Mine Water and the Environment (2021) 40:442–455
https://doi.org/10.1007/s10230-020-00742-2

TECHNICAL ARTICLE

Water Source Discrimination in a Multiaquifer Mine Using a Comprehensive Stepwise Discriminant Method

Chunlu Jiang[1] · Yanqing An[1] · Liugen Zheng[1] · Wangwang Huang[1]

Received: 21 December 2019 / Accepted: 18 November 2020 / Published online: 5 January 2021

Abstract

Previously proposed discriminant methods cannot accurately identify water sources in a complex multiaquifer mine. Based on statistical analysis of the water chemistry of samples collected at the Xinji no. 2 multiaquifer coal mine, a comprehensive stepwise discriminant method was proposed for this purpose. Characteristic ion contrast and ion proportional coefficients were applied to aquifers with distinct chemical characteristics to establish a characteristic index discrimination system. Aquifers with small differences in water chemistry were identified by the Fisher discriminant method. Different methods (first simple ones, followed by more complex ones) were used to distinguish the water sources of different aquifers. This approach enabled us to identify water sources for the Xinji no. 2 mine and should be tried for other sites with similar hydrogeological conditions.

Keywords Fisher discriminant method · Water chemistry · Water inrush · Xinji mine

Introduction

China is a coal-based country, with a coal output of 3.68 billion tons in 2018, accounting for 47.2% of the world's total coal production and 63.4% of China's energy consumption. The hydrogeological conditions of coal deposits in China are complex and diverse (Liu and Cao 2011), and mine water disasters can be severe and greatly affect safe production (Xu et al. 2018; Yang et al. 2018). When a water hazard develops in a mine, it not only causes huge economic losses but can also cause casualties. Therefore, accurately determining the source of the water involved in water inrush events is a key issue for solving and preventing mine water disasters.

✉ Chunlu Jiang
ahuclj@ahu.edu.cn

Yanqing An
anyanqing2017@163.com

Liugen Zheng
lgzheng@ustc.edu.cn

Wangwang Huang
1512165283@qq.com

[1] School of Resource and Environmental Engineering, Anhui Province Engineering Laboratory for Mine Ecological Remediation, Anhui University, Hefei 230601, People's Republic of China

When analyzing mine water inrush events, one must comprehensively consider the hydrogeological conditions, structural conditions, water level, water temperature, and water chemistry data of the mining area (Liu et al. 2018). Hydrochemical methods have been widely used (Chen et al. 2013; Chidambaram et al. 2013; Cloutier et al. 2008; Kumar et al. 2013; Matter et al. 2006; Petitta et al. 2010; Wang et al. 2016), and there are many methods to investigate mine water inrush events using water chemistry combined with related mathematical models, such as the Fisher discriminant (Chen et al. 2009b; Huang and Chen 2011; Sun et al. 2017), Bayes discriminant model (Chen et al. 2009a; Yan et al. 2020), distance discriminant (Chen et al. 2009c; Wang et al. 2011a, b; Zhou et al. 2010), grey relational analysis (Yue 2002; Qiu et al. 2017), fuzzy clustering analysis (Yin et al. 2006; Yang et al. 2012), neural network (Xu et al. 2007), and support vector machine (Jiang and Liang 2006; Yan and Bai 2009) models to achieve water source discrimination. In addition, when there are many influential factors in a mine water inrush event, some scholars have used principal component analysis (PCA) and a single discriminant method to reduce the model dimensions, reduce the influence of redundant information, and improve the model's validity by eliminating overlapping information from multiple variables (Gong and Lu 2014; Huang and Wang 2018; Ju et al. 2018; Wang et al. 2017; Zhang et al. 2018). Such methods include PCA-Fisher and the PCA-Bayes models.

In situations where the hydrogeological conditions of the mine are relatively simple and there are a limited number of different aquifers, the use of a single method such as these can achieve good discrimination. However, when the hydrogeological conditions are more complicated and there are multiple distinct aquifers, such discriminant models are less effective and do not meet the demands for safe production. A new method for identifying mine water sources was needed.

Study Area

The Xinji no. 2 mine is located in the north of Anhui Province on the alluvial plain north of the Huaihe River. The surface is flat, and the elevation is mostly between 22 and 26 m. The terrain is high in the west and low in the east. Structurally, this area belongs to the southeastern edge of the North China plate and is located in the middle of the southern flank of the Huai'an synclinorium, in the south-central Fufeng nappe structure. The main tectonic lineation is oriented northwest-southeast (Fig. 1).

The mine is located in a thrust nappe tectonic system that has three parts: an overlying allochthonous system, the main nappe surface, and an underlying autochthonous system (Fig. 2). The allochthon is mainly composed of lower Proterozoic gneiss and Cambrian limestone. The autochthon consists of Ordovician limestone, Taiyuan Formation limestone, and coal strata. According to the characteristics of the aquifer media and groundwater burial conditions, the six aquifers affecting safe coal production in the area can be divided into three types from top to bottom: porous, fractured, and karst confined aquifers.

The unconsolidated Cenozoic stratum is a porous, confined aquifer composed of Neogene and Quaternary unconsolidated sediments. Lithologically, these layers are mainly composed of silt, fine sand, medium sand, clay, and sandy clay. A clay gravel layer and a small amount of marl are present locally and represent sedimentary facies

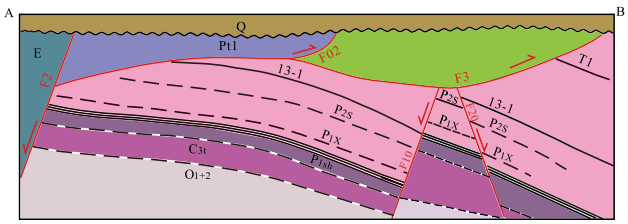

Fig. 2 Cross section of the Xinji no. 2 coal mine along line A–B

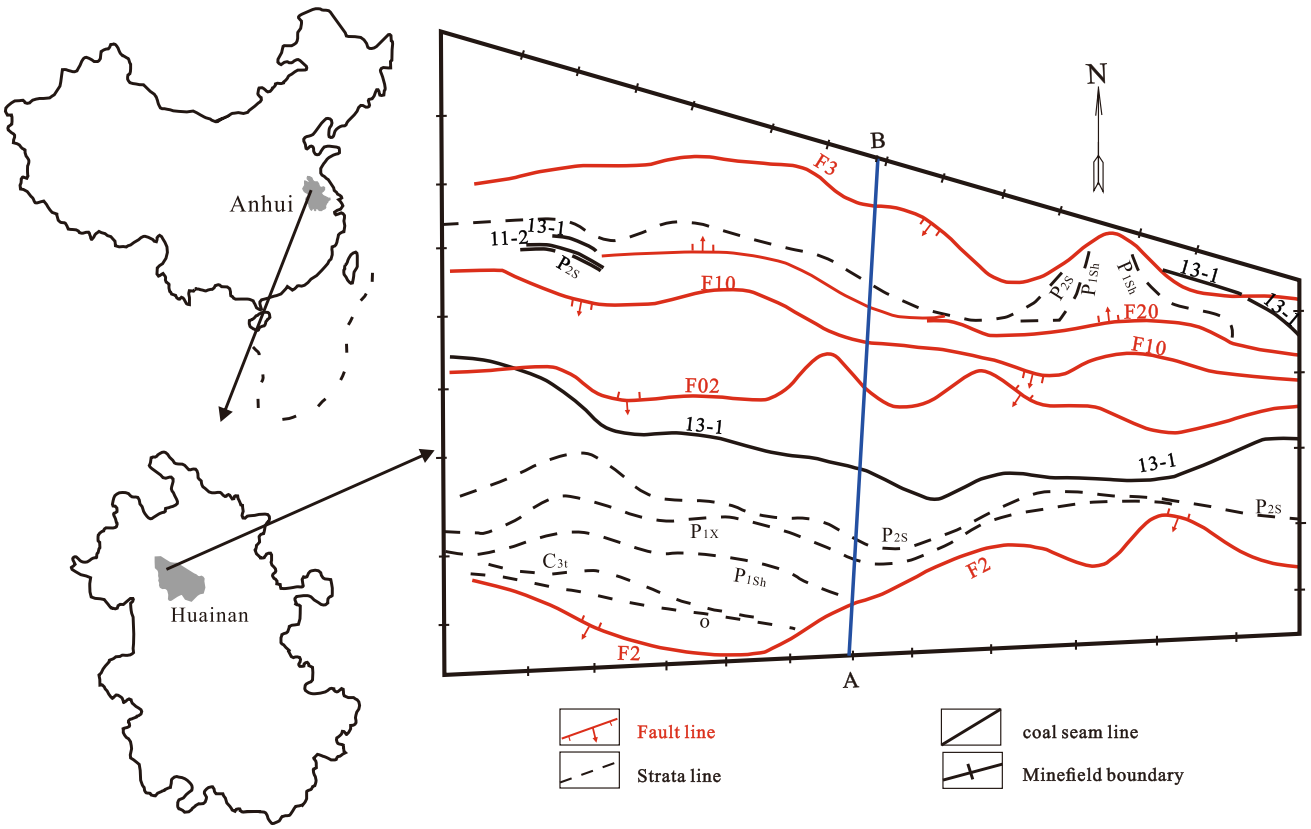

Fig. 1 General map of the study area location and geology

of lakes and rivers, respectively. The water content of this aquifer is moderate to abundant, and the total thickness of the stratum is 110–136.37 m, with an average of 123.7 m (Fu et al. 2004).

The lower Proterozoic gneiss within the nappe structure represents a fractured, confined aquifer that is distributed in the east–west direction between faults F2 and F02; this unit is superimposed on the coal measures and Cambrian strata. The lithology is mainly gray-green hornblende gneiss, light red migmatitic granite-gneiss, hornblende gneiss, plagioclase-hornblende schist, and granitic gneiss. According to the mine's pump test data, the unit water inflow is 0.104 L $(s\ m)^{-1}$, and the overall water abundance is weak.

The Cambrian limestone within the nappe structure represents a confined, karst aquifer. This unit is distributed in the middle and north of the mining area in a band shape with an east–west trend and is superimposed on the coal measures. Lithologically, this unit consists of limestone: dolomitic, oolitic, siliceous, and argillaceous limestones, along with some dolomite, mudstone, sandy mudstone, siltstone, and sandstone. The unit water inflow is 0.00026–0.974 L $(s\ m)^{-1}$ and the water abundance is discontinuous (Jin and Zhang 2002).

The confined, fissured Permian sandstone aquifer is distributed between the main recoverable coal seams, argillaceous rocks, and siltstones. It is dominated by medium- and fine-grained sandstones, with argillaceous rocks and siltstones that vary in thickness. The fissures are poorly developed, and their distribution is uneven; thus, the water yield properties of this unit are weak and related mainly to static reserves (Zhang 2001) According to the mine's pumping test data, the unit water inflow is 0.00143–0.0186 L $(s\ m)^{-1}$.

The confined, fissured Taiyuan Formation limestone karst aquifer is located in the lower part of the Permian sandstone aquifer, has an average thickness of 111.09 m, and contains 12 to 13 layers of limestone. The average total thickness of limestone is 58.6 m, accounting for 52.8% of this unit. The development of karst fissures is heterogeneous, and the water abundance is low to moderate; however, water can be locally abundant. Thus, this unit is classified as a strong regional aquifer. The water abundance depends mainly on the degree of karst development (Wu et al. 2012).

The fissured Ordovician limestone karst aquifer is located in the lower part of the Taiyuan Formation limestone aquifer; an outcrop of the Ordovician strata at the surface is cut by F2 or the nappe structure. This unit has an average exposed thickness of 114.3 m and consists of grayish limestone and gray dolomitic limestone. Its fissures are relatively well developed, and the rocks exhibit honeycomb corrosion holes. According to the mine field's pump test data, the unit water inflow is 0.00037–0.722 L $(s\ m)^{-1}$, and its water abundance is weak to medium.

According to the hydrogeological classification report, the F2 fault on the southern boundary of the mine is the regional aquiclude. The F02 and F3 faults are the main nappe structural planes of the regional faults, which are generally compressive, non-conductive faults. The F10 and F20 faults have weak water conductance and are relatively water-tight. The above-mentioned faults are all large, providing a safe waterproof coal pillar, which under normal circumstances will not become a hydraulic connection channel between the aquifers.

Methodology

Sampling

Data were collected from the water quality analysis ledger of the Xinji Energy Co.'s no. 2 mine; a total of 183 water samples were used to establish the model. The dataset included 23 water samples from the unconsolidated Cenozoic layer, 22 from the nappe gneiss, 12 from the nappe Cambrian limestone, 46 from the Permian sandstone, 66 from the Carboniferous Taiyuan Formation limestone, and 14 from the Ordovician limestone. Water samples were also collected from the underground construction and mine water inrush point, two unconsolidated Cenozoic layer, five Permian sandstone water samples, and six Carboniferous Taiyuan Formation limestone water samples were collected from underground sites to test the applicability of the discriminant model. The analyzed parameters of each water sample included: pH, Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, and HCO_3^-.

Analysis

Each water sample was collected using a pre-cleaned 500 mL high-density polyethylene plastic bottle, and the sample bottle was repeatedly washed with the original water sample 3–5 times before sampling. After the water sample was taken back to the laboratory, it was filtered through a 0.45 μm microporous membrane. The water sample for cation analysis was acidified with 1:4 HNO_3 to adjust the pH to 2, and Na^+, K^+, Ca^{2+}, and Mg^{2+} were measured using inductively coupled plasma atomic emission spectrometry (ICP-AES; IRIS Intrepid IIXSP, Thermo Fisher Scientific, USA). Ion chromatography (ICS-1500, Dion, USA) was used to measure the Cl^-, SO_4^{2-}, and HCO_3^- concentrations by double indicator titration. The experimental water was ultrapure water, and the experimental vessels were soaked in 6 $mol{\cdot}L^{-1}$ HNO_3 solution for 24 h and then washed and dried with deionized water. To ensure accuracy, conservation calculations (positive and negative charges) were performed for each experimental test. Since the $Na^+ + K^+$, Ca^{2+}, Mg^{2+},

Table 1 Statistics on water chemical indexes of various aquifers in Xinji No. 2 coal mine

	pH	TDS mg L^{-1}	$Na^+ + K^+$	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-
Unconsolidated Cenozoic layer water								
Average value	–	325.4	52.6	58.0	16.7	19.0	9.6	349.8
Standard deviation	–	41.0	11.9	9.3	3.3	9.9	11.0	31.6
Coefficient of variation	–	0.13	0.23	0.16	0.20	0.52	1.15	0.09
Nappe gneiss water								
Average value	8.0	1092.3	341.8	57.9	28.2	388.6	116.2	311.0
Standard deviation	0.2	184.8	57.7	14.5	9.1	68.6	25.2	57.3
Coefficient of variation	0.03	0.17	0.17	0.25	0.32	0.18	0.22	0.18
Nappe Cambrain limestone water								
Average value	7.9	1667.4	516.7	83.7	41.9	665.2	194.0	322.1
Standard deviation	0.4	165.6	44.0	14.7	10.4	95.7	61.4	101.0
Coefficient of variation	0.04	0.10	0.09	0.18	0.25	0.14	0.32	0.31
Permian sandstone water								
Average value	8.0	2106.8	803.8	35.3	26.9	808.5	73.5	661.4
Standard deviation	0.4	475.1	195.6	36.9	30.3	291.4	82.3	361.1
Coefficient of variation	0.05	0.23	0.24	1.04	1.13	0.36	1.12	0.55
Taiyuan formation limestone water								
Average value	8.2	2392.8	959.4	21.4	14.9	925.6	45.6	787.8
Standard deviation	0.5	417.7	199.6	16.2	15.7	169.1	65.2	355.3
Coefficient of variation	0.06	0.17	0.21	0.76	1.05	0.18	1.43	0.45
Ordovician limestone water								
Average value	7.3	4949.2	1514.1	302.3	92.5	2631.4	290.3	237.3
Standard deviation	0.4	207.1	61.2	30.4	8.6	193.6	114.4	17.5
Coefficient of variation	0.05	0.04	0.04	0.10	0.09	0.07	0.39	0.07

Cl^-, SO_4^{2-}, and HCO_3^- parameters were measured values, the error was required to be within ± 5%.

Hydrochemical Analysis

Main Chemical Constituents

The 183 water samples were analyzed to determine the chemical properties of each aquifer (Table 1). The coefficient of variation is the ratio of the standard deviation to the average value and reflects the degree of dispersion in the unit mean. The average pH of the Ordovician limestone water is 7.3, which is neutral. The mean pH values of the nappe gneiss, nappe Cambrian limestone, Permian sandstone, and Carboniferous Taiyuan Formation limestone water samples were 8.0, 7.9, 8.1, and 8.2, respectively.

The average total dissolved solid (TDS) value of the unconsolidated Cenozoic layer water was 325.4 mol L^{-1}. Among the anions, the mass concentration of HCO_3^- was the highest, the mass concentrations of Cl^- and SO_4^{2-} were low, and the mass concentrations of the cations exhibited the following order: $Ca^{2+} > Na^+ + K^+ > Mg^{2+}$. Except for SO_4^{2-} and Cl^-, the coefficients of variation for the unconsolidated layer water were small. The average TDS value in the Ordovician limestone water was 4949.2 mol L^{-1}. The Cl^- content was much greater than the HCO_3^- and $SO_4^{2-}$and the $Na^+ + K^+$ content was much greater than the Ca^{2+} and Mg^{2+}. The coefficient of variation for each indicator in the Ordovician limestone water was small. Lower coefficients of variation indicate that the water samples have more consistent chemical compositions (Salinas-Garcia et al. 1997). The mean TDS values of the nappe gneiss, nappe Cambrian limestone, Permian sandstone, and Taiyuan Formation limestone water samples were 1092.3, 1667.4, 2106.8, and 2392.8 mol L^{-1}, respectively, between the values of the unconsolidated Cenozoic and the Ordovician limestone water samples. The anions are mainly Cl^- and HCO_3^-, and the cations are mainly $Na^+ + K^+$. The coefficients of variation of Ca^{2+}, Mg^{2+}, and SO_4^{2-} in the Permian sandstone and Taiyuan Formation limestone water samples were high, whereas the other ions had lower coefficients of variation. The coefficients of variation of the ions in the nappe gneiss and Cambrian limestone water were low.

The box plots in Figs. 3 and 4 show the characteristics of the original data distribution. Figure 3 compares the same parameter for different aquifers, and Fig. 4 compares multiple parameters for the same aquifer. In Fig. 3, the

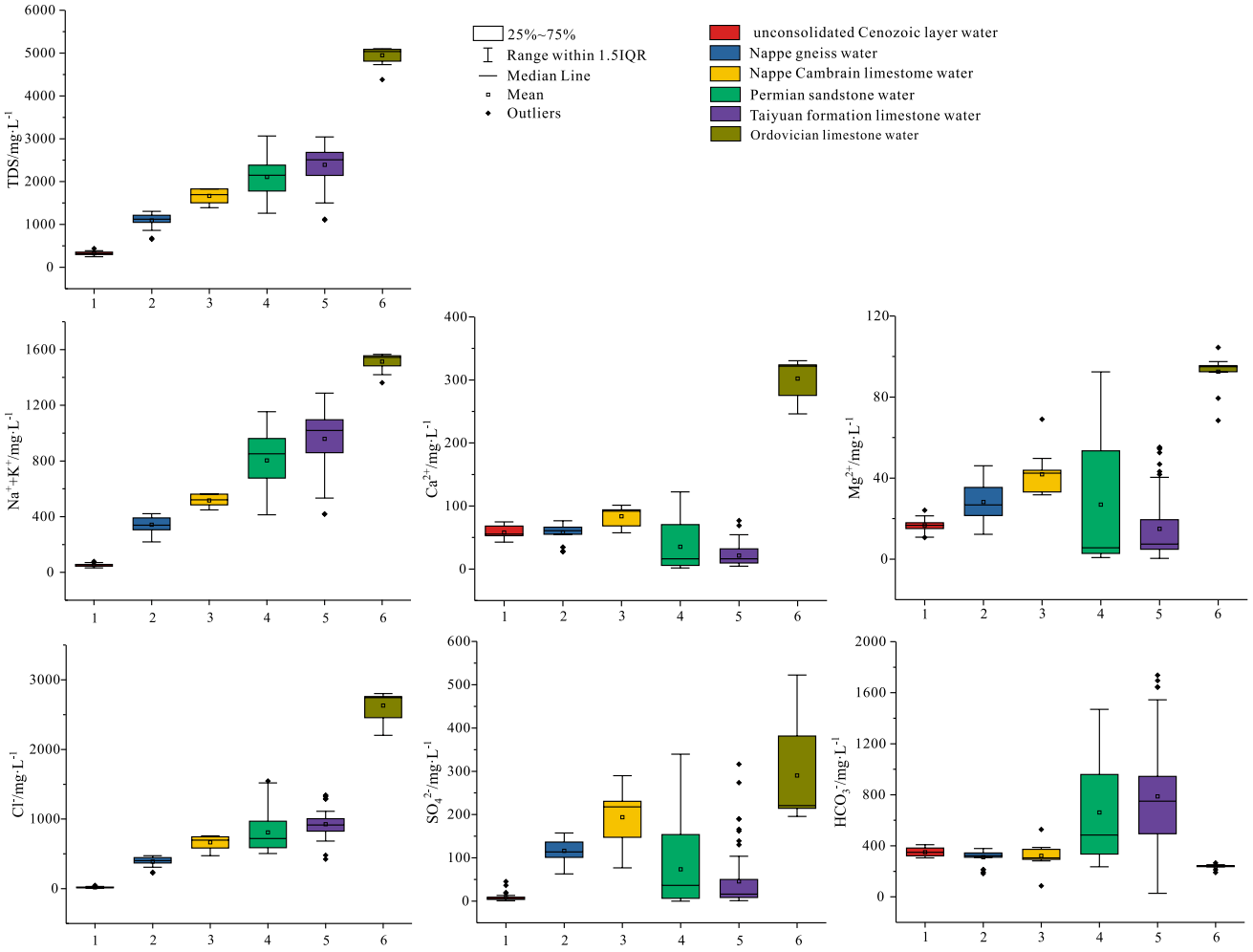

Fig. 3 Comparison of major hydrogeochemical parameters for different aquifers

unconsolidated Cenozoic layer water has significantly less TDS, Na^+, and Cl^- contents than the other aquifers, while the Ordovician limestone water is significantly higher than the other aquifers in terms of TDS, Na^+, and Cl^-. The various formation waters cannot be completely distinguished based on these parameters because the indicator contents of the nappe gneiss water and nappe Cambrian limestone water are similar and because the indicator contents of the Permian sandstone water and Carboniferous Taiyuan Formation limestone water are similar. The coal measure strata water (from the Permian sandstone and Taiyuan Formation limestone) generally has higher levels of TDS, Na^+, Cl^-, and HCO_3^- than the nappe (gneiss and Cambrian limestone) water. In contrast, the Ca^{2+} content distribution is exactly the opposite: the nappe water has higher levels than the coal measure strata water. In the unconsolidated Cenozoic layer, the HCO_3^- content was higher than the Cl^- and SO_4^{2-} contents, and the cation content exhibited the following order: $Ca^{2+} > Na^+ > Mg^{2+}$ (Fig. 4). In the other aquifers, the Na^+ concentrations were significantly higher than the Ca^{2+} and Mg^{2+}, and differences exist in the anion content distribution. The coal measure strata and the nappe water had similar Cl^- and HCO_3^- concentrations, exceeding the SO_4^{2-} concentrations, while the Cl^- content was significantly higher than the HCO_3^- and SO_4^{2-} in the Ordovician limestone water.

Hydrochemical Characteristics

Because the chemical composition of groundwater is complicated, Piper and Durov diagrams were used to analyze the ion distribution and water chemistry type of each aquifer (Fig. 5). It can be seen from the Piper diagram that the cations of the unconsolidated Cenozoic layer water are concentrated below the center of the left triangle: the Ca^{2+} content is highest, the Na^+ content is second, and the anions are concentrated in the lower right corner of the right triangle. In the Durov diagram, the anion concentration is

biased towards the HCO_3^- vertex, indicating that the samples had a high HCO_3^- content. The water samples were mostly the HCO_3–Ca·Na type, with a few of the HCO_3–Ca type.

The anions and cations were more concentrated in the Ordovician limestone water. On the Piper diagram, the anions and cations are distributed in the lower right corner of each triangle. The distribution characteristics are more obvious on the Durov diagram. The cations are mostly biased towards the Na^+ end-member, and the anions are mainly biased towards the Cl^- end-member. This indicates that the chemical type of the Ordovician limestone water is entirely the Cl–Na type. The chemical characteristics of the unconsolidated Cenozoic layer and the Ordovician limestone water were obviously different, which may be related to the burial depth. Water composition varies strongly in the shallow, unconsolidated Cenozoic layer and the runoff and leaching effects are stronger, while the deeper Ordovician limestone exhibits the opposite pattern (Li et al. 2017). The cations in the nappe strata, the Permian sandstone, and the Taiyuan Formation limestone water were mainly Na^+, while the anion distribution was more dispersed. The Durov diagram more intuitively shows that the cations are biased towards the Na^+ end-member. The anions are mostly concentrated on the Cl–HCO_3 line and are more biased towards the Cl^- vertex, with HCO_3^- as the second most abundant anion. The water chemistry types are mostly Cl·HCO_3-Na type, while a few samples exhibit the Cl–Na and HCO_3·Cl–Na types, which may be related to the different mineralogy and its effects on the water chemistry in each aquifer (Armengol et al. 2017).

Hydrogeochemical Process

The ion proportional coefficient is often used to study the water chemistry, and $\gamma Cl^-/\gamma Ca^{2+}$ is used to describe hydrodynamic characteristics. Cl^- is usually enriched in areas with weak hydrodynamic conditions, while Ca^{2+} is the main ion in low-salinity water. Thus, the higher the $\gamma Cl^-/\gamma Ca^{2+}$ value, the weaker the region's hydrodynamic conditions (Su et al. 2018), the slower the groundwater flow, the weaker the changes in the water flow, and the weaker the leaching effect. The $\gamma Cl^-/\gamma Ca^{2+}$ values in the unconsolidated Cenozoic layer, the nappe gneiss, the nappe Cambrian limestone, the Permian sandstone, the Carboniferous Taiyuan Formation limestone, and the Ordovician limestone water were 0.2, 3.9, 4.5, 41.9, 37.6, and 4.9, respectively. These values indicate that the hydrodynamic conditions of the different aquifers in the study area are quite different: the hydrodynamic conditions of the unconsolidated Cenozoic layer are the strongest, those of the Permian sandstone and the Taiyuan Formation limestone are the weakest, and the remaining aquifers are between these two aquifer sets.

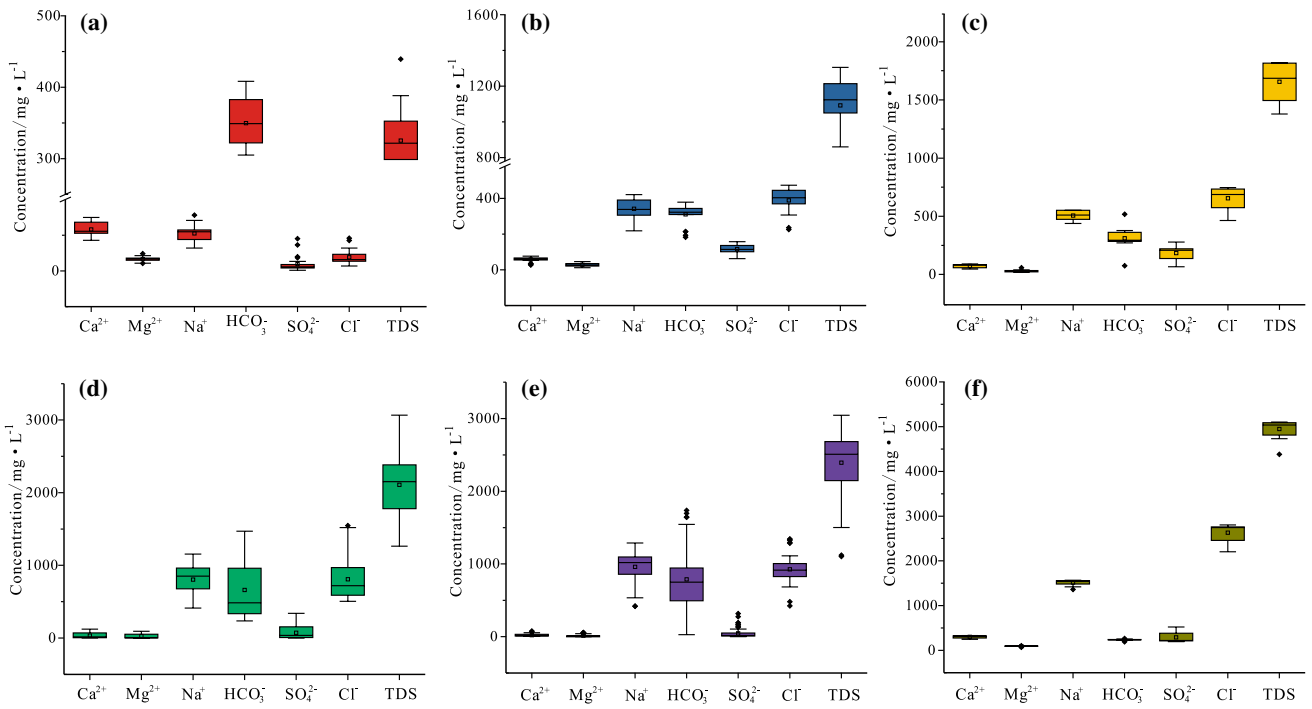

Fig. 4 Boxplots of major hydrochemical parameters for different aquifers: **a** unconsolidated Cenozoic layer water; **b** Nappe gneiss water; **c** Nappe Cambrian limestone water; **d** Permian sandstone water; **e** Taiyuan formation limestone water; **f** Ordovician limestone water

The $\gamma Na^+/\gamma Cl^-$ coefficient can be used to characterize the leaching and accumulation strength of mineral salts resulting from water–rock interactions. This coefficient is also called the genetic coefficient of groundwater. Since the chemical properties of Cl^- are stable, whereas Na^+ concentrations can be affected by adsorption, precipitation, and ion exchange; this coefficient is important for judging the Na^+ source in groundwater. Under natural conditions, the same amount of Na^+ and Cl^- are mobilized by the dissolution of rock salt (Jia et al. 2015). Therefore, when the $\gamma Na^+/\gamma Cl^-$ value is close to 1, the leaching of salt-bearing formations plays a leading role in groundwater composition. In Fig. 6, the $\gamma Na^+/\gamma Cl^-$ values of the unconsolidated Cenozoic layer water are all above the 2:1 line, those of the Ordovician limestone water are all below the 1:1 line, those of the nappe water are mostly between the 1:1 and 2:1 lines, and those of the coal measure strata water are relatively dispersed. This indicates that the Na^+ content of the unconsolidated Cenozoic layer water is high and that the Na^+ is derived from other sources in addition to rock salt dissolution. The Na^+ in the Ordovician limestone water is mainly derived from the dissolution of rock salt. Analysis of the $\gamma Cl^-/\gamma Ca^{2+}$ results reveals that the unconsolidated Cenozoic layer water has stronger hydrodynamic conditions and greater changes in water chemical composition. Therefore, one of the reasons for the high Na^+ content of the unconsolidated Cenozoic layer water sample may be cation exchange, that is, Ca^{2+} in water substituting for Na^+.

The $\gamma HCO_3^-/\gamma Cl^-$ value similarly reflects the anion evolution process and changes in the component distribution ratio (Mondal et al. 2010) and can also reflect the degree of groundwater concentration. In Fig. 7, the $\gamma HCO_3^-/\gamma Cl^-$ values of the unconsolidated Cenozoic layer water are far above the 1:1 line, whereas those of the Ordovician limestone water are far below the 1:1 line. This indicates that little concentration is occurring in the Cenozoic layer and more concentration is occurring in the Ordovician limestone.

The saturation index (SI) reflects the saturation state of mineral phases in groundwater (Deutsch 1997; Stumm and Morgan 1995). Due to errors introduced during water quality sampling and analysis, and the calculation of mineral equilibrium constants and ion activity, the calculated result for the SI of a mineral is inevitably uncertain. Therefore, an aqueous solution and a mineral are commonly considered in equilibrium when the SI is 0 ± 0.5 (Qian et al. 2005).

The nappe Cambrian, the Carboniferous Taiyuan Formation, and the Ordovician strata are all limestone formations. Theoretically, due to the leaching of limestone, the content of HCO_3^- in the three aquifers should be high (Wang et al. 2019). Figure 8a and b show the SI of calcite and halite in the limestone water. The calcite SI of only a few Taiyuan Formation limestone water samples was less than − 0.5, while the calcite SI of the nappe Cambrian limestone water and the Ordovician limestone water were almost all in equilibrium or in a supersaturated state. The halite SI of all of the various limestone water samples was far less than 0 and unsaturated. In these three aquifers, calcite tends to precipitate, while halite, if present, dissolves. Due to the long-term leaching of salt-bearing rocks, the anion content of the three aquifers is dominated by Cl^-.

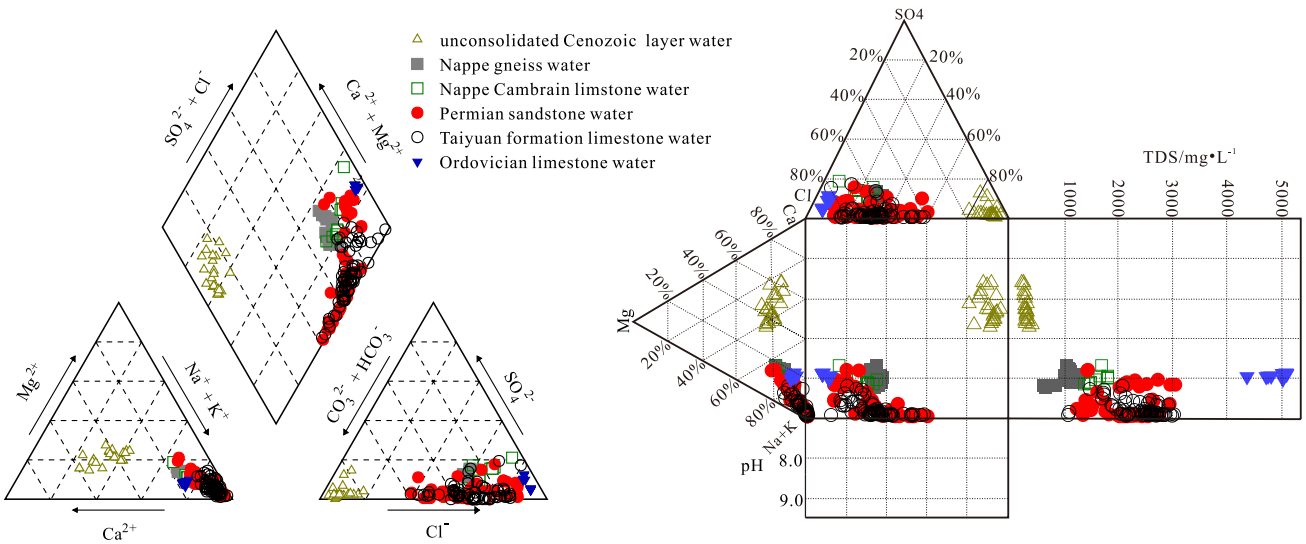

Fig. 5 Piper and Durov diagrams for different aquifers

Water Source Discrimination Model

Determination of the Discriminant Index and Threshold

1. Unconsolidated Cenozoic layer and Ordovician limestone water.

The TDS and water chemistry types of these two aquifers are different, with distinct TDS, Cl^-, and $Na^+ + K^+$ concentrations and can be effectively distinguished from the other aquifer water samples, as shown in Figs. 3, 9, and Table 2. Therefore, the above three indicators were selected as discriminating indicators for the unconsolidated Cenozoic layer water and Ordovician limestone water.

2. Nappe water and coal measure strata water.

The nappe gneiss and nappe Cambrian limestone water are similar in terms of conventional ions and water chemistry types, while the Permian sandstone water and the Carboniferous Taiyuan Formation limestone water are similar. Therefore, the first two were collectively classified as nappe water. Similarly, the Permian sandstone and Taiyuan Formation limestone water were classified collectively as coal measure strata water. This study found that the characteristic ion ratio coefficients ($\gamma Cl^-/\gamma Ca^{2+}$) of the nappe and coal measure strata water are significantly different (Table 3) and can be used to distinguish the two.

3. Nappe gneiss water and nappe Cambrian limestone water.

The nappe gneiss water and the nappe Cambrian limestone water have distinct characteristics in terms of TDS and Cl^- content. Thus, these indicators can be used to distinguish the two, as shown in Fig. 10 and Table 4.

Fisher Discriminant Model

1. Fisher discriminant principle.

The Fisher model is a linear discriminant method with no specific requirements for overall distribution. The basic principle of the Fisher method is to project the data points with high dimensions to a space with lower dimensions (such as one-dimensional lines) and solve high-dimensional problems using a 1D method. 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

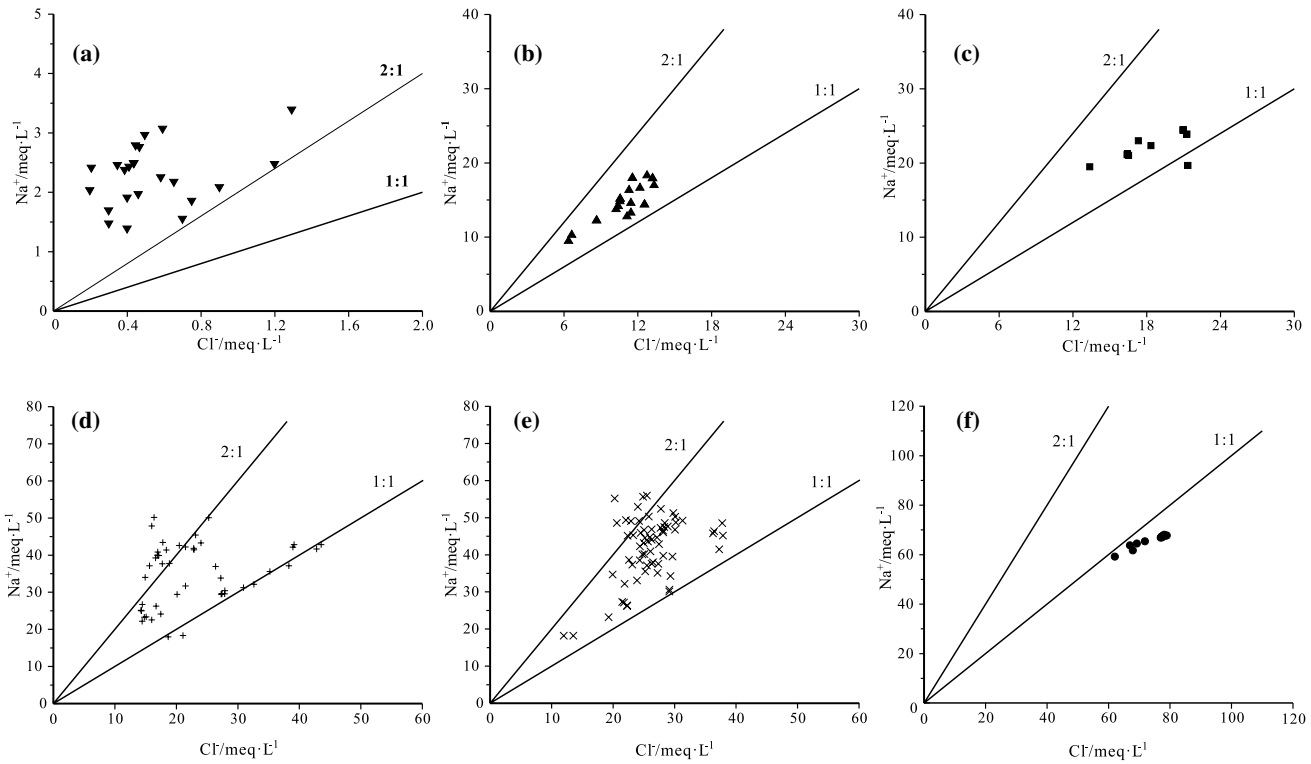

Fig. 6 Calculation results of $\gamma Na^+/\gamma Cl^-$: **a** unconsolidated Cenozoic layer water; **b** Nappe gneiss water; **c** Nappe Cambrian limestone water; **d** Permian sandstone water; **e** Taiyuan formation limestone water; **f** Ordovician limestone water

between-class distance and minimum inner-class distance. Then, a new sample can be identified which group it belongs to by comparing the function score with the central value of each group (Chen et al. 2009b; Huang and Chen 2011; Huang and Wang 2018).

2. Fisher discriminant function.

The conventional water chemical index and ion proportional coefficient of the Permian sandstone water and Taiyuan Formation limestone water are similar, and these waters are indistinguishable via conventional means based on these two aspects. To this end, seven indicators, namely, $Na^+ + K^+$, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, HCO_3^-, and TDS, were selected as the discriminant factor between the sandstone and limestone water, and the Fisher discriminant method was used for discrimination. The Fisher discriminant function was determined as follows:

$$F = 0.025X_1 + 0.049X_2 + 0.072X_3 - 0.007X_4 + 0.002X_5 - 0.002X_6 - 0.009X_7 + 3.067 \quad (1)$$

In the formula, X_1, X_2, X_3, X_4, X_5, X_6 and X_7 are $Na^+ + K^+$, Ca^{2+}, Mg^{2+}, Cl^-, SO_4^{2-}, HCO_3^-, and TDS, respectively. Table 5 shows the central values of the typical discriminant functions in each group. The function scores were 0.798 in the center of the Permian sandstone water and − 0.556 in the center of the Taiyuan Formation limestone water. On this basis, it can be determined which group a new sample belongs to by comparing the distance between the water sample function value and the central value of these two types of water sample groups.

Comprehensive Stepwise Discrimination

The comprehensive stepwise discriminant method is a combination of the characteristic ion contrast, ion proportional coefficient, and Fisher discriminant methods. The different methods are used to distinguish different aquifer water sources. First, simple methods are applied, then more complex methods, to gradually determine the water source type, as shown in Fig. 11.

The specific steps are as follows:

(a) In the first step, for a single aquifer water sample to be discriminated, the characteristic ion contrast method is used to judge whether the water sample is from the unconsolidated Cenozoic layer or the Ordovician limestone. If the result can be determined, the discrimination is stopped; otherwise, the next step is performed.

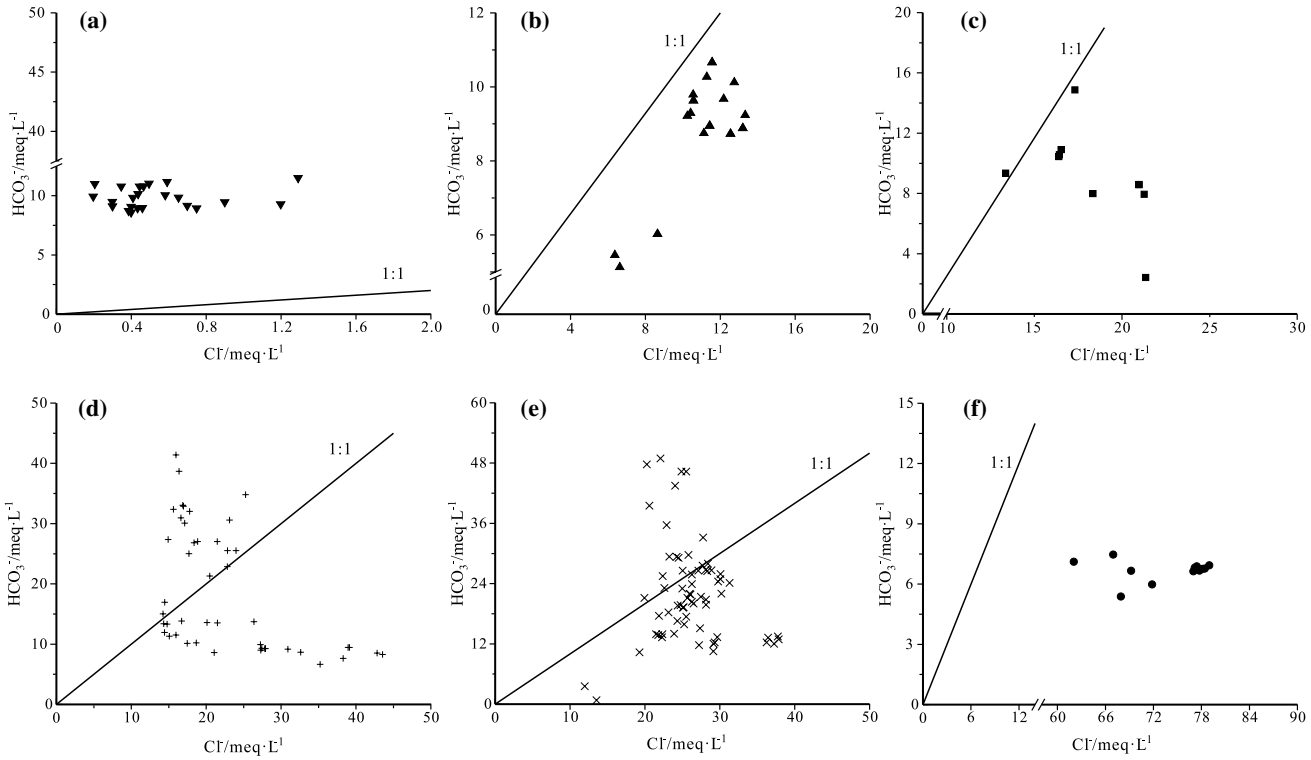

Fig. 7 Calculation results of $\gamma HCO_3^-/\gamma Cl^-$: **a** unconsolidated Cenozoic layer water; **b** Nappe gneiss water; **c** Nappe Cambrian limestone water; **d** Permian sandstone water; **e** Taiyuan formation limestone water; **f** Ordovician limestone water

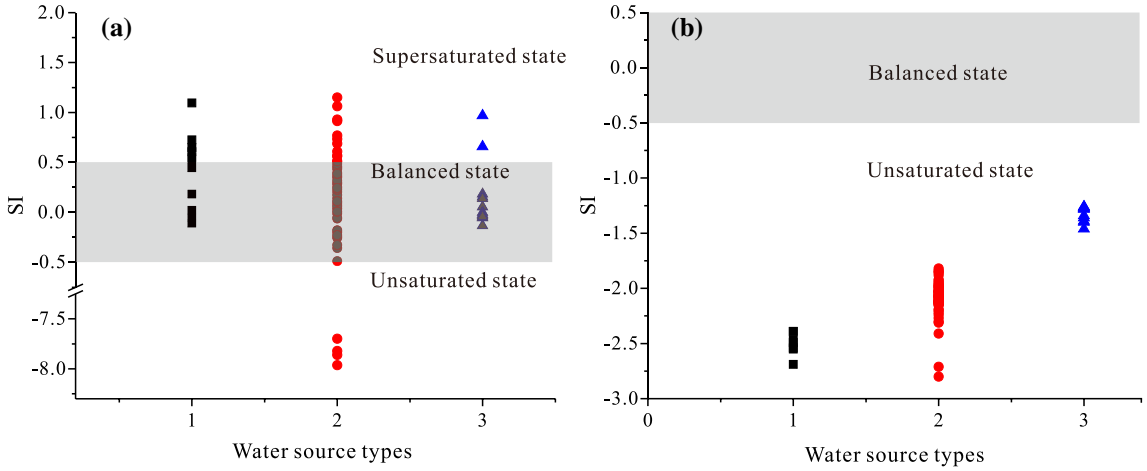

Fig. 8 Saturation index calculation results: **a** calcite; **b** halite. Water source types ID: 1, Nappe cambrain limestone water; 2, Carboniferous Taiyuan formation limestone water; 3, Ordovician limestone water

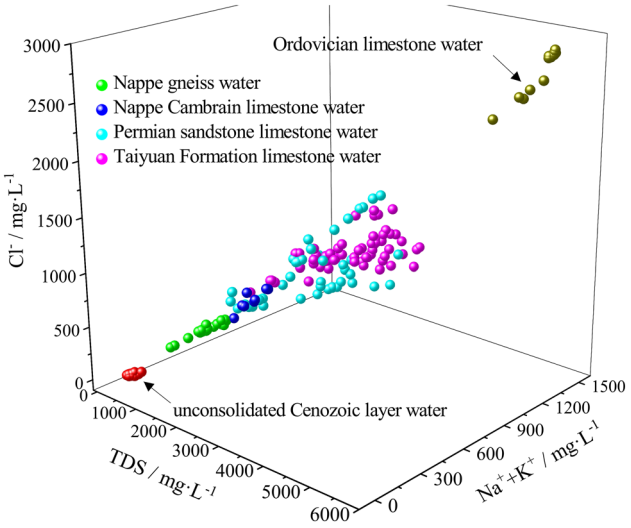

Fig. 9 Characteristic ion contrast between unconsolidated Cenozoic water and Ordovician limestone water

Table 2 Characteristic ion range of unconsolidated Cenozoic aquifer and Ordovician limestone aquifer water

Characteristic ion		Characteristic value/mg·L^{-1}	
		Unconsolidated Cenozoic layer water	Ordovician limestone water
TDS	Average value	325.4	4949.2
	Minimum value	247.2	4381.1
	Maximum value	439.5	5103.0
Cl^-	Average value	19.0	2628.2
	Minimum value	7.0	2202.5
	Maximum value	45.8	2803.3
$Na^+ + K^+$	Average value	52.6	1521.4
	Minimum value	32.0	1361.8
	Maximum value	78.1	1565.9

(b) In the second step, the ion proportional coefficient method is used to determine whether the water sample is nappe water or coal measure strata water.

(c) In the third step, if the water sample was determined to be nappe water, the characteristic ion contrast method is used to determine whether the water sample is from the nappe gneiss or nappe Cambrian limestone. If the water sample is determined to be coal measure strata water, the Fisher method is used to determine whether the water sample is from the Permian sandstone or Taiyuan Formation limestone.

Thus, after these three steps, a water sample from a single aquifer can be classified as being derived from the unconsolidated Cenozoic layer, the nappe gneiss, the nappe Cambrian limestone, the Permian sandstone, the Carboniferous Taiyuan Formation limestone, or the Ordovician limestone.

Discriminant Effect Test

To investigate the validity of the discriminant method, the error discriminant rate was calculated using the back-generation estimation method based on the model sample. For n_i training samples, $x_{\alpha}^{(i)} = \left(x_{\alpha 1}^{(i)}, \ldots, x_{\alpha p}^{(i)}\right)^T$, ($\alpha = 1, 2, \ldots, n_i$; $i = 1, 2, \ldots, k$), from the overall G_i capacity. All modeling samples were taken as new samples, which were sequentially substituted into the established discriminant model. This process is called back-discrimination. If the total number of error

Table 3 Ion proportional coefficients range of nappe water and coal measure strata water

Characteristic ion proportional coefficients	Characteristic value	
	Nappe water	Coal measure strata water
$\gamma Cl^-/\gamma Ca^{2+}$		
Average value	4.1	36.9
Minimum value	3.4	5.3
Maximum value	5.0	209.7

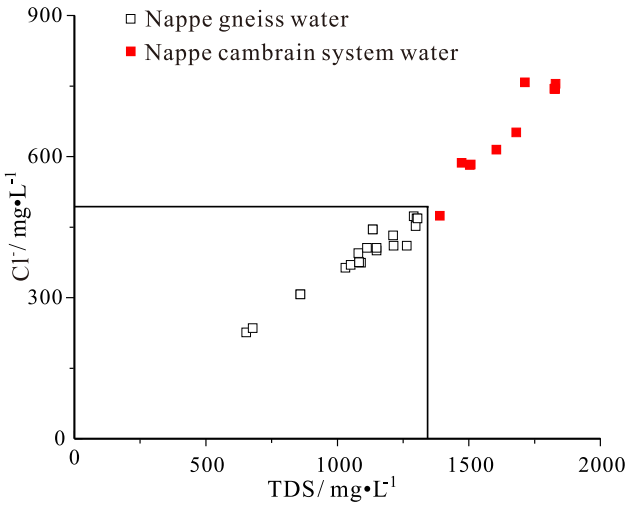

Fig. 10 TDS-Cl^- relationship diagram of nappe water

Table 4 Characteristic ion range of nappe water

Characteristic ion	Characteristic value/mg L^{-1}	
	Nappe gneiss water	Nappe cambrain limestone water
TDS		
Average value	1092.3	1667.4
Minimum value	653.7	1389.8
Maximum value	1305.1	1829.3
Cl^-		
Average value	388.6	665.2
Minimum value	226.1	474.3
Maximum value	473.3	757.9

discriminations is N, the back-generation of the error discrimination rate η is estimated as:

$$\eta = \frac{N}{n_1 + n_2 + \ldots + n_m} \quad (2)$$

The validity test results of the discriminant model are shown in Table 6. The overall discriminant accuracy rate of the model was 86.3%, and for the unconsolidated Cenozoic layer, nappe gneiss, nappe Cambrian limestone, and Ordovician limestone water, the accuracy exceeded 91%.

Discriminant Model Application

The 13 recently collected water samples were taken as test samples and substituted into the established discriminant model. The final discrimination results of the 13 samples were consistent with the 10 actual sampling horizons (Table 7). The determination effect of the unconsolidated Cenozoic layer water was 100%, while it was 72.7% for the coal measure strata water.

Judging from these results and application, the discriminant model has a good overall discriminant effect. However, compared with the other aquifers, it was relatively poor in discriminating the Permian sandstone water and the Taiyuan formation limestone water. Similar inferences can be obtained from the analysis of the water chemistry and hydrogeochemical effects of the aquifers described above. Although the Taiyuan formation limestone water is dominated by thin-layer limestone, it is also sandwiched by thin argillaceous sandstones, which are similar in composition to the Permian Sandstone. In addition, the hydrogeochemical effects of the two aquifers are similar. Therefore, the hydrogeochemical characteristics of the two aquifers are similar, which makes identification of these water sources more challenging.

Conclusions

Due to the influence of buried depth, concentration, leaching and water flow changes, the TDS value of the Ordovician limestone water in the study area is relatively high, the anions are mainly Cl^-, the cations are mainly Na^+, and the water chemistry type is entirely Cl–Na. The unconsolidated Cenozoic layer water has lower TDS values; the anions are dominated by HCO_3^-; the cation Ca^{2+} is dominant, followed by Na^+; the water chemistry type is mainly HCO_3–Ca·Na; and there are a few samples with the HCO_3–Ca types. The water chemistry characteristics of the sandstone, limestone, nappe gneiss, and nappe Cambrian limestone water were

Table 5 Central values of Fisher discriminant function

Water source types	Central values
Permian sandstone water	0.798
Taiyuan formation limestone water	− 0.556

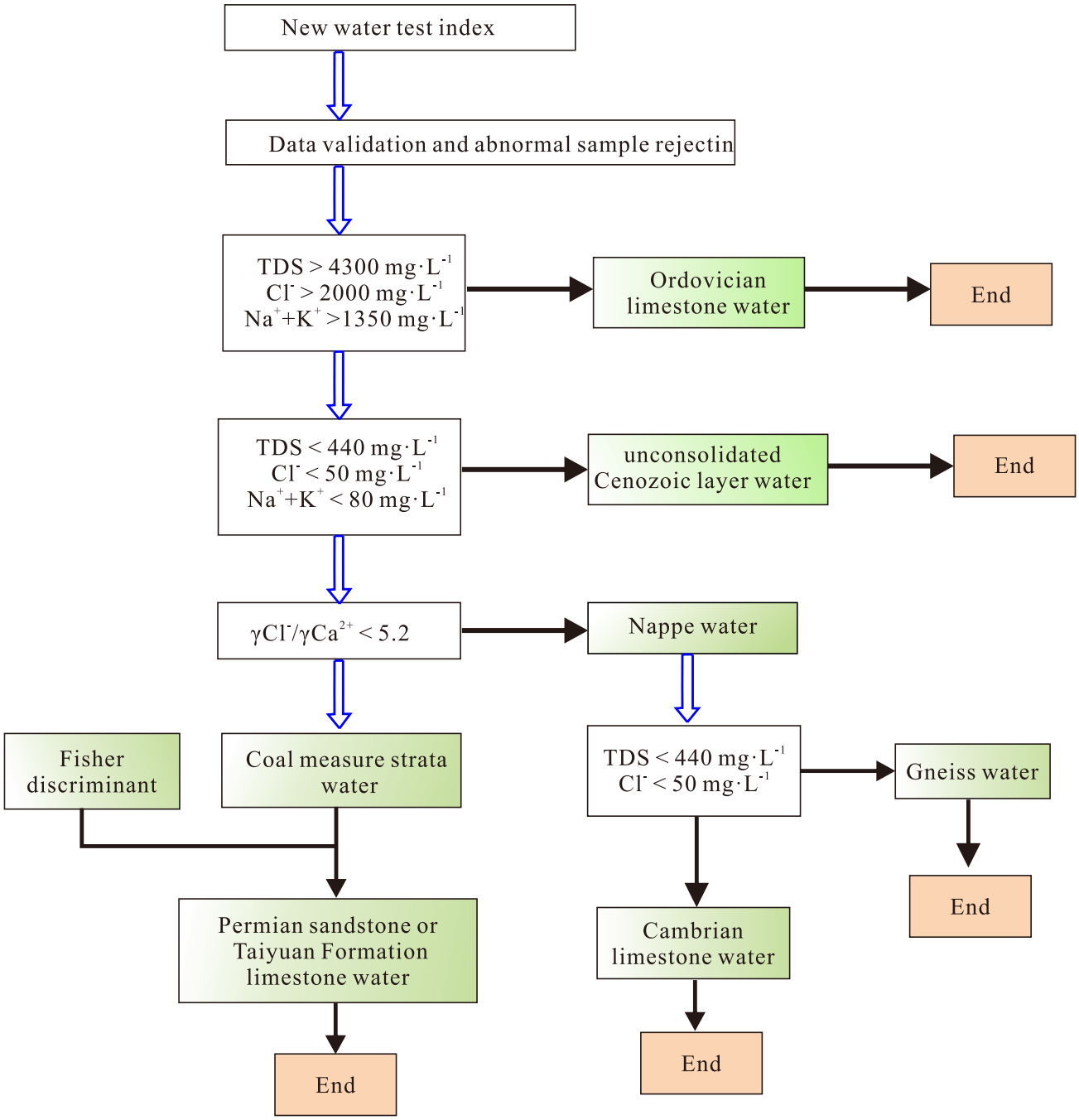

Fig. 11 Flow diagram of comprehensive stepwise discriminant method

between those of the unconsolidated layer and Ordovician limestone water; these aquifers are characterized by Cl^- and HCO_3^- as the main anions, Na^+ as the main cation, and are dominantly of the Cl• HCO_3–Na type.

Using the characteristic ion contrast, the ion proportional coefficient, and the Fisher discriminant methods, a comprehensive stepwise discriminant model was established for water source identification of the six aquifers in the Xinji no. 2 Mine. Different methods (first simple ones followed by complex ones) were used for the various aquifers to determine the water source. The modeled sample back-calculation method was used to test the discrimination effect using recently collected samples. The results show that the established model is quite accurate and can provide a new method for mine water source identification under similar conditions.

Table 6 Results of sample resubstituting

Water source types	Number of modeling samples	Samples number of discriminating correct	Accuracy/% (%)
unconsolidated Cenozoic layer water	23	23	100
Nappe gneiss water	22	22	100
Nappe Cambrain limestone water	12	11	91.7
Permian sandstone water	46	36	78.3
Taiyuan formation limestone water	66	52	78.8
Ordovician limestone water	14	14	100
Total	183	158	86.3

Table 7 Discrimination of test samples

Serial number	Water quality analysis results/mg L^{-1}							Original water source type	Discriminant result
	TDS	$Na^+ + K^+$	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-		
1	3843.1	1146.3	234.8	125.6	1850.0	261.2	450.4	I	I
2	2378.5	956.5	27.7	19.0	915.5	42.3	834.8	I	***II***
3	3364.0	1063.7	169.5	96.7	1592.2	199.8	484.5	I	I
4	1395.2	589.1	14.7	7.7	605.1	45.8	330.9	I	I
5	1245.7	460.7	28.4	12.9	447.3	0	592.6	I	I
6	2621.1	1058.5	15.3	6.8	999.7	189.7	702.2	II	II
7	2556.5	1028.5	17.6	8.7	921.7	189.8	780.3	II	II
8	1896.5	740.3	27.3	19.4	776.4	21.1	624.2	II	***I***
9	2508.4	1003.0	16.8	7.3	911.1	74.9	750.6	II	II
10	2461.0	1013.0	16.0	8.8	932.3	30.7	848.2	II	II
11	1619.2	658.4	16.8	8.9	550.9	59.3	650.5	II	***I***
12	334.5	41.8	62.8	13.5	43.2	28.5	289.5	III	III
13	334.9	75.5	33.8	12.6	53.7	33.1	252.3	III	III

I, Permian Sandstone water; II, Taiyuan formation limestone water; III, unconsolidated Cenozoic layer water. The bold and italics indicate error discriminant in the discrimination results

Acknowledgements Our deepest gratitude goes to the editors and anonymous reviewers for their careful work and thoughtful suggestions that helped improve this paper substantially. The authors also gratefully acknowledge the financial support of the National Natural Science Foundation of China (41602310) and China Postdoctoral Science Foundation (2017M611044).

References

Armengol S, Manzano M, Bea SA, Martinez S (2017) Identifying and quantifying geochemical and mixing processes in the Matanza-Riachuelo aquifer system, Argentina. Sci Total Environ 599:1417–1432

Chen HJ, Li XB, Liu AH (2009) Studies of water source determination method of mine water inrush based on Bayes' multigroup stepwise discriminant analysis theory. Rock Soil Mech 30(12):3655–3670 **(in Chinese)**

Chen HJ, Li XB, Liu AH, Dong LJ, Liu ZX (2009) Forecast method of water inrush quantity from coal floor based on distance discriminant analysis theory. J China Coal Soc 34(4):487–491 **(in Chinese)**

Chen HJ, Li XB, Liu AH, Peng SQ (2009) Identifying of mine water inrush sources by Fisher discriminant analysis method. J Cent South Univ 40(4):1114–1120 **(in Chinese)**

Chen LW, Yin XX, Gui HR (2013) Water-rock interaction tracing and analysis of deep quifers in the mining area using isotope and hydrochemistry methods. Acta Geol Sin 87(7):1021–1030 **(in Chinese)**

Chidambaram S, Anandhan P, Prasanna MV, Srinivasamoorthy K, Vasanthavigar M (2013) Major ion chemistry and identification of hydrogeochemical processes controlling groundwater in and around Neyveli lignite mines, Tamil Nadu, South India. Arab J Geosci 6(9):3451–3467

Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. J Hydrol 353(3–4):294–313

Deutsch WJ (1997) Groundwater geochemistry fundamentals and applications to contamination. CRC Press, New Jersey, pp 61–67

Fu XJ, Lai DJ, Wang XM, He QL (2004) Brief comments on water inflow mechanism of coal mining facetested under water body. Coal Sci Technol 32(4):69–71 **(in Chinese)**

Gong FQ, Lu JT (2014) Recognition method of mine water inrush sources based on the principal element analysis and distance discrimination analysis. J Min Safety Eng 31(2):236–242 **(in Chinese)**

Huang PH, Chen JS (2011) Fisher indentify and mixing model based on multivariate statistical analysis of mine water inrush sources. J China Coal Soc 36(S1):131–136 **(in Chinese)**

Huang PH, Wang XY (2018) Piper-PCA-Fisher recognition model of water inrush source: a case study of the Jiaozuo mining area. Geofluids 2018:1–10

Jia ZX, Zang HF, Zhen XQ (2015) The origin of Na^+, Cl^- and thermal source of karst groundwater in the stagnant area of Liulin spring basin. Carsol Sin 34(6):570–576 **(in Chinese)**

Jiang AN, Liang B (2006) The particle swarm optimization support vectors machine method of identifying standard components of ions of groundwater. J China Coal Soc 31(3):310–313 **(in Chinese)**

Jin LF, Zhang WB (2002) Nappe cambrain system geology and hydrological geology characteristic in Xinji mine area. Coal Techno 21(9):64–65 **(in Chinese)**

Ju QD, Hu YB, Zhang SY (2018) Mine water inrush source identification method based on principal component analysis and Bayesian discriminant. Coal Eng 50(12):90–94 **(in Chinese)**

Kumar M, Herbert R, Ramanathan A, Rao MS, Kim K, Deka JP, Kumar B (2013) Hydrogeochemical zonation for groundwater management in the area with diversified geological and land-use setup. Chem Erde-Geochem 73(3):267–274

Li X, Chen WF, Wang LQ, Xia FX, Zhang YB, Yuan ML (2017) An analysis of hydrochemical characteristics and environmental isotopic characteristics of the groundwater in the bedrock mountain area in Northern Songxian County, Henan Province. Acta Geosci Sin 38(3):403–412 **(in Chinese)**

Liu HH, Cao YQ (2011) Technologies of preventing coal mine Water hazards for sustainable development in north China. Geotech Geol Eng 29(1):1–5

Liu Q, Sun YJ, Xu ZM, Xu G (2018) Application of the comprehensive identification model in analyzing the source of water inrush. Arab J Geosci 11(9):1–10

Matter JM, Waber HN, Loew S, Matter A (2006) Recharge areas and geochemical evolution of groundwater in an alluvial aquifer system in the Sultanate of Oman. Hydrogeol J 14(1–2):203–224

Mondal NC, Singh VP, Singh VS, Saxena VK (2010) Determining the interaction between groundwater and saline water through groundwater major ions chemistry. J Hydrol 388(1–2):100–111

Petitta M, Mugnozza GS, Barbieri M, Fasani GB, Esposito C (2010) Hydrodynamic and isotopic investigations for evaluating the mechanisms and amount of groundwater seepage through a rockslide dam. Hydrol Process 24(24):3510–3520

Qian H, Ma ZY, Li PY (2005) Hydrogeochemistry. Geological Publishing House, Beijing, pp 45–50

Qiu M, Shi LQ, Teng C, Zhao Y (2017) Assessment of water inrush risk using the fuzzy delphi analytic hierarchy process and grey relational analysis in the Liangzhuang coal mine, China. Mine Water Environ 36(1):39–50

Salinas-Garcia JR, Hons FM, Matocha JE (1997) Soil carbon and nitrogen dynamics as affected by long-term tillage and nitrogen fertilization. Biol Fert Soils 25(2):182–188

Stumm W, Morgan JJ (1995) Aquatic chemistry, 3 Edit. Wiley, New York City, pp 398–416

Su CT, Nie FY, Zou SZ, Zhao GS, Luo F, Huang QB, Ba JJ, Li XP, Liang JP, Yang Y (2018) Hydrochemical characteristics and formation mechanism of strontium-rich groundwater in Xintian County, Hunan Province. Geoscience 32(3):554–564 **(in Chinese)**

Sun FX, Wei JC, Wang YP, Liu CE (2017) Recognition method of mine water source based on Fisher's discriminant analysis and centroid distance evaluation. Coal Geol Explor 45(1):80–84 **(in Chinese)**

Wang XY, Ji HY, Wang Q, Liu XM, Huang D, Yao XP, Chen GS (2016) Divisions based on groundwater chemical characteristics and discrimination of water inrush sources in the Pingdingshan coalfield. Environ Earth Sci 75(10):1–11

Wang J, Li XB, Cui T, Yang JL (2011) Application of distance discriminant analysis method to headstream recognition of water-bursting source. Procedia Eng 26:374–381

Wang BB, Lu GP, Hu XN, Ou H (2019) Hydrochemical characterization of thermal spring waters in the deep fault region in western Guangdong. Environ Chem 38(5):1150–1160 **(in Chinese)**

Wang XY, Xu T, Huang D (2011) Application of distance discriminants in identifying water inrush resource in similar coalmine. J China Coal Soc 36(8):1354–1358 **(in Chinese)**

Wang Y, Zhou MR, Yan PC, He CY, Liu D (2017) Identification of coalmine water inrush source with PCA-BP model based on laser-induced fluorescence technology. Spectrosc Spect Anal 37(3):978–983

Wu W, Hu YB, Wang HZ (2012) Coal Seam's characteristics of hydrogeology in Xinji No. 2 mine and calculation of water yield. Coal Technol 31(1):157–159 **(in Chinese)**

Xu ZM, Sun YJ, Gao S, Zhao XM, Duan RQ, Yao MH, Liu Q (2018) Groundwater source discrimination and proportion determination of mine inflow using ion analyses: a case study from the Longmen coal mine, Henan Province, China. Mine Water Environ 37(2):385–392

Xu ZJ, Yang YG, Tang L (2007) Application of BP neural network in evaluation of water source in mine. Saf Coal Mine 38(2):4–6 **(in Chinese)**

Yan ZG, Bai HB (2009) MMH Support vector machines model for recognizing multi-headstream of water inrush in mine. China J Rock Mech Eng 28(2):324–329 **(in Chinese)**

Yan BQ, Ren FH, Cai MF, Qiao C (2020) Bayesian model based on Markov chain Monte Carlo for identifyingmine water sources in submarine gold mining. J Clean Prod 253(1):1–10

Yang WF, Shen DY, Ji YB, Wang Y (2012) Discrimination of mine water bursting source based on fuzzy system. Appl Mech Mater 1810(365):644–648

Yang BB, Yuan JH, Duan LH (2018) Development of a system to assess vulnerability of flooding from water in karst aquifers induced by mining. Environ Earth Sci 77(3):1–13

Yin XX, Xu GQ, Gui HR, Chen LW (2006) Analyzing for sources of inrush-water in Wanbei mining area by systemic clustering and stepwise distinguishing. Coal Geol Explor 34(2):58–61 **(in Chinese)**

Yue M (2002) Related analysis of gray system applied to determinate water resources of mine water inrush. Coal Sci Technol 30(4):37–39 **(in Chinese)**

Zhang WB (2001) Features of lower horse strata and its significances for retain waterproof pillar in no. 2 coal mine of Xinji. Coal Geol China 13(3):40–41 **(in Chinese)**

Zhang M, Liu QM, Zhang YT (2018) Fisher discrimination model for sources of mine water inrush based on PCA analysis. Coal Technol 37(3):172–174 **(in Chinese)**

Zhou J, Shi XZ, Wang HY (2010) Water-bursting source determination of mine based on distance discriminant analysis model. J China Coal Soc 35(2):278–282 **(in Chinese)**