Research Article

Piper-PCA-Fisher Recognition Model of Water Inrush Source: A Case Study of the Jiaozuo Mining Area

Pinghua Huang ^{1,2} and Xinyi Wang^{1,2}

¹School of Resources and Environment Engineering, Henan Polytechnic University, Henan, Jiaozuo 454000, China

²Collaborative Innovation Center of Coalbed Methane and Shale Gas for Central Plains Economic Region, Henan, Jiaozuo 454000, China

Correspondence should be addressed to Pinghua Huang; hph2001@hpu.edu.cn

Received 12 September 2017; Revised 14 January 2018; Accepted 30 January 2018; Published 26 February 2018

Academic Editor: Cinzia Federico

Copyright © 2018 Pinghua Huang and Xinyi Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Source discrimination of mine water plays an important role in guiding mine water prevention in mine water management. To accurately determine water inrush source from a mine in the Jiaozuo mining area, a Piper trilinear diagram based on hydrochemical experimental data of stratified underground water in the area was utilized to determine typical water samples. Additionally, principal component analysis (PCA) was used for dimensionality reduction of conventional hydrochemical variables, after which mutually independent variables were extracted. The Piper-PCA-Fisher water inrush source recognition model was established by combining the Piper trilinear diagram and Fisher discrimination theory. Screened typical samples were used to conduct back-discriminate verification of the model. Results showed that 28 typical water samples in different aquifers were determined through the Piper trilinear diagram as a water sample set for training. Before PCA was carried out, the first five factors covered 98.92% of the information quantity of the original data and could effectively represent the data information of the original samples. During the one-by-one rediscrimination process of 28 groups of training samples using the Piper-PCA-Fisher water inrush source model, 100% correct discrimination rate was achieved. During the prediction and discrimination process of 13 samples, one water sample was misdiscriminated; hence, the correct prediscrimination rate was 92.3%. Compared with the traditional Fisher water source recognition model, the Piper-PCA-Fisher water source recognition model established in this study had higher accuracy in both rediscrimination and prediscrimination processes. Thus it had a strong ability to discriminate water inrush sources.

1. Introduction

The threat posed to coal resources by mine water inrush accidents is prominent, usually causing local and even complete submergence of the mining area. Consequently, mine production efficiency will degrade or stagnate and bring about enormous economic loss. Jiaozuo mining area is a famous large water-filling mine in China and a typical North China-type coal field. Water damage caused by large-capacity and frequent water inrush occurs during the production process, resulting in submerging accidents and increasing the water drainage cost for coal output, as well as the production cost of per ton of coal. Water damage can also lead to major injuries, property loss, and loss of life. The main cause of water inrush in coal mines is rock fractured channels coming into

contact with aquifers and mining roadways. A large volume of groundwater from a coal seam floor or a roof aquifer rushes into the roadway.

Water-conducting passages and water-bearing structures buried before the roadway drivage face may easily lead to water bursts in coal mines. Therefore, rapid and accurate recognition of water inrush source has important practical significance in water prevention and control in coal mines.

Methods of water inrush source recognition used in mines in recent years include hydrochemical characteristic component analysis, isotope analysis, artificial tracing, multivariate statistical analysis, multiclass clustering functions, environmental isotope [1–12], and other relatively typical analysis methods. The following methods are applied mostly in multivariate statistical analysis: hierarchical clustering

linear discriminant method based on Fisher theory [13–15] and hydrochemical concentration forecast method based on artificial neural networks [16–18]. The distance discriminant analysis model was established to effectively predict the sources of mine water inrush at the base of the Jiaozuo mining area through measured data by selecting six discrimination factors [19]. With the progress of production systems, extensive experimental studies have been carried out on the application of some nonlinear mathematical statistics analysis methods in recent years. For instance, Pen et al. [20] utilized fuzzy comprehensive judgment method to carry out water inrush source recognition in mines. Yang et al. [21] carried out a corresponding research on water inrush grading of the Hebi Mining Bureau. Yan et al. [22] used support vector machine theory to analyze the water inrush source of mines. Xu et al. [23] applied neural network to perform practical work on water inrush sources, and Zhang et al. [24] made a corresponding application of quantification theory to discriminate the mine water inrush source.

Numerous scholars have carried out considerable research on water inrush source models and obtained great success in their practical application. However, present discriminant methods have not considered the complicated information superposition problem between hydrochemical data, a problem that results in misdiscrimination of the established model in the practical application process, and their discrimination accuracy still needs further improvement. Therefore, in this paper, a new discriminant method (Piper-PCA-Fisher) is presented, which extracts and compresses the information of hydrochemical data, transforms original data into mutually independent new data without information superposition, and combines with the mathematical classification method to establish the discriminating model of water inrush source. In this way, the high-accuracy water source recognition model can be trained out. The Piper-PCA-Fisher water inrush source recognition model has important practical value in the Jiaozuo mining area, and it can also provide theoretical guidance for the prevention and control of water damage in other North China coal mines.

2. Description of the Study Area

Jiaozuo mining area is in the northwestern part of Henan Province, north of the Taihang Mountains and south of the Yellow River. This area has a temperate continental monsoon climate with four distinctive seasons and an annual average temperature of approximately 13.2°C [1]. Influenced by meteorological conditions and topographic factors, the annual average rainfall in the northern mountainous area is about 700 mm, while that in the southern piedmont plain is about 600 mm. Rainfall mainly occurs in summer and autumn. The ground level of the northwestern mountainous area is 200–1790 m. The ground has elevations and depressions. The ground level of the piedmont alluvial plain in the southeast is 80–200 m. This topographic feature forms a noticeable control effect on underground water runoff.

The overall structural feature of strata in Jiaozuo mining area is monoclinic morphology with a tendency toward the southeast and an orientation at the southwest. Dip angle is 6°–16° and that of the local part is 25°–30°. This area has faulted structures, mostly normal faults with few folds. The principal fracture is a strong runoff zone of underground karst water with mature karstic development that intersects with a secondary fault to form a horizontally and longitudinally staggered stereoscopic fracture network. As a result, underground water in this area has close connections in both vertical and horizontal directions and hydraulic connections of different degrees between multiple aquifers to form complicated hydrogeological conditions (Figure 1).

From top to bottom, the main aquifers can be divided into four categories according to the lithology, thickness, water features, and burial conditions of the stratum. The first category is the Quaternary aquifer made up of Quaternary sandstone, clay, calcareous nodules, and the conglomerate bottom, which is the main aquifer for the segment. The second category is the coal-bearing sandstone aquifer that consists of sandstone, siltstone, and shale layers with a permeability (K) of 0.1–0.3 m/d. The third category is the carboniferous limestone aquifer composed of Taiyuan limestone. The fourth category is the Ordovician limestone aquifer that is the sedimentary basement of the coal measure strata. Jiaozuo coal-mining area mainly contains Ordovician limestone aquifer groundwater (type I), carboniferous limestone aquifer groundwater (type II), coal sandstone aquifer groundwater (type III), and Quaternary system aquifer groundwater (type IV).

3. Material and Methods

3.1. Data Analysis of Modeling Experiment. Water samples collected in the study area included underground water at different aquifers in the coal-mining area (Figure 1). 38 groups of water samples were collected from the coal-mining district, including 8 groups from type I, 10 groups from type II, 9 groups from type III, and 11 groups from type IV. Water samples were collected in clean 550 ml plastic bottles to determine hydrochemical ions. 13 groups of data (numbers A1–A13) in this paper were quoted from Wang et al. [19] and were used to test the model. Table 1 shows the results of water sample analysis. The chemical composition of the water samples was determined in the State Key Laboratory of Hydrology, Henan Polytechnic University, using a Shimadzu CTO-10Avp ion chromatograph and ICP-MS with a relative error of 1%, while HCO_3^- was determined using dilute sulfuric acid-methyl orange titration.

3.2. Piper Trilinear Diagram Analysis. Underground hydrochemical components at aquifers change with the movement of groundwater. Different aquifers always have different hydrochemical features, and hydrochemical components at the same aquifer will also be noticeably different because of varying hydrogeological conditions. However, their chemical components will maintain a dynamic equilibrium through a series of physical and chemical reactions. Therefore, water

FIGURE 1: Hydrogeological map of the Jiaozuo coal-mining district in China.

samples at the same aquifer will present the same hydrochemical features after hydrochemical analysis. During the practical analysis process, hydrochemical components at the same aquifer sometimes vary greatly because samples are influenced by the hydraulic connection between different aquifers, the movement of underground water, and so on. Hence, we must determine typical water samples and find the water sample that can best represent hydrochemical components at this aquifer to establish a high-accuracy water intrusion source discrimination model. The Piper trilinear diagram of water samples can be used to determine typical water samples. If water samples significantly deviate from the formation center in the Piper trilinear diagram, these samples should be discriminated as abnormal and excluded in this study.

3.3. Principal Component Analysis. Principal component analysis (PCA) is a multivariate statistical analysis method whose basic idea is data dimension reduction. Multiple observational variables with the original existence of information superposition are transformed into several mutually

irrelevant aggregate variables through orthogonal transformation to extract feature information. Multiple correlation variables must also be simplified with the least information loss possible [25].

3.4. Fisher Mathematical Principle. The basic idea of Fisher discriminant analysis is projection. Specifically, it projects high-dimensional points onto low-dimensional space and uses univariate analysis of variance to establish a linear discriminant function per criteria of maximum between-class distance and minimum inner-class distance. Sample class can be discriminated per corresponding criterion. Fisher discriminant analysis can be used to skillfully avoid the “curse of dimension” and solve high-dimensional problems using a 1D method.

3.4.1. Solving Method of Fisher Discriminant Analysis. Suppose there are m totalities $G_1, G_2, \dots, G_m$, corresponding mean values, and correlation coefficient matrixes $\mu^{(1)}, \mu^{(2)}, \dots, \mu^{(m)}; V^{(1)}, V^{(2)}, \dots, V^{(m)}$, respectively.

TABLE 1: Conventional hydrochemical ions in aquifers of Jiaozuo mining and samples to be measured (mg/L).

No.	Aquifer	Na ⁺ + K ⁺	Mg ²⁺	Ca ²⁺	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻
1	I	24.2	25.3	74	284.95	18.84	44
2	I	97.6	9.2	6.5	42.4	7	145
3	I	27.1	24.9	64.5	290.2	7	25
4	I	18.3	15.3	59.1	212.2	5	33
5	I	22.2	24	65.6	264.3	19.95	34.4
6	I	15.71	26.14	81.96	286.67	25.81	63.86
7	I	95.6	8.2	6.2	42.3	8	155
8	I	15.7	26.02	78.6	309.27	16.18	65.52
9	II	13.3	14.6	67	185.5	7.5	74.5
10	II	172.3	2.2	2.9	367.7	11	19
11	II	15.4	26.3	61.7	260.9	7.5	35.5
12	II	4.2	27.8	62	295.4	9	22.5
13	II	4.7	27.1	68.6	255.2	6.5	43.5
14	II	58	30.5	145.4	319.7	69.5	217
15	II	12.13	24.6	69.3	299.51	9.09	33.4
16	II	15.29	29.33	61.23	309.85	12.16	47.46
17	II	13.3	26.9	63.5	282.52	11.19	43.85
18	II	13.9	23.9	67.1	280.69	11.19	41.9
19	III	2.8	46.7	98.2	495.3	12.5	3
20	III	303.6	0.5	4.3	685.2	15	1.5
21	III	157.5	13.8	54.1	535.7	27.5	2
22	III	205.2	7	41.1	659.4	61	133
23	III	176.9	28.7	85	697.31	25.19	15.4
24	III	228.1	1.1	3.1	638.7	43.84	23.5
25	III	206.3	3	26.29	316.25	11.67	38.58
26	III	202.1	6.88	17.23	239.1	14.11	6.49
27	III	206.6	2.07	39.1	505.25	27.24	0
28	IV	266.6	6.1	24.9	503.9	26.5	69.5
29	IV	261	7.1	23	500	27.5	61.2
30	IV	60.3	21.4	54.5	193.8	33	85
31	IV	55.7	27.1	74.3	302.56	15.4	51.1
32	IV	55	25.5	72.7	294.71	15.4	47.3
33	IV	55.1	31.8	86.5	384.31	22.4	57.8
34	IV	59.25	31.1	99.2	361.12	29.85	83
35	IV	55.2	39.1	106.7	402.1	40.1	69.8
36	IV	62.6	21.8	56.3	196.78	24.6	74.2
37	IV	59.3	16.5	76.4	299.51	12.59	28.09
38	IV	256.1	5.3	23.4	507.3	24.1	68.2
A1	I	34.3	23.4	64.9	301.8	9.5	36.5
A2	I	30.3	16.2	72.5	248.7	16	52.5
A3	II	12.1	24.7	63	266.13	7.35	37.8
A4	II	12.7	39	67.1	281.57	8.82	46.5
A5	II	12	24.9	68.72	282.16	11.7	43.77
A6	II	7.4	24.8	55.2	256.6	5	19.5
A7	II	17.6	23.9	66.7	285.8	6.5	27.5
A8	III	263.3	2.2	4.7	630.3	8	2.5
A9	III	170.2	16.5	49.3	319.74	10.87	36.5
A10	III	177.3	18.4	52	366.78	56.06	40.41

TABLE 1: Continued.

No.	Aquifer	Na ⁺ + K ⁺	Mg ²⁺	Ca ²⁺	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻
A11	IV	54.61	35.3	90.5	361.7	16.04	86.8
A12	IV	63.3	20.6	98.2	354.4	20.24	53.2
A13	IV	62.2	38.7	84.21	280.49	56.6	119.74

Note. Ordovician limestone aquifer groundwater (simplified as type I), carboniferous limestone aquifer groundwater (simplified as type II), coal sandstone aquifer groundwater (simplified as type III), and Quaternary system aquifer groundwater (simplified as type IV). The water samples (numbers 1–38) are used to rebuild the model. The data (numbers A1–A13) in Table 1 are not used to build the model, which is only used to validate the model.

Samples with capacity n_i are extracted from totality G_i as follows:

$$X_{\alpha}^{(i)} = (X_{\alpha 1}^{(i)}, X_{\alpha 2}^{(i)}, \dots, X_{\alpha p}^{(i)})^T. \quad (1)$$

Then, $u^T X_{\alpha}^{(i)} = (u_1 X_{\alpha 1}^{(i)}, u_2 X_{\alpha 2}^{(i)}, \dots, u_p X_{\alpha p}^{(i)})$ is a projection of $X_{\alpha}^{(i)}$ on the axis, where

$$\alpha = 1, 2, \dots, n_i; \quad i = 1, 2, \dots, m; \quad (2)$$

vector $u = (u_1, u_2, \dots, u_p)^T$ represents one direction in p -dimensional space; and $Y = u^T X$ is the inner product of u and X .

The projection of X on the u axis is

$$\begin{aligned} \bar{X}^{(i)} &= \frac{1}{n_i} \sum_{\alpha=1}^{n_i} X_{\alpha}^{(i)}; \\ \bar{X} &= \frac{1}{n} \sum_{i=1}^m \sum_{\alpha=1}^{n_i} X_{\alpha}^{(i)}, \end{aligned} \quad (3)$$

where $n = \sum_{i=1}^m n_i$ and $\bar{X}^{(i)}$ and $\bar{X}$ are the sample mean and total sample mean, respectively. Therefore, the inner-group difference is as follows:

$$\begin{aligned} e &= \sum_{i=1}^m \sum_{\alpha=1}^{n_i} (u^T X_{\alpha}^{(i)} - u^T \bar{X}^{(i)})^2 \\ &= u^T \left\{ \sum_{i=1}^m \left[\sum_{\alpha=1}^{n_i} (X_{\alpha}^{(i)} - \bar{X}^{(i)}) (X_{\alpha}^{(i)} - \bar{X}^{(i)})^T \right] \right\} u \\ &= u^T \left(\sum_{i=1}^m S_i \right) u = u^T W u, \end{aligned} \quad (4)$$

where S_i is the sample dispersion matrix of n_i samples $X_{\alpha}^{(i)}$ ($\alpha = 1, 2, \dots, n_i$) in G_i , and the intergroup difference is

$$\begin{aligned} b &= \sum_{i=1}^m \sum_{\alpha=1}^{n_i} (u^T \bar{X}^{(i)} - u^T \bar{X})^2 \\ &= u^T \left[\sum_{i=1}^m (\bar{X}^{(i)} - \bar{X}) (\bar{X}^{(i)} - \bar{X})^T \right] u = u^T B u. \end{aligned} \quad (5)$$

The equation is $\Phi = b/e = u^T B u / u^T W u$. To make the maximum Φ and make the solution unique, $u^T W u = 1$ is set.

Therefore, the problem is transformed into solving u , which causes $u^T B u$ to reach the maximum under $u^T W u = 1$. The Lagrange multiplier method is used, and the following is set:

$$F = u^T B u - \lambda (u^T W u - 1). \quad (6)$$

The partial differential of the above equation is solved and set as 0; that is,

$$\frac{\partial F}{\partial u} = 2B u - 2\lambda W u = 0. \quad (7)$$

Through further arrangement, the following equation is obtained:

$$(W^{-1} B - \lambda I) u = 0. \quad (8)$$

The equation shows that λ should be the maximum eigenvalue of $W^{-1} B$, and u is the eigenvector corresponding to $\lambda_{\max}$. Hence, the Fisher discriminant function can be solved.

3.4.2. Verification of Fisher Discrimination Effect. To ascertain whether the above criterion was suitable, back-substitution estimation method was used for rediscrimination to estimate the misdiscrimination rate. For training samples with capacity n_i from totality G_i (where $\alpha = 1, 2, \dots, n_i$; $I = 1, 2, \dots, m$), all training samples were successively substituted into the established discriminant function and the corresponding criterion was used for water source recognition. Total misdiscrimination number was N , and the misdiscrimination rate η through back-substitution estimation was as follows:

$$\eta = \frac{N}{n_1 + n_2 + \dots + n_m}. \quad (9)$$

4. Results and Discussion

4.1. Screening of Typical Water Samples. The Piper trilinear diagram (Figure 2) of Ordovician water samples showed that water samples numbers 2 and 7 significantly deviated from the formation center. Hence, these samples were discriminated as abnormal and excluded in this study. The remaining six groups were screened out as typical water samples in the Ordovician aquifer. The cation content in the remaining six groups of water samples, such as Ca^{2+} , Mg^{2+} , and Na^+ , was comparatively stable. Anion content was

FIGURE 2: Piper trilinear diagram of the water samples from aquifer I, II, III, IV.

relatively large, except for the variation range of HCO_3^- . The variation ranges of both Cl^- and SO_4^{2-} were small. The main water sample types of aquifer formation were Ca-Mg-HCO_3 , $\text{Ca-Mg-HCO}_3\text{-SO}_4$, and Na-SO_4 because the Ordovician limestone aquifer runoffs had favorable discharge conditions. Influenced by the dissolution filtration effect, carbonate-type ores centering on calcite and dolomite formed a hydrochemical type with cations, mainly Ca^{2+} and Mg^{2+} , and anions, mainly HCO_3^- . Meanwhile, influenced by iron pyrite (Fe_2S) at the bottom of the coal strata in this area, oxidation caused the water pH to decline and SO_4^{2-} ions to subsequently enter the Ordovician limestone aquifer and increase the SO_4^{2-} content of ions.

Piper trilinear diagram of Archaean limestone water samples (Figure 2) showed that Ca-Mg-HCO_3 , which belongs to typical carbonatite karst water, is the main hydrochemical type in carboniferous Archaean aquifer formation. In addition, water samples numbers 9, 10, and 14 were far from the formation center and clearly beyond the Ordovician aquifer formation. Thus, these samples were excluded, and

the remaining seven groups were screened out as typical water samples of Ordovician aquifer.

Water samples numbers 28, 29, and 38 of the Quaternary aquifer (Figure 2) were very far from the formation center and were thus excluded. The remaining eight groups were regarded as typical water samples of the Quaternary aquifer.

Piper trilinear diagram analysis of water samples in the sandstone aquifer (Figure 2) showed that the main hydrochemical types of water samples in this formation were Na-Ca-HCO_3 , Na-HCO_3 , Ca-Mg-HCO_3 , and $\text{Na-HCO}_3\text{-SO}_4$. The sample number 19 was excluded. Clay content was high and underground water flowed slowly, which contributed to the sufficient exchange and accumulation of Na^+ in the Permian sandstone aquifer formation. Hydrochemical types centering on Na-HCO_3 were formed in underground water. The distributions of Na^+ , K^+ , Ca^{2+} , and HCO_3^- were uniform. Na^+ and K^+ ions were noticeably higher than those in other aquifer formations, and the distributions of Mg^{2+} , Cl^- , and SO_4^{2-} ions were concentrated.

Therefore, 28 typical water samples (1, 3, 4, 5, 6, 8, 11, 12, 13, 15, 16, 17, 18, 20, 21, 23, 24, 25, 26, 27, 30, 31, 32, 33, 34, 35, 36, and 37) in Table 1 were finally confirmed as water samples for training the water inrush source recognition model in the next step.

Considering the difference and data effectiveness of components for different water sources, six components, namely, $K^+ + Na^+$ (X_1), Mg^{2+} (X_2), Ca^{2+} (X_3), HCO_3^- (X_4), Cl^- (X_5), and SO_4^{2-} (X_6), were selected and compressed into five variables, namely, Y_1 , Y_2 , Y_3 , Y_4 , and Y_5 , through PCA. These variables were then used as indexes of the water inrush source recognition model of the mine to establish the Fisher water source recognition model.

4.2. PCA Treatment of Data. PCA was carried out for the sample data in Table 1. The six water source components had definite correlations, where the correlation coefficient between Mg^{2+} (X_2) and Ca^{2+} (X_3) was 0.901, and a noticeable information superposition existed between data. PCA treatment of sample data was necessary to establish the accuracy of the water inrush source discrimination model. The first five factors covered most of the information quantity of the original data and occupied about 98.92%. Therefore, the five given principal components could effectively represent the original data information of samples.

According to the PCA matrix, relational expressions of the five extracted principal components Y_1 , Y_2 , Y_3 , Y_4 , and Y_5 with the original variables X_1 , X_2 , X_3 , X_4 , X_5 , and X_6 were obtained as follows:

$$\begin{aligned}
 Y_1 &= -0.935X_1 + 0.898X_2 + 0.847X_3 - 0.701X_4 \\
 &\quad - 0.201X_5 + 0.753X_6, \\
 Y_2 &= 0.220X_1 + 0.222X_2 + 0.307X_3 + 0.501X_4 \\
 &\quad + 0.920X_5 + 0.375X_6, \\
 Y_3 &= -0.048X_1 + 0.313X_2 + 0.359X_3 + 0.463X_4 \\
 &\quad - 0.244X_5 - 0.472X_6, \\
 Y_4 &= 0.177X_1 - 0.038X_2 - 0.011X_3 + 0.143X_4 \\
 &\quad - 0.231X_5 + 0.258X_6, \\
 Y_5 &= 0.154X_1 - 0.132X_2 - 0.236X_3 + 0.111X_4 \\
 &\quad - 0.013X_5 + 0.026X_6.
 \end{aligned} \tag{10}$$

4.3. Fisher Discriminant Analysis

(a) Piper-PCA-Fisher Model. The data of principal components Y_1 , Y_2 , Y_3 , Y_4 , and Y_5 obtained through PCA method were taken as input variables of the Fisher discriminant analysis model to make calculation for the Fisher discriminant analysis. The following Fisher discriminant functions were solved:

TABLE 2: Central values of the discriminant function in each category.

Classification	I	II	III	IV
F1	-3.126	-4.828	7.812	-0.266
F2	-0.130	-1.178	-0.704	1.744
F3	-0.362	0.198	0.002	0.097

The first discriminant function:

$$F_1 = -0.027Y_1 + 0.065Y_2 - 0.006Y_3 - 0.137Y_4 + 0.348Y_5 - 5.832. \tag{11}$$

The second discriminant function:

$$F_2 = -0.013Y_1 + 0.049Y_2 - 0.013Y_3 - 0.006Y_4 + 0.150Y_5 - 4.153. \tag{12}$$

The third discriminant function:

$$F_3 = -0.020Y_1 + 0.055Y_2 - 0.007Y_3 - 0.270Y_4 + 0.079Y_5 - 3.163. \tag{13}$$

Table 2 shows the central values of the first, second, and third discriminant functions in the distribution of water sources. Taking the first discriminant function as an example, the central values of the type I, II, III, and IV water sources were -3.126, -4.828, 7.812, and -0.266, respectively. Water source recognition was implemented by comparing distances from functional values of the water samples to be discriminated against the central values of the distribution of water sources.

(b) Fisher Model. The data of hydrochemical component were taken as input variables of the Fisher discriminant analysis model to make calculation for the Fisher discriminant analysis. The solved discriminant functions were as follows:

$$\begin{aligned}
 f_1 &= 0.092Z_1 + 0.505Z_2 + 0.078Z_3 + 0.09Z_4 \\
 &\quad - 0.272Z_5 + 0.079Z_6 - 12.311, \\
 f_2 &= 0.092Z_1 + 0.489Z_2 + 0.094Z_3 + 0.014Z_4 \\
 &\quad - 0.274Z_5 + 0.073Z_6 - 13.676, \\
 f_3 &= 0.165Z_1 + 0.742Z_2 + 0.096Z_3 + 0.004Z_4 \\
 &\quad - 0.089Z_5 - 0.001Z_6 - 23.318, \\
 f_4 &= 0.169Z_1 + 0.814Z_2 + 0.113Z_3 - 0.006Z_4 \\
 &\quad - 0.148Z_5 + 0.027Z_6 - 21.087,
 \end{aligned} \tag{14}$$

where f_1 , f_2 , f_3 , and f_4 were the respective Fisher discriminant functions of types I, II, III, and IV; Z_1 , Z_2 , Z_3 , Z_4 , Z_5 , and Z_6 represent the contents of $K^+ + Na^+$, Mg^{2+} , Ca^{2+} , HCO_3^- , Cl^- , and SO_4^{2-} , respectively; and the final item of discriminant function was a constant.

TABLE 3: Classification result of water inrush source discriminant model.

No.	Y1	Y2	Y3	Y4	Y5	Piper-PCA-Fisher
A1	-142.07	206.29	149.16	56.63	-17.06	I
A2	-90.40	191.53	116.11	50.59	-9.54	I
A3	-95.34	181.75	133.35	48.50	-17.15	II
A4	-84.16	198.67	141.95	53.22	-19.93	II
A5	-97.84	197.80	139.01	51.25	-17.83	II
A6	-104.09	164.55	135.61	42.21	-18.16	II
A7	-119.44	189.13	148.34	49.75	-17.23	II
A8	-681.79	383.94	278.43	135.57	-28.77	III
A9	-301.40	240.12	142.85	82.84	-0.91	III
A10	-343.16	309.54	152.98	81.43	-5.34	III
A11	-134.12	276.15	163.50	80.42	-17.51	I
A12	63.3	20.6	98.2	354.4	20.24	IV
A13	62.2	38.7	84.21	280.49	56.6	IV

4.4. Verification of Water Inrush Source Recognition Model. To verify the accuracy of the established PCA-based Fisher water source recognition model, 28 groups of training samples in Table 1 were rediscriminated one by one. The results showed that all water samples were discriminated correctly, with a discrimination rate of 100%. Meanwhile, compared with the Fisher water source recognition model established without data processing through PCA, the traditional Fisher water source recognition model incurred multiple errors in its rediscrimination steps; the correct discrimination rate was less than 90%. Therefore, the Fisher water source recognition model based on PCA was more accurate, had a higher degree of stability, and could meet the actual requirements of water inrush water source recognition.

In practical application, 13 water samples to be discriminated in the Jiaozuo mining area were substituted into the trained Piper-PCA-Fisher water source recognition model for discrimination (Table 3). Except for Quaternary water sample number 11, which was misdiscriminated as Ordovician water, the prediscrimination results showed that the prediction results of other water samples complied with actual classification, and prediscrimination success rate was 92.3%. However, the traditional Fisher water source recognition model misdiscriminated repeatedly for water samples, resulting in a prediscrimination success rate of less than 80%. Through comprehensive comparison, the Piper-PCA-Fisher water source recognition model was more accurate and had more extensive applicability than the other models.

4.5. Application of Water Inrush Source Recognition Model. The identification method of water inrush source can be used in coal mines and can be carried out according to the following steps.

Step 1. Collect water sample data and aquifer information from the coal mine and select the representative water samples based on the Piper diagram.

Step 2. Analyze the hydrochemical ion concentration of the representative water samples through PCA and obtain the key principal component data.

Step 3. Use the Fisher classification model to train the key principal component data and establish the water inrush source recognition model of the mine.

Step 4. Test the water sample using the established water inrush source recognition model.

Step 5. Present the identification results of the water inrush source.

During the application process of the Piper-PCA-Fisher water source recognition model in the Jiaozuo mining area, as proposed in this study, misdiscrimination only appeared in one water sample from the Quaternary water source. The prediscrimination success rate was 92.3%, which indicated that the established water source recognition model was successful. The misdiscrimination of the water sample was primarily caused by the hydraulic connection between the Quaternary and Ordovician aquifers in the mining area.

During the model establishing process, the whole mining area was taken as the study object. During the application process, it might be impossible to implement accurate water inrush water source recognition for individual mines. Hydrogeological conditions and perfect underground hydrochemical database in aquifers must be sufficiently studied to apply the model to a single mine.

In this study, discrimination of mine water inrush water sources was based on finite data and was affected by data randomness, representativeness, and accuracy. Thus, we must extensively collect measured data, establish a corresponding training sample database, and enhance the applicability of this model.

5. Conclusion

Stratified sampling and experimental analysis of water quality were carried out based on the hydrogeological conditions of the mining area. The Piper trilinear diagram was then utilized to analyze and extract typical water samples. Finally, the

Fisher water source recognition model based on PCA method was established. The following conclusions were drawn:

(1) The overall hydrochemical features between aquifers were significantly different. The difference of hydrochemical type between individual water samples at the same aquifer was noticeable, and thus exclusion was needed.

(2) Compared with the traditional Fisher water source recognition model, the Piper-PCA-Fisher water source recognition model established in this work had higher accuracy in both rediscrimination and prediscrimination processes.

(3) The Piper-PCA-Fisher water source recognition model could meet the requirements of modern water inrush waterproof development.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (Grant 41672240), Science and Technology Key Research Project of the Education Department of Henan, China (nos. 13A170313, 14A510022), China Postdoctoral Science Foundation (2017M612395), Innovation Scientists and Technicians Troop Construction Projects of Henan Province (Grant CXTD2016053), Henan Province's Technological Innovation Team of Colleges and Universities (Grant 15IRTSTHN027), Fundamental Research Funds for the Universities of Henan Province (NSFRF1611), Scientists and Technicians Projects of Henan Province (Grant 172107000004).

References

- [1] P.-H. Huang, J.-S. Chen, and C. Ning, "The analysis of hydrogen and oxygen isotopes in the ground water of Jiaozuo mine area," *Journal of the China Coal Society*, vol. 37, no. 5, pp. 770–775, 2012.
- [2] G. Y. Pan, S. N. Wang, and X. Y. Sun, "Application of isotopic technique in identification of mine water inrush source," *Mining Safety & Environmental Protection*, vol. 2008, no. 4, pp. 7–9, 2009.
- [3] Z. Chen, D. Li, and F. Jiang, "The prediction model of ground water inrush from floor in jiaozuo coal mine," *Coal Geology & Exploration*, vol. 4, pp. 38–40, 1996.
- [4] X. Xu, B. Guo, and G. Z. Wang, "Application of artificial neural network for recognition of multiple water sources in mine," *Journal of Safety Science and Technology*, vol. 12, no. 1, pp. 181–185, 2016.
- [5] M. Qiu, L. Shi, C. Teng, and Y. Zhou, "Assessment of water inrush risk using the fuzzy delphi analytic hierarchy process and grey relational analysis in the liangzhuang Coal Mine, China," *Mine Water and the Environment*, vol. 36, no. 1, pp. 39–50, 2017.
- [6] T. E. Keskin, M. Düğenci, and F. Kaçaroğlu, "Prediction of water pollution sources using artificial neural networks in the study areas of Sivas, Karabük and Bartın (Turkey)," *Environmental Earth Sciences*, vol. 73, no. 9, pp. 5333–5347, 2015.
- [7] G. Li, Z. Meng, X. Wang, and J. Yang, "Hydrochemical prediction of mine water inrush at the Xinli Mine, China," *Mine Water and the Environment*, vol. 36, no. 1, pp. 78–86, 2017.
- [8] Y. Wang, M.-R. Zhou, P.-C. Yan, C.-Y. He, and D. Liu, "Identification of coalmine water inrush source with pca-bp model based on laser-induced fluorescence technology," *Spectroscopy and Spectral Analysis*, vol. 37, no. 3, pp. 978–983, 2017.
- [9] J. Wu, S. Xu, R. Zhou, and Y. Qin, "Scenario analysis of mine water inrush hazard using Bayesian networks," *Safety Science*, vol. 89, pp. 231–239, 2016.
- [10] Q. Wu, Y. Liu, X. Wu, S. Liu, W. Sun, and Y. Zeng, "Assessment of groundwater inrush from underlying aquifers in Tunbai coal mine, Shanxi province, China," *Environmental Earth Sciences*, vol. 75, no. 9, article no. 737, 2016.
- [11] P.-C. Yan, M.-R. Zhou, Q.-M. Liu, R. Wang, and J. Liu, "Research on the source identification of mine water inrush based on LIF technology and PLS-DA algorithm," *Spectroscopy and Spectral Analysis*, vol. 36, no. 9, pp. 2858–2862, 2016.
- [12] J. Qian, L. Wang, L. Ma, Y. Lu, W. Zhao, and Y. Zhang, "Multivariate statistical analysis of water chemistry in evaluating groundwater geochemical evolution and aquifer connectivity near a large coal mine, Anhui, China," *Environmental Earth Sciences*, vol. 75, no. 9, article no. 747, 2016.
- [13] P.-H. Huang and J.-S. Chen, "Fisher identify and mixing model based on multivariate statistical analysis of mine water inrush sources," *Journal of the China Coal Society*, vol. 36, no. S1, pp. 131–136, 2011.
- [14] J. T. Lu, X. B. Li, and F. Q. Gong, "Recognizing of mine water inrush sources based on principal components analysis and fisher discrimination analysis method," *China Safety Science Journal*, vol. 22, no. 7, pp. 109–115, 2012.
- [15] H. Chen J, X. B. Li, A. H. Liu et al., "Identifying of mine water inrush sources by Fisher discriminant analysis method," *Journal of Central South University*, vol. 40, pp. 1114–1120, 2009.
- [16] M. Alizamir and S. Sobhanardakani, "A comparison of performance of artificial neural networks for prediction of heavy metals concentration in groundwater resources of toyserkan plain," *Avicenna Journal of Environmental Health Engineering*, vol. 4, no. 1, 2017.
- [17] J. S. Alagha, M. A. M. Said, and Y. Mogheir, "Modeling of nitrate concentration in groundwater using artificial intelligence approach-a case study of Gaza coastal aquifer," *Environmental Modeling & Assessment*, vol. 186, no. 1, pp. 35–45, 2014.
- [18] I. N. Daliakopoulos, P. Coulibaly, and I. K. Tsanis, "Groundwater level forecasting using artificial neural networks," *Journal of Hydrology*, vol. 309, no. 1–4, pp. 229–240, 2005.
- [19] X.-Y. Wang, T. Xu, and D. Huang, "Application of distance discriminance in identifying water inrush resource in similar coalmine," *Journal of the China Coal Society*, vol. 36, no. 8, pp. 1354–1358, 2011.
- [20] X. D. Pen, Y. H. Guo, Y. W. Jie et al., "Application and discussion of fuzzy comprehensive evaluation in identifying mine inrush water source," *Mining safety and environmental protection*, vol. 33, no. 3, pp. 57–59, 2006.
- [21] Y. G. Yang, B. T. Li, and K. Q. Shang, "Fuzzy comprehensive evaluation of water gush and its prediction in coal mines of hebi mining bureau," *Journal of China University of Mining & Technology*, vol. 27, no. 2, pp. 204–208, 1998.
- [22] Z.-G. Yan, P.-J. Du, and D.-Z. Guo, "SVM models for analysing the headstreams of mine water inrush," *Journal of the China Coal Society*, vol. 32, no. 8, pp. 842–847, 2007.

- [23] Z. J. Xu, Y. G. Yang, and L. Tang, "Application of BP neural network in evaluation of water source in mine," in *Safety in Coal Mines*, pp. 4–6, 2 edition, 2007.
- [24] X. L. Zhang, Z. X. Zhang, and S. P. Peng, "Application of the second theory of quantification in identifying gushing water sources of coal mines," *Journal of China University of Mining & Technology*, vol. 32, no. 3, pp. 251–254, 2003.
- [25] I. M. Farnham, K. J. Stetzenbach, A. K. Singh, and K. H. Johannesson, "Deciphering groundwater flow systems in Oasis Valley, Nevada, using trace element chemistry, multivariate statistics, and geographical information system," *Mathematical Geology*, vol. 32, no. 8, pp. 943–968, 2000.