Water Bodies, Railways, and Main Roadways in China" (from now on referred to as the "Codes"), the most widely used empirical equations for medium hard rock conditions are presented in Eq. 1 of Table 1 (State Administration of Work Safety 2017). However, China's eastern and western parts have very different geological conditions and mining technology. Wu et al. (2016) calculated the development height of medium-hard rock WCFZ using Eq. 2 from the coal mine water prevention and control manual (Wu et al. 2013) and evaluated the roof caving degree of the Taigemiao mining area in the northern Ordos Basin, and achieved good results (Table 1). Xing et al. (2017) summarized the nonlinear statistical relationship between the height of the WCFZ and the mining height, mining depth, and coalface length in Jurassic coal mines and developed a formula to predict the height of the WCFZ. These empirical formulas are prone, however, to poor predictive accuracy. Our challenge was to improve the predictive accuracy.

In the laboratory, similar material model testing is a widely used research method. Miao et al. (2011) used a similar material model test to illustrate that if the vertical distance from the primary key stratum to the coal seam is lesser than a certain value, the developed fracture zone would extend above the key stratum to the top of bedrock, resulting in a higher WCFZ height than the calculated empirical method in Bulianta coal mine. Zhu et al. (2020) established a soil-rock model using numerical simulation to study the WCFZ height under the Q_{21} loess at the Jinjitan coal mine. In terms of theoretical analysis, Liu et al. (2018) used Pu's theory and rock mass limit equilibrium theory to calculate the height of the WCFZ in the Jurassic coal field, and used micro-resistivity scanning imaging and logging technology to study the height of the WCFZ in different Jurassic coalfield layers in western China. Liu et al. (2019a) determined the WCFZ height in soil-rock composite structure overburdens in northern Shaanxi Province by monitoring the quantity of running water in boreholes monitoring and distributed optical fiber sensing technology.

In recent years, multiple regression analysis (Chen and Zhu 2020; He et al. 2020; Hu et al. 2021; Liu et al. 2019a, 2019b; Wang et al. 2020; Wu et al. 2020; Xing et al. 2017), random forest regression (Zhao and Wu 2018), neural network (Wu et al. 2017), PSO-SVR (Xue et al. 2020), MPGA-SVR (Guo et al. 2020), and some other data analysis methods have been used in hydrogeological studies, and, to a certain extent, to improve the predictive accuracy of the empirical equation for the WCFZ height. However, due to the diverse geological and engineering conditions of China's mining areas, the predictive method's adaptability is low.

To improve the accuracy and applicability of the predictive method, the northern Ordos Basin was selected as a study area. The overlying strata of the coal seam are medium hard rocks, and the engineering and geological conditions

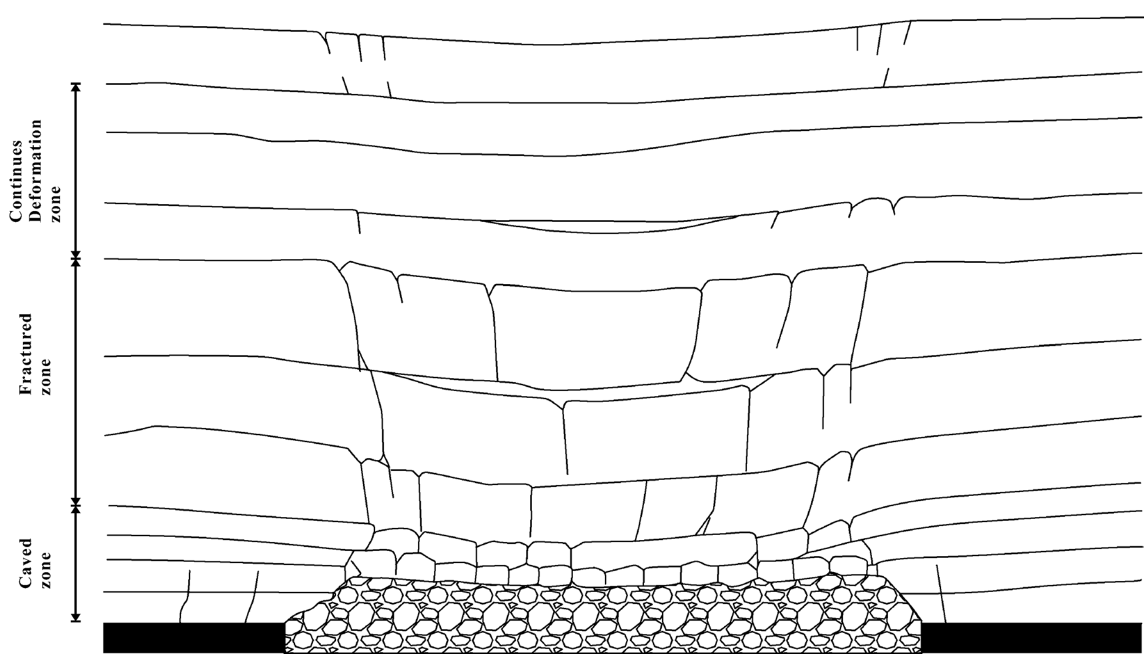

Fig. 1 Zones of overburden failure due to mining

are similar. The height of the WCFZ and its predictive variables were determined using data collected from 14 coal mines in the northern Ordos Basin. The multiple regression model was used to obtain the fitting formula and the information entropy was used to improve each factor's weight in a WCFZ height predictive method. The 31,114 working face of the Jinjie coal mine was used as a case study. The spatial characteristics of the WCFZ height were analyzed by GIS technology in the study area, and the WCFZ height was measured by borehole TV. Finally, the feasibility and precision of this approach were confirmed by error analysis and a comparison of predicted and measured values.

Material and Methods

Information Entropy

Entropy is a measure of the variance of random variables in information theory. It can be used to determine the amount of information stored in the data as well as the discrete degree of the data. A greater discrete degree of data is associated with less information entropy, more information deviation, more contained information, and more significant output results (Patricia 2006; Yang et al. 2017, 2020b). The information entropy of the factors is (Li et al. 2013; Wu et al. 2015):

$$E_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} U_{ij} \ln(U_{ij}) \quad (1)$$

where $U_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}}$, x_{ij} is the value of the jth factor of the ith group of the height of the WCFZ in the measured data. There are 24 data sets in this study, so the maximum value of i is 24; there are three factor sets in this paper, so the maximum value of j is 3. The weight coefficient of the jth factor is:

$$w_j = \frac{1 - E_j}{\sum_{j=1}^{p} \left(1 - E_j\right)} \quad (2)$$

Improved Multiple Regression Model

Multiple regression models includes multiple linear and nonlinear regression models. The multiple linear regression model can be defined as (Wang et al. 2020; Wu et al. 2020):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \quad (3)$$

where y, the dependent variable, is the height of the WCFZ, x_1 to x_p are independent variables, such as the thickness of the coal seam, buried depth of coal seam, inclined length of working face,β_0, β_1, β_2, ..., β_p are $p+1$ unknown parameters ($p \geq 2$), β_0 is a regression constant, β_1, β_2, ..., β_p are regression coefficients, and ε is a random error.

To improve the accuracy of the multiple regression model, it needs to be tested for significance using the F or t tests. The F test statistic is used when $F > F_\alpha(p, \ n-p-1)$; the multiple regression model is considered significant at the significance level α, which means that the multiple regression model is available.

$$F = \frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2} \times \frac{n-p-1}{p} \quad (4)$$

where n is the number of measured data sets, y_i is the measured height of the WCFZ (m), $\hat{y}_i$ is the fitting height of the WCFZ of samples (m), and $\bar{y}$ is the average height of the WCFZ of samples (m).

The t-test statistic is used when $\left|t_j\right| \geq t_{\frac{\alpha}{2}}(n-p-1)$, the independent variable x_j significantly affects the dependent variable y; that is, the influencing factors have a significant impact on the predictive object.

$$t_j = \frac{\hat{\beta}_j}{\sqrt{c_{jj}}\hat{\sigma}} \quad (5)$$

where $c_{jj} = \frac{Var(\hat{\beta}_j)}{\sigma^2}$ and $\hat{\sigma} = \sqrt{\frac{1}{n-p-1} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)}$ stands for the acceptable deviation of the controlled variable.

The goodness of fit is then used to test the regression equation's degree of fit to the samples' observed values. The coefficient of determination, R^2, ranges between 0 and 1: the closer to 1, the better the fit; the closer to 0, the worse the fit.

Table 1 Empirical equations to calculate the height of the WCFZ

Lithology of overburden	Equation (1)	Equation (2)
Medium-hard	$H_{li} = \frac{100 \sum M}{1.6 \sum M + 3.6} \pm 5.6$	$H_{li} = \frac{100M}{0.26M + 6.88} \pm 11.49$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n} \left(y_i - \overline{y}_i\right)^2} \quad (6)$$

In this paper, the information entropy theory was used to determine the weight coefficient w_j of each factor. Therefore, the improved multiple regression model is as follows (Liu et al. 2019b):

$$y = \beta_0 + w_1\beta_1 x_1 + w_2\beta_2 x_2 + \cdots + w_p\beta_p x_p + \varepsilon \quad (7)$$

Measured Data

To evaluate the influence of the factors on the height of the WCFZ, 24 groups of measured data and relevant factors were collected from coal mines in the Shandong, Shenfu, Yushen, and Yuheng mining areas of the northern Ordos Basin (Lv 2014; Wei et al. 2016; Xing et al. 2017; Xue et al. 2020). The mining area locations are shown in Fig. 2 and the collected data are listed in Table 2.

Results and Discussion

Multiple Regression Predictive Model

With the data in Table 2, a multiple linear regression model was used to determine the relationship between the height of the WCFZ (H, in m), and its factors, and a formula for this relationship was obtained. The height of the WCFZ can be calculated with:

$$H = 3.24 + 12.07M + 0.29b - 0.07l \quad (8)$$

where M is the thickness of the coal seam (m), b is the buried depth of the coal seam (m), and l is the inclined length of the working face (m).

SPSS software was used to calculate the significance probability value (P value) of the regression equation. The P value is $0.000 \le \alpha = 0.05$, which indicates that the multiple linear equation is highly significant. The t statistic and P-value of the multiple linear regression models were then statistically analyzed (Table 3).

It can be seen from Table 3 that at a significance level of $\alpha = 0.05$, the thickness and buried depth of the coal seam significantly affected the height of the WCFZ, while the inclined length of the working face was not significant. Although the regression equation has high significance in the multiple linear regression model, some factors were not significant effect. Therefore, it was necessary to use multiple nonlinear models to predict the height of the WCFZ.

The relationship between the height of the WCFZ and its influencing factors are analyzed and plotted in Fig. 3. In SPSS software, 11 basic models were used to determine the relationship between the height of the WCFZ and its influencing factors. The R^2 and significance of the models are listed in Table 4.

According to each model's R^2 and significance, the optimal relationship between each factor and the WCFZ height was determined.

$$\begin{aligned} H &= e^{5.45 - \frac{4.03}{M}} \\ H &= 0.334b^{1.06} \\ H &= -155.282 + 2.298l - 0.005l^2 \end{aligned} \quad (9)$$

After the multiple regression model equations were established, information entropy was applied to improve the weight coefficient of each influencing factor. Information entropy was used to measure each factor's information content and evaluate its usefulness in predicting the WCFZ height. For example, if one of the factors has the same value in each measured data set, then that factor does not provide useful information about the WCFZ's height; since it does not affect the WCFZ height, the information entropy is 1. On the other hand, if it significantly affects the measured WCFZ data, then the factor's effect is more significant. The information entropy and calculated weight coefficient of factors based on the information entropy theory are shown in Table 5. A multiple nonlinear regression model was established to predict the WCFZ height with information entropy:

$$H = 0.39e^{5.45 - \frac{4.03}{M}} + 0.13b^{1.06} + 0.52l - 0.0011l^2 - 35.47 \quad (10)$$

Error Analysis and Reliability of the Results

The empirical equations in the "Codes" and coal mine water prevention and control manual (Wu et al. 2013) have both been used to predict the height of the WCFZ in the western mining area. To compare the accuracy of the equation based on the multiple regression model and information entropy with other empirical equations, the relative error of measured data from 24 working faces was calculated (Table 6).

The standard form of presentation for the error analysis of the Code and paper equation for the height of the WCFZ and presented in Table 6 would generally be as shown on Fig. 4, which compares the relative reliability of the predictive model to the Code equations. We find that Eq. 10 does a good job of predicting the actual height of the WCFZ, whereas Code Eqs. 1 and 2 substantially and unsafely underpredicted the measured height of the WCFZ.

The commonly used evaluation indicators for regression models are mean error (ME), mean absolute error (MAE),

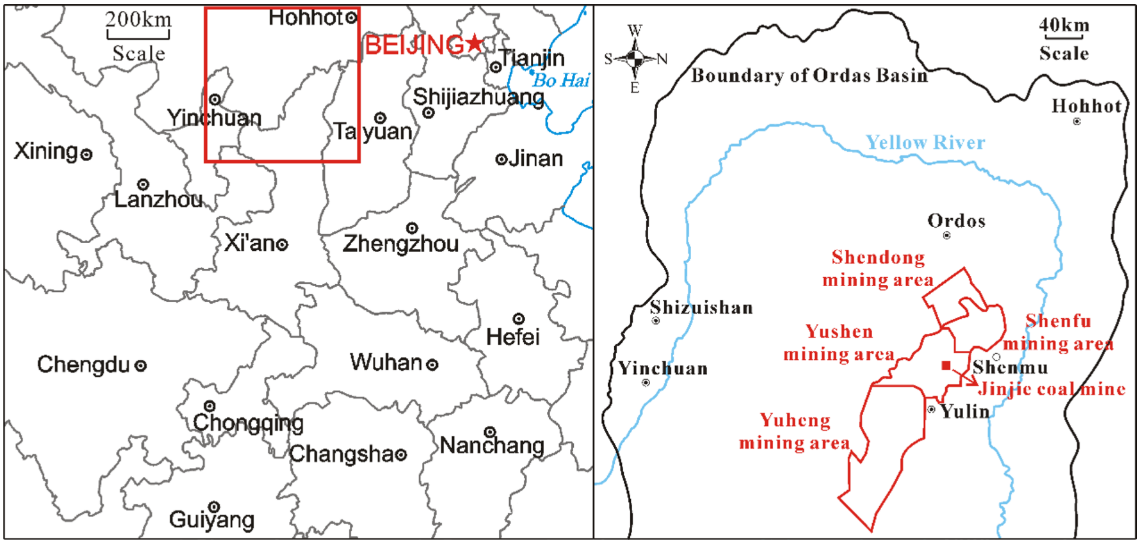

Fig. 2 Distribution of mine area in the northern Ordos Basin

Table 2 Measured data of the height of the WCFZ in the northern Ordos Basin

No	Coal mine	Working face	Thickness of coal seam /m	Buried depth of coal seam /m	Inclined length of working face /m	The measured height of the WCFZ/m	Mine area
1	Cuncaota	43,115	2.3	101.6	360.0	25.2	Shendong
2	Wulanmulun	12,403	2.0	130.0	310.0	35.7	Shendong
3	Wulanmulun	12,403	2.5	130.0	310.0	62.9	Shendong
4	Huoluowan	22,101	2.4	139.0	/	37.8	Shendong
5	Zhuanlongwan	23,103	4.5	145.8	260.0	92.1	Shendong
6	Bulianta	12,406	4.4	180.0	310.0	74.0	Shendong
7	Bulianta	12,406	4.4	181.7	310.0	89.5	Shendong
8	Bulianta	12,511	7.6	242.1	319.1	104.5	Shendong
9	Jinjie	31,104	3.4	110.0	369.0	45.7	Yushen
10	Yushuwan	10,204	5.0	275.8	255.0	117.8	Yushen
11	Yushuwan	10,204	5.0	278.5	255.0	130.5	Yushen
12	Yushuwan	10,204	5.0	286.9	255.0	138.9	Yushen
13	Yushuwan	10,203	5.0	279.3	297.0	137.3	Yushen
14	Hanglaiwan	30,101	4.5	243.5	300.0	108.3	Yushen
15	Longde	205	4.0	210.0	182.0	76.9	Yushen
16	Jinjitan	101	5.5	269.5	300.0	108.6	Yushen
17	Yuyang	1307	2.8	157.0	125.0	68.6	Yuheng
18	Yuyang	2301	3.6	143.5	130.0	70.0	Yuheng
19	Yuyang	2304	3.5	188.0	200.0	84.8	Yuheng
20	Yuyang	2304	3.5	208.0	200.0	96.3	Yuheng
21	Zhangjiamao	15,204	6.0	89.7	300.0	69.2	Shenfu
22	Zhangjiamao	N15203	5.5	164.7	295.2	108.0	Shenfu
23	Niingtiaota	N1114	5.5	188.9	250.0	145.2	Shenfu
24	Daliuta	52,306	6.8	180.0	301.0	137.3	Shenfu

Table 3 The significance (Sig.) of the coefficient of multiple linear regression model

Factor	Thickness of coal seam (m)	Buried depth of coal seam (m)	Inclined length of working face (m)	Constant
Sig	0.001	0.000	0.242	0.879
t	3.986	4.190	− 1.208	0.155

mean absolute percentage error (MAPE), standard error of estimate (SEE), root mean squared error (RMSE), and normalized root mean squared error (NRMSE). As shown in Table 7, the ME for the 24 points is − 2.72 m for Eq. 10. The height of the WCFZ is underestimated, but it is close enough for use in a safety analysis. The NRMSE is 13.42%, which is a little higher than the 10% value generally considered desirable, but for this geotechnical analysis with a significant range of geological conditions, is considered to acceptable. For Code Eqs. 1 and 2, the ME are − 50.00 m and − 36.46 m, respectively. They both seriously underestimated the actual WCFZ height, and both were unacceptable, with NRMSEs of 49.77% and 37.52%, respectively, which is far higher than 10%. That is unacceptable for use in safety analysis. Therefore, the height of the WCFZ predicted using the multiple regression model and information entropy was closer to the measured value than the heights predicted by the other empirical equations. It can be used to forecast the height of the WCFZ in this area.

The equations in the Code only consider the influence of the thickness of the coal seam on the height of the WCFZ. The current research has found that the height of the WCFZ is affected by various factors, such as the geological conditions of the mining area and coal mining technology (Chen and Zhu 2020; He et al. 2020). The data based on Code Eqs. 1 and 2 are derived from blast-winning technology and conventionally-mechanized coal winning technology in the 1950s and 1980s. Currently, coal mines are all fully mechanized. The influence of coal mining technology on the height of the WCFZ is different because the technology has changed. We calculated the weights of the factors influencing the WCFZ using information entropy and used the multiple regression model to obtain the fitting equation. Compared with the Code equations, the equation obtained in this paper for calculating the height of the WCFZ is more scientific and reasonable.

The factor of safety is very important for mine safety design. It was found that the predicted value was less than the measured value in 14 of the 24 cases in Eq. 10, so the probability of failure is 0.58. For engineering design of temporary structures, the acceptable probability of success is 0.95–0.99 (Harr 1987) and the corresponding probability of failure is 0.01–0.05. For mine safety analysis, a 95% probability of safety against failures is often used for design. To improve the safety of geotechnical structures, the standard method is to apply a factor of safety against failure, typically ≈ 1.3 for temporary underground structures in rock (Hoek et al. 1995). Through calculation, it was found that using a factor of safety of 1.33 reduced the risk of failure to 0.042, which meets the 95.8% reliability requirements.

Equation 10 only applies to predicting the height of the WCFZ in the Shandong, Shenfu, Yushen, and Yuheng mining areas in the northern Ordos Basin. Further work is needed to extend this work to other mining areas; data must be collected from mines with similar engineering and geological conditions, as was done in this study.

Case Study Based on GIS and Borehole TV

The Jinjie coal mine is located in the Yushen mining area of the northern Ordos Basin. The administrative division belongs to Shenmu City, Shaanxi Province, China. It is located at the northern end of the Loess Plateau in northern Shaanxi and the Mu Us Desert's southeastern edge. It covers an area of ≈ 141.78 km^2 and has an approved production capacity of 18 Mt/a. The study area is the 31,114 working face of the Jinjie coal mine, located in the middle of the first panel, and the working face is arranged in the north–south direction. The strike and inclined length of the 31,114 working face are 5257.5 m and 369 m, respectively. The coal seam inclines to the northwest with a dip angle of 1°.

The lithology in the studied area is composed of Quaternary Holocene (Q_4^{eol}), Middle Quaternary Lishi Formation (Q_{2l}), Middle Jurassic Zhiluo Formation (J_{2z}), and Middle Jurassic Yan'an Formation (J_{2y}). The weathered bedrock aquifer of the Zhiluo Formation is the main water inrush source in the Jinjie coal mine.

We predicted the height of the WCFZ in the study area with the fitting formula based on the multiple regression model and information entropy. Initially, using the study area's drilling exploration data, basic data such as the thickness of the coal seam, buried depth of the coal seam, and inclined length of the working face were statistically obtained. Then, the kriging method in the geostatistics analyst of ArcGIS was used to generate the basic layers of the coal seam thickness, the buried depth of the coal seam, and the inclined length of the working face. The raster calculator in the spatial analyst of ArcGIS was used to carry out the basic layer overlay operation on the multiple regression model improved with information entropy. Finally, as shown in Fig. 5, the map for the height prediction of the WCFZ of the study area was constructed.

The predicted height of the WCFZ in the study area was 49.20–54.54 m and tended to be low in the middle and

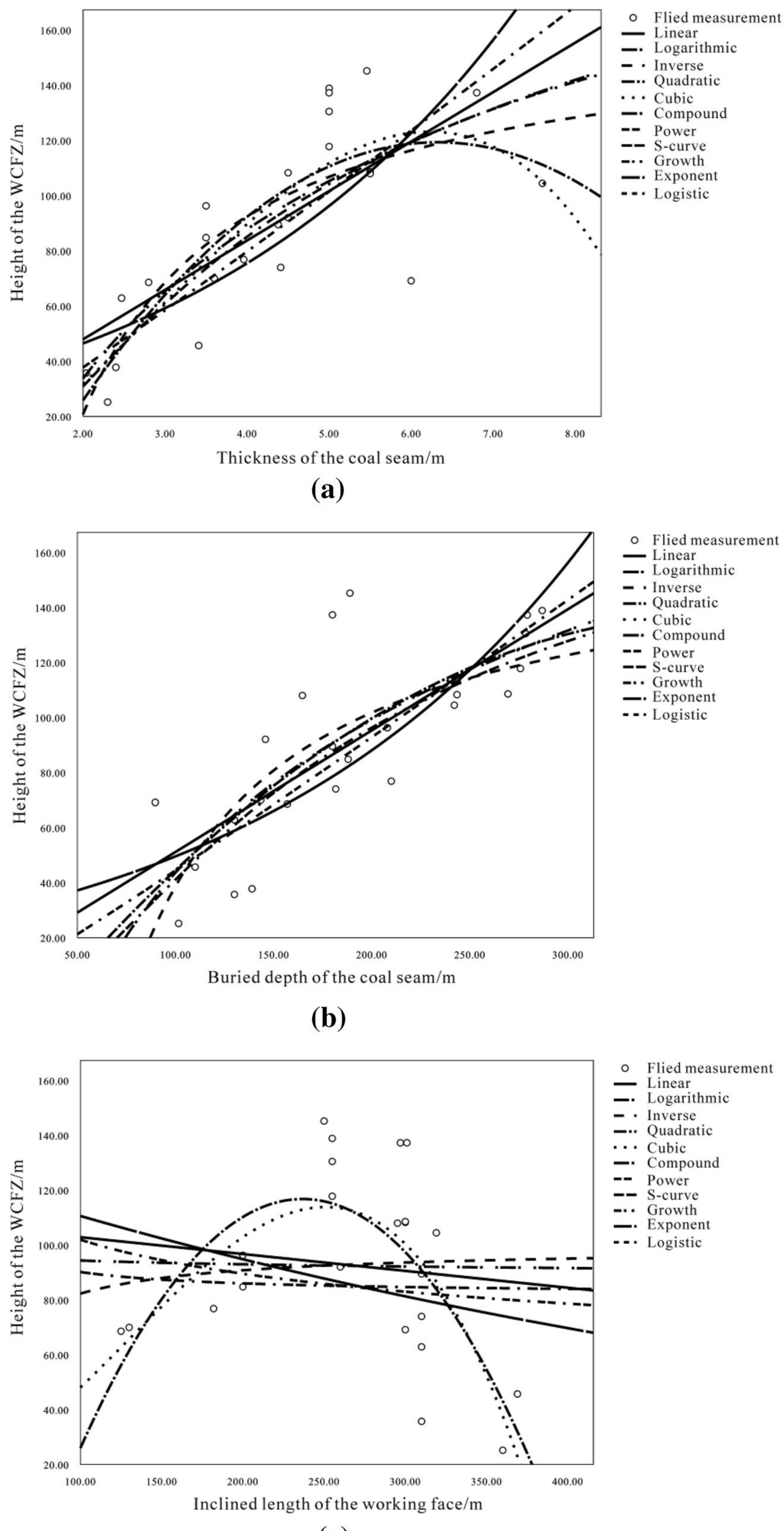

Fig. 3 Relationship between the height of the WCFZ and its influencing factors in nonlinear regression models

Table 4 R^2 and significance of nonlinear regression models

Model	Thickness of coal seam		Buried depth of coal seam		Inclined length of working face	
	R^2	Sig	R^2	Sig	R^2	Sig
Linear	0.546	0.000	0.595	0.000	0.014	0.594
Logarithmic	0.621	0.000	0.605	0.000	0.000	0.939
Inverse	0.639	0.000	0.571	0.000	0.005	0.750
Quadratic	0.672	0.000	0.609	0.000	0.478	0.002
Cubic	0.679	0.000	0.609	0.000	0.495	0.004
Power	0.661	0.000	0.591	0.000	0.014	0.585
Compound	0.558	0.000	0.561	0.000	0.049	0.309
S-curve	0.712	0.000	0.577	0.000	0.001	0.894
Logistic	0.558	0.000	0.561	0.000	0.049	0.309
Growth	0.558	0.000	0.561	0.000	0.049	0.309
Exponent	0.558	0.000	0.561	0.000	0.049	0.309

Table 5 Information entropy and weight coefficient of factors

Factor	Thickness of coal seam	Buried depth of coal seam	Inclined length of working face
Information entropy	0.9834	0.9841	0.9904
Weight coefficient	0.3938	0.3778	0.2284

high on both sides (Fig. 5). From the open-off cut to about 2240 m of mining, the maximum height was predicted to be 52–53 m, and then gradually decreasing to 49–50 m. There were two reasons for this: one is that the buried depth of the coal seam in this area was shallower, which also decreased the mine pressure slightly; the other is that the coal seam was thinner than earlier. Due to these factors, it is estimated that the development height of the WCFZ, which has been mined for about 4170 m, will gradually increase to 53–55 m.

To determine the development of the WCFZ in the study area, the height of the WCFZ was measured by borehole TV during coal mining at the 31,114 working face of the Jinjie mine from October to December 2019. A yygd-TG50 downhole drilling TV detection system directly displayed the development of the overburden rock fractures as a video recording (Lai et al. 2006; Xu et al. 2017). We used the measured maximum mining-induced fracture development height on October 21, 2019, when the working face was 143 m from the observation drillhole, as an example, and then calculated the maximum possible development height of the WCFZ.

The layout of the observation drillhole was carried out as mining progressed. As shown in Fig. 6, the field observation drillhole was at the 30th joint of the 31,115 ventilation roadway. The lithology of the observation drillhole section is shown in Fig. 7.

As the working face advance from November 21 to 24, the number of large mining-induced fractures from the roof of the coal seam significantly increased at a distance of 39.1–42.4 m. On December 12, 2019, it was 36.15 m from the drillhole, the opening and dislocation of the rock formation was very obvious, and the falling of broken rock blocks was very common. According to the observed data, it is clear that the developed height of the WCFZ at the observation drillhole is at least 42.4 m. Since the predicted value of the WCFZ using the multiple regression model and GIS method at this location was 49.6 m, this verified the applicability of the method in the northern Ordos Basin.

Conclusions

Based on the multiple regression model and information entropy, a method was proposed for predicting the height of the WCFZ in the northern Ordos Basin. Standard error analysis was used to determine the reliability of the predictive model and once the appropriate factor of safety of 1.33 was applied to the model, it met the 95.8% reliability requirements.

This method comprehensively considers the thickness of the coal seam, the buried depth of the coal seam, and the inclined length of the working face as influencing factors on the development height of the WCFZ in the northern Ordos Basin. Using the information entropy to calculate the weights of these three factors, we found that the thickness of the coal seam in this area has the greatest weight, the buried depth of the coal seam is second, and the inclined length of the working face has the least weight.

Taking the 31,114 working face of the Jinjie coal mine as a case study, the spatial characteristics of the WCFZ in

Table 6 Predictive height and measured height of the WCFZ

Coal mine	Working face	Measured height of the WCFZ/m	Code Eq. (1)		Code Eq. (2)		Equation (10)—paper	
			Calculated value/m	Relative error/%	Calculated value/m	Relative error/%	Predictive value/m	Relative error/%
Cuncaota	43,115	25.2	31.59	25.37	30.76	22.05	42.34	68.01
Wulanmulun	12,403	35.7	29.41	17.61	27.03	24.29	54.75	53.37
Wulanmulun	12,403	62.9	32.89	47.70	32.20	47.22	60.76	3.40
Huoluowan	22,101	37.8	32.26	14.66	31.98	15.39	41.00	8.47
Zhuanlongwan	23,103	92.1	41.67	54.76	55.90	39.30	88.00	4.45
Bulianta	12,406	74.0	41.35	44.12	54.84	25.90	88.30	19.32
Bulianta	12,406	89.5	41.35	53.80	54.84	38.73	88.62	0.98
Bulianta	12,511	104.5	48.22	53.85	85.82	17.88	115.62	10.64
Jinjie	31,104	45.7	37.61	17.70	43.79	4.18	53.34	16.71
Yushuwan	10,204	117.8	43.10	63.41	61.12	48.11	116.38	1.21
Yushuwan	10,204	130.5	43.10	66.97	61.12	53.16	116.90	10.42
Yushuwan	10,204	138.9	43.10	68.97	61.12	55.99	118.52	14.67
Yushuwan	10,203	137.3	43.10	68.61	61.12	55.48	113.39	17.41
Hanglaiwan	30,101	108.3	41.67	61.53	55.90	48.38	102.62	5.25
Longde	205	76.9	40.00	47.98	50.51	34.32	93.51	21.59
Jinjitan	101	108.6	44.35	59.16	66.19	39.06	114.17	5.13
Yuyang	1307	68.6	34.65	49.48	36.80	46.35	61.51	10.34
Yuyang	2301	70.0	38.46	45.05	46.06	34.20	68.31	2.42
Yuyang	2304	84.8	38.04	55.14	44.93	47.02	86.69	2.23
Yuyang	2304	96.3	38.04	60.49	44.93	53.34	90.48	6.04
Zhangjiamao	15,204	69.2	45.45	34.31	71.09	2.73	83.18	20.20
Zhangjiamao	N15203	108.0	44.35	58.93	66.19	38.72	94.89	12.14
Niingtiaota	N1114	145.2	44.35	69.45	66.19	54.42	103.04	29.04
Daliuta	52,306	137.3	46.96	65.80	78.63	42.73	103.53	24.60

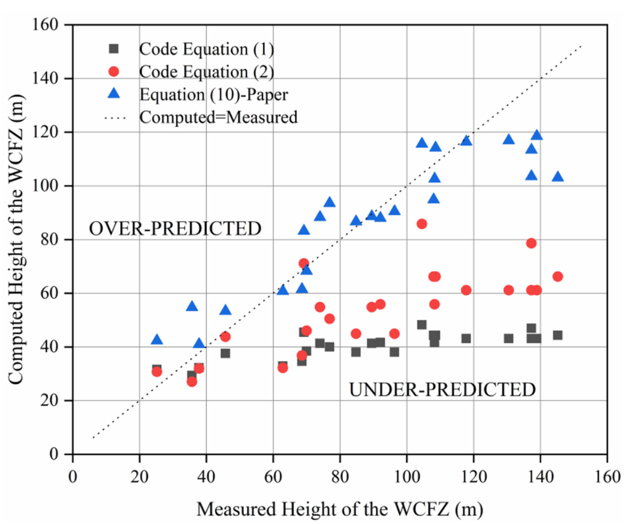

Fig. 4 The standard form of presentation for the error analysis of the predictive equation for the height of the WCFZ

Table 7 Error analysis of the predictive model for the height of the WCFZ

	Code Eq. (1)	Code Eq. (2)	Equation (10)-Paper
ME/m	− 50.00	− 36.46	− 2.72
MAE/m	50.53	37.08	11.93
MAPE/%	50.20	37.04	15.34
SEE/m	61.07	46.03	16.47
RMSE/m	59.72	45.02	16.10
NRMSE/%	49.77	37.52	13.42

the study area were examined using GIS technology; the predicted height of the WCFZ in the study area was determined to be 49.20–54.54 m, and the development height of the WCFZ was expected to be low in the middle and high

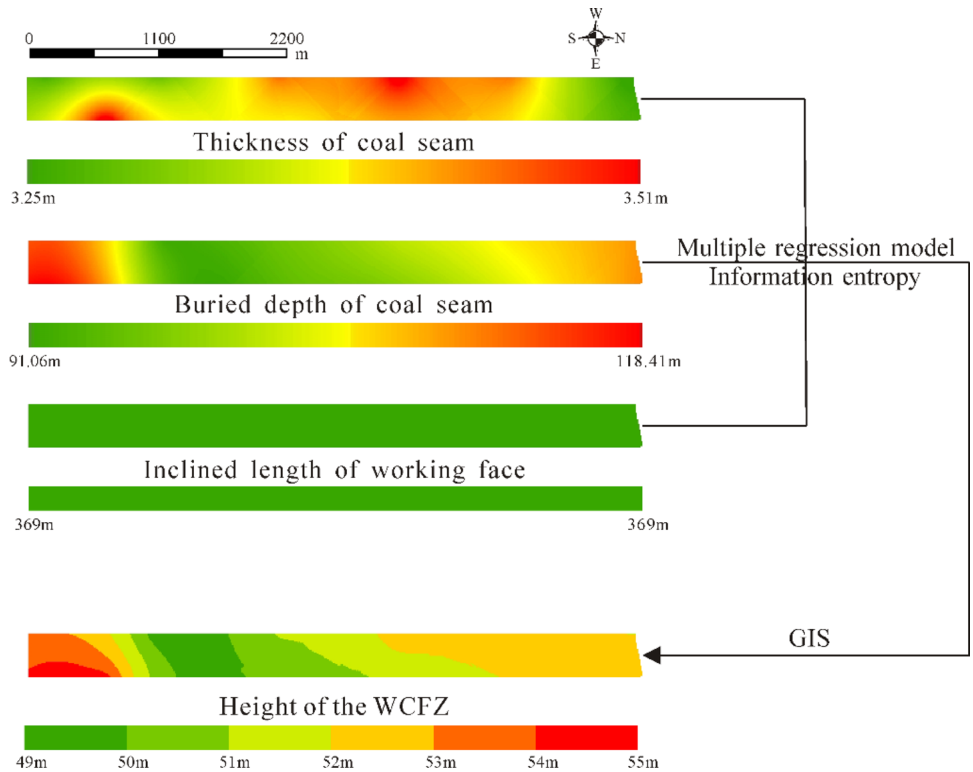

Fig. 5 Layer for the prediction of the height of the WCFZ

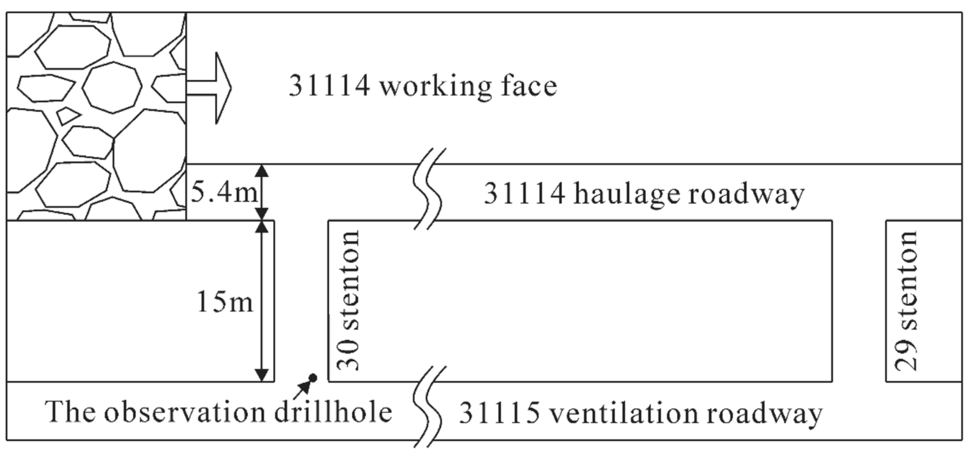

Fig. 6 Layout of the 31,114 working face and the observation drill hole

on both sides. The height of the WCFZ was measured with a borehole TV, and the measured value was 42.4 m, while the predicted value was 49.6 m, which verifies the applicability of the method in this area.

Thus, the proposed predictive method provides a basis for predicting mine water hazards control in in the northern Ordos Basin area and should greatly aid coal mining and the conservation of groundwater resources.

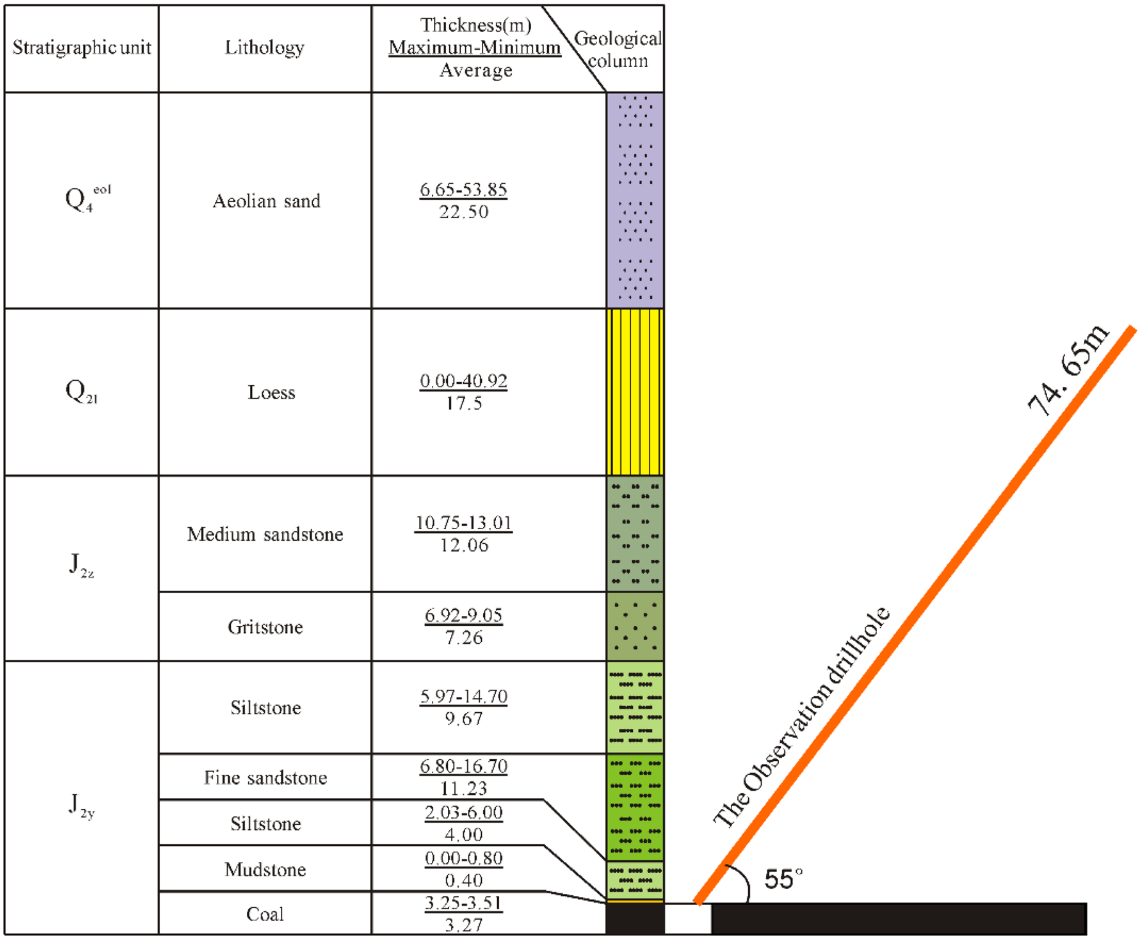

Fig. 7 Schematic diagram of the lithology of the observation drill hole section

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10230-021-00805-y.

Acknowledgements Financial support for this work is provided by the National Key R&D Program of China (2017YFC0804101); the National Natural Science Foundation of China (40802076); the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

Chen Y, Zhu SY (2020) Determination of caved and water-conducting fractured zones of "two soft and one hard" unstable coal seam. Acta Geod Geophys 55:451–475. https://doi.org/10.1007/s40328-020-00300-w

Gui HR, Lin ML (2016) Types of water hazards in China coalmines and regional characteristics. Nat Hazards 84:1501–1512. https://doi.org/10.1007/s11069-016-2488-5

Guo CF, Yang Z, Li S, Lou JF (2020) Predicting the water-conducting fracture zone (WCFZ) height using an MPGA-SVR approach. Sustainability 12(5):1809. https://doi.org/10.3390/su12051809

Harr ME (1987) Reliability-based design in civil engineering. McGraw-Hill, New York

He X, Zhao YX, Zhang C, Han PH (2020) A model to estimate the height of the water-conducting fracture zone for longwall panels in western China. Mine Water Environ 39:823–838. https://doi.org/10.1007/s10230-020-00726-2

Hoek E, Kaiser PK, Bawden WF (1995) Support of underground excavations in hard rock. Balkema, Rotterdam

Hu WY, Zhao CH (2021) Evolution of water hazard control technology in China's coal mines. Mine Water Environ. https://doi.org/10.1007/s10230-020-00744-0

Hu YB, Li WP, Liu SL, Wang QQ (2021) Prediction of floor failure depth in deep coal mines by regression analysis of the multi-factor influence index. Mine Water Environ. https://doi.org/10.1007/s10230-021-00769-z

Lai XP, Cai MF, Ren FH, Xie MW, Esaki T (2006) Assessment of rock mass characteristics and the excavation disturbed zone in the Lingxin Coal Mine beneath the Xitian river. China Int J Rock Mech Min 43(4):572–581. https://doi.org/10.1016/j.ijrmms.2005.10.005

Li PY (2018) Mine water problems and solutions in China. Mine Water Environ 37:217–221. https://doi.org/10.1007/s10230-018-0543-z

Li PY, Qian H, Wu JH, Chen J (2013) Sensitivity analysis of TOPSIS method in water quality assessment: I sensitivity to the parameter weights. Environ Monit Assess 185:2453–2461. https://doi.org/10.1007/s10661-012-2723-9

Liu TQ (1981) Surface movements, overburden failure and its application. China Coal Industry Press, Beijing

Liu SL, Li WP, Wang QQ, Pei YB (2018) Investigation on mining induced fractured zone height developed in different layers above Jurassic coal seam in western China. Arab J Geosci 11(2):30. https://doi.org/10.1007/s12517-018-3383-z

Liu Y, Liu QM, Li WP, Li T, He JH (2019a) Height of water-conducting fractured zone in coal mining in the soil–rock composite structure overburdens. Environ Earth Sci 78:242. https://doi.org/10.1007/s12665-019-8239-7

Liu Y, Yuan SC, Yang BB, Liu JW, Ye ZY (2019b) Predicting the height of the water-conducting fractured zone using multiple regression analysis and GIS. Environ Earth Sci 78:422. https://doi.org/10.1007/s12665-019-8429-3

Lv WH (2014) Measure and simulation for development height of water conducted crack zone in overburden roof. J Xi'an Univ Sci Technol 34(3):309–313. https://doi.org/10.13800/j.cnki.xakjdxxb.2014.03.011(inChinese)

Miao XX, Cui XM, Wang JA, Xu JL (2011) The height of fractured water-conducting zone in undermined rock strata. Eng Geol 120(1–4):32–39. https://doi.org/10.1016/j.enggeo.2011.03.009

Patricia AB (2006) Generating agricultural landscapes for alternative futures analysis: a multiple attribute decision-making model. Trans GIS 10(1):103–120. https://doi.org/10.1111/j.1467-9671.2006.00246.x

Peng SS, Li HM, Zhou YF (2015) Study of strata control in shendong and zhungeer mining areas. Science Press, Beijing (**in Chinese**)

State Administration of Work Safety (2017) Codes of coal mining and protective coal pillar design under buildings, water bodies, railways, coal mines and roadways. China Coal Industry Press, Beijing (**in Chinese**)

Sun Q, Zhang JX, Zhang Q, Zhao X (2017) Analysis and prevention of geo-environmental hazards with high-intensive coal mining: a case study in China's western eco-environment frangible area. Energies 10:786. https://doi.org/10.3390/en10060786

Wang D, Wu JH, Wang YH, Ji YJ (2020) Finding high-quality groundwater resources to reduce the hydatidosis incidence in the Shiqu county of Sichuan province, China: analysis, assessment, and management. Expo Health 12(2):307–322. https://doi.org/10.1007/s12403-019-00314-y

Wei JC, Wu FZ, Xie DL, Yin HY, Guo JB, Xiao LL, Zhi HF, Zhai SH (2016) Development characteristic of water flowing fractured zone under semi-cemented medium-low strength country rock. J China Coal Soc 41(4):974–983. https://doi.org/10.13225/j.cnki.jccs.2015.0750 (**in Chinese**)

Wu Q, Zhao SQ, Dong SN, Li JS (2013) Coal mine water prevention and control manual. China Coal Industry Press, Beijing (**in Chinese**)

Wu JH, Li PY, Qian H, Chen J (2015) On the sensitivity of entropy weight to sample statistics in assessing water quality: statistical analysis based on large stochastic samples. Environ Earth Sci 74(3):2185–2195. https://doi.org/10.1007/s12665-015-4208-y

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. https://doi.org/10.13225/j.cnki.jccs.2015.1210 (**in Chinese**)

Wu Q, Shen JJ, Liu WT, Wang Y (2017) A RBFNN-based method for the prediction of the developed height of a water-conductive fractured zone for fully mechanized mining with sublevel caving. Arab J Geosci 10:172. https://doi.org/10.1007/s12517-017-2959-3

Wu JH, Li PY, Wang D, Ren XF, Wei MJ (2020) Statistical and multivariate statistical techniques to trace the sources and affecting factors of groundwater pollution in a rapidly growing city on the Chinese Loess Plateau. Hum Ecol Risk Assess 26(6):1603–1621. https://doi.org/10.1080/10807039.2019.1594156

Xing ML, Li WP, Yin JH (2017) Study on height prediction of water flowing fractured zone caused by full-mechanized caving mining in Jurassic coalfield. Safety Coal Mines 48:39–42. https://doi.org/10.13347/j.cnki.mkaq.2017.09.011 (**in Chinese**)

Xu NW, Dai F, Li B, Zhu YG, Zhao T, Yang DS (2017) Comprehensive evaluation of excavation-damaged zones in the deep underground caverns of the Houziyan hydropower station, southwest China. Bull Eng Geol Environ 76:275–293. https://doi.org/10.1007/s10064-016-0858-6

Xue JK, Wang H, Zhao CH, Yang J, Zhou ZF, Li DB (2020) Prediction of the height of water-conducting fracture zone and water-filling model of roof aquifer in Jurassic coalfield in Ordos Basin. J Min Safety Eng 37(6):1341–1347. https://doi.org/10.13545/j.cnki.jmse.2020.06.017 (**in Chinese**)

Yang BB, Sui WH, Duan LH (2017) Risk assessment of water inrush in an underground coal mine based on GIS and fuzzy set theory. Mine Water Environ 36:617–627. https://doi.org/10.1007/s10230-017-0457-1

Yang QC, Li ZJ, Xie C, Liang J, Ma HY (2020a) Risk assessment of groundwater hydrochemistry for irrigation suitability in Ordos Basin, China. Nat Hazards 101:309–325. https://doi.org/10.1007/s11069-018-3451-4

Yang P, Yang WF, Zhang XQ, Wang ZR, Yang ML (2020b) Dynamic evolution of overlying strata stress induced by mining and mine water disaster identification based on information entropy. J China Coal Soc. https://doi.org/10.13225/j.cnki.jccs.2020.0706 (**in Chinese**)

Yang J, Dong SN, Wang H, Li GQ, Wang TT, Wang QM (2021) Mine water source discrimination based on hydrogeochemical characteristics in the northern Ordos Basin. Mine Water Environ, China. https://doi.org/10.1007/s10230-020-00723-5

Zhao DK, Wu Q (2018) An approach to predict the height of fractured water-conducting zone of coal roof strata using random forest regression. Sci Rep-UK 8:18986. https://doi.org/10.1038/s41598-018-29418-2

Zhu TE, Li WP, Wang QQ, Hu YB, Fan KF, Du JF (2020) Study on the height of the mining-induced water-conducting fracture zone under the Q_{2l} loess cover of the Jurassic coal seam in northern Shaanxi, China. Mine Water Environ 39:57–67. https://doi.org/10.1007/s10230-020-00656-z