An Interdisciplinary Response to Mine Water Challenges - Sui, Sun & Wang (eds)
©2014 China University of Mining and Technology Press, Xuzhou, ISBN 978-7-5646-2437-8

Comparative Study among Diverse Statistical Methods used in Groundwater Source Identification in Coal Mine Area

Song Chen
School of Earth Sciences and Engineering, Suzhou University, Suzhou, Anhui China
szxychensong@163.com

Abstract Major ions were analyzed for twenty six groundwater samples collected from diverse aquifers in typical coal mine, northern Anhui province, China. Hydro-geochemical characteristics were analyzed using a statistical approach to identify sources. The results showed that the groundwater solutes are mainly controlled by silicate weathering, otherwise the dissolution of evaporite and carbonate also played a role. Diverse statistical methods such as Fisher discriminant analysis, fuzzy recognition and BP neural networks were used to constitute the water source identification model, and the effective levels of discrimination were 80.8%, 92.3% and 92.3%.
Keywords groundwater, statistical methods, source identification coal mine

Introduction

Deep groundwater always provides some information, which is inherited from the aquifer through water-rock interactions over a long period of time. Many studieshave revealed regional hydro-chemical process characteristics that are important for the evaluation and utilization of groundwater resources (Sun et al. 2011, Ramkumar et al. 2013). In addition, deep groundwater also presents a threat to coal mine safety from potential inrushes, so accurate groundwater source identification could reduce the risk in this regard. Hydro-chemical analyses provide a means for constructing a source identification model for groundwater.

Multivariate statistical analysis methods have been widely used in water source identification, including cluster analysis, factor analysis, discriminant analysis, fuzzy recognition and back propagation (BP) neural networks (Gui et al. 2007, Chen et al. 2013). However, comparative studies of diverse statistical methods used in groundwater source identification are limited.

The purpose of this study is to evaluate groundwater evolution for diverse aquifers, using hydro-chemical and statistical analysis. The major targets are to (1) constitute the groundwater recognition model, (2) compare different statistical methods for their effectiveness at source identification.

Materials and methods

Our study focused on groundwater from the Qidong coal mine, collected from diverse aquifers. The Coal Mine is located in the southern part of the Huaibei coalfield, which constituted by 23 active underground coal mines. Huaibei coalfield is one of the major coalfields in China, being located in the northern Anhui province, China. The basement of coal mine in the district is composed by Archean and Proterozoic metamorphic rock, with cover strata are stable sedimentation between late-Proterozoic and Permian.

A total of twenty-six groundwater samples were collected. The following samples were collected from the following aquifers that occur at the mine and present a threat to mining (Gui et al. 2007): seven samples from Quaternary aquifer (QA), eleven samples from Coal bearing aquifer (CA) and eight samples from Limestone aquifer (LA). The Quaternary aquifer is constituted by yellow mudstone, sandstone and conglomerate, with a depth ranging from 280 m to 300 m. The Coal bearing aquifer is characterized by mudstone, siltstone and sandstone, with a depth between 300 m and 700 m. Limestone aquifers are mainly composed of limestone with clastic rocks, which belong to Taiyuan formation and Ordovician.

Water samples were collected via drainage holes in tunnels, and then filtered through 0.45 μm pore-size membrane into polyethylene bottles that had been cleaned using trace element cleaning procedures. All twenty-six samples were analyzed for the following major ions: Na, Mg, Ca, HCO_3^-, Cl^-, SO_4^{2-}.

Results and discussion

Major ion chemistry

The analytical results for all samples are listed in Table1. In general, The anions Cl^- , SO_4^{2-} and HCO_3^- concentrations of the groundwater range from 21 mg/L to 283 mg/L, 4 mg/L to 2064 mg/L and188 mg/L to 630 mg/L, with means of 211 mg/L, 721 mg/L and 397 mg/L, respectively. The cation Na^+, Ca^{2+} and Mg^{2+} concentrations of the groundwater range from 1.0 mg/L to 940 mg/L, 3.0 mg/L to 539 mg/L, and 7.0 mg/L to 144 mg/L, with averages of 344 mg/L, 147 mg/L and 68 mg/L, respectively.

The QA groundwaters classified as $Na\text{-}SO_4$ or $Na\text{-}HCO_3$ type, however the LA groundwaters are mainly $Ca\text{-}SO_4$ or $Ca\text{-}HCO_3$ types. The concentrations of Na^+ decrease in the order: LA<QA<CA, whereas, conversely, the concentrations of Ca^{2+} and Mg^{2+} increased from CA to LA.

Table 1 *Major solute composition (in mg/L) of groundwater from diverse aquifers in coal mine*

Sample No.	Na^+	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-	Aquifers	Water Type
1	308	112	83	223	586	444	QA	SO_4-Na
2	307	104	78	216	560	440	QA	SO_4-Na
3	289	111	92	227	568	455	QA	SO_4-Na
4	352	66	53	197	398	519	QA	HCO_3-Na
5	405	49	31	201	457	452	QA	SO_4-Na
6	467	3	7	227	273	525	QA	HCO_3-Na
7	259	61	57	201	423	255	QA	SO_4-Na
8	405	51	35	227	428	470	CA	SO_4-Na
9	353	84	57	230	496	396	CA	SO_4-Na
10	383	6	11	188	22	546	CA	HCO_3-Na
11	940	21	17	230	1234	618	CA	SO_4-Na
12	384	48	51	221	396	506	CA	HCO_3-Na
13	274	167	85	248	661	399	CA	SO_4-Na
14	204	203	97	244	659	390	CA	SO_4-Ca
15	206	201	102	249	663	399	CA	SO_4-Ca
16	349	5	9	161	4	630	CA	HCO_3-Na
17	307	24	21	203	130	430	CA	HCO_3-Na
18	210	200	101	249	663	401	CA	SO_4-Ca
19	1	111	38	78	33	296	LA	HCO_3-Ca

Continued Table 1

Sample No.	Na^+	Ca^{2+}	Mg^{2+}	Cl^-	SO_4^{2-}	HCO_3^-	Aquifers	Water Type
20	27	117	16	21	76	321	LA	HCO_3-Ca
21	377	539	141	283	2064	242	LA	SO_4-Ca
22	319	451	135	251	1767	225	LA	SO_4-Ca
23	331	446	144	259	1804	226	LA	SO_4-Ca
24	519	234	111	238	1595	188	LA	SO_4-Na
25	477	202	98	205	1393	263	LA	SO_4-Na
26	487	196	106	204	1398	298	LA	SO_4-Na
Average	343.93	146.64	68.21	210.85	721.07	397.43		

Source identification model

The water source identification model could be constructed based on the diverse hydro-chemical characteristics. Previous studies showed that discriminant analysis, fuzzy recognition and BP neural networks were usually good methods to identify the source and type of groundwater. Thus, this study constituted the source identification model using the diverse statistical methods based on the hydro-chemical data.

Fisher discriminant analysis

Based on presented hydro-chemical data, the Fisher discriminant analysis was carried out using DPS statistical software, and the results are present in functions F1 and F2 and Table 2. Two Fisher discriminant functions were obtained and are expressed as follows:

$$F_1=0.0151[Na^+]+0.0225[Ca^{2+}]+0.0227[Mg^{2+}]-0.057[Cl^-]-0.0025[SO_4^{2-}]-0.0148[HCO_3^-]+9.7094$$

$$F_2=0.0693[Na^+]+0.0871[Ca^{2+}]+0.1264[Mg^{2+}]-0.0367[Cl^-]-0.0335[SO_4^{2-}]-0.0243[HCO_3^-]-3.6548$$

The F_1 and F_2 are the first and second discriminant functions, respectively, thus the value could be calculated by the discriminant function. The groundwater sample could be recognized by comparing the distance between central values and calculating value(table 2). In total, 100% variance was explained by the first and second functions, indicating that the two functions could differentiate between all samples, based on the tested major ions alone. The discriminant results for the aquifers are shown in Table 3. Five samples are incorrectly classified, thus the error rate could be calculated as 19.23%. Overall, the efficiency of Fisher discriminant analysis is considered to be acceptable.

***Table 2** Fisher discriminant equation and various categories of function value*

Parameter	F_1	F_2
Constant	9.7094	-3.6584
Percentage of variance	98.29%	1.71%
QA	-2.1012	-0.6985
CA	-2.4332	0.4251
LA	5.1842	0.0266

Table 3 *Results of discriminate aquifer and actual aquifer for groundwater in Qidong coal mine*

Sample	Actual	Discriminate	Probability	Sample	Actual	Discriminate	Probability
1	1	1	0.68	14	2	1	0.52
2	1	1	0.72	15	2	2	0.52
3	1	1	0.65	16	2	2	0.79
4	1	2	0.51	17	2	2	0.65
5	1	1	0.78	18	2	2	0.51
6	1	2	0.61	19	3	3	1
7	1	1	0.84	20	3	3	1
8	2	1	0.62	21	3	3	1
9	2	2	0.68	22	3	3	1
10	2	2	0.99	23	3	3	1
11	2	2	0.59	24	3	3	1
12	2	2	0.60	25	3	3	1
13	2	1	0.54	26	3	3	1

Fuzzy recognition

The fuzzy recognition model, also established by DPS statistical software, was run for all the samples, with Na^{+}, Ca^{2+}, Mg^{2+}, Cl^{-}, SO_4^{2}, HCO_3^{-} and CO_3^{2-} selected as variables. The maximum value, minimum value and corresponding parameter were calculated. The fuzzy recognition results are shown in Table 4. Of all twenty six water samples only 2 samples discriminate erroneously, an effectiveness of 92.3%. The fuzzy mathematics method is therefore a better discriminant model than Fisher discriminant analysis for this work.

Table4 *Results of fuzzy recognition for groundwater in Qidong coal mine*

No.	Actual	Recognition	No.	Actual	Recognition	No.	Actual	Recognition
1	1	1	10	2	2	19	3	3
2	1	1	11	2	2	20	3	3
3	1	1	12	2	1	21	3	3
4	1	1	13	2	2	22	3	3
5	1	1	14	2	2	23	3	3
6	1	1	15	2	2	24	3	3
7	1	1	16	2	2	25	3	3
8	2	1	17	2	2	26	3	3
9	2	2	18	2	2			

BP neural network

The BP neural network method of source identification was used on the same samples as the foregoing methods, what obtained through DPS software. Samples labeled as being from QA, CA and LA were set to values of 1, 2 and 3 respectively. Thus, for the sample data to be classified, the predicted result could be given through inputting the hydro-chemical data and application the discriminant model. The predicted results are given in fig. 1 Only samples 5 and 12 are false of all twenty-six samples, an effective discrimination rate of 92.3%. As can be seen, the BP neural network identification and fuzzy discriminant methods have the same accuracy rate (92.3%).

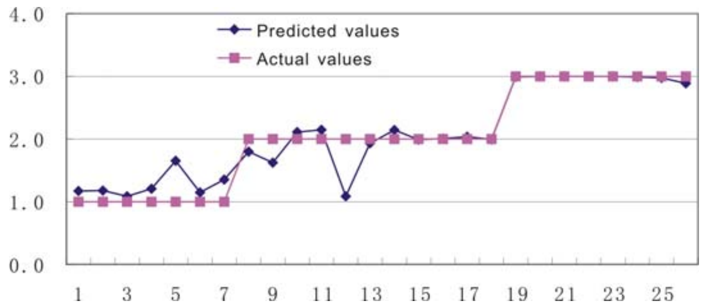

Fig.1 *BP neural network analysis of groundwater in Qidong coal mine*

Conclusions

The concentrations of Cl^- are low for all groundwater samples, whereas those of SO_4^{2-} and HCO_3^- with high concentrations. The cations show similar trends to the anions. Diverse statistical methods such as Fisher discriminant analysis, fuzzy recognition and BP neural network were used to constitute water source identification models, and the effective rates of discrimination were 80.77%, 92.3% and 92.3% respectively.

Acknowledgements

The study was supported by the Anhui province Department of education Natural Science Foundation (KJ2013A249), the Program for Innovative Research Team in Suzhou University (2013kjtd01) and the National Nature Science Foundation of China (40873015).

References

Chen S, Fu XF, Gui HR, Sun LH (2013) Multivariate statistical analysis of the hydro-geochemical characteristics for Mining groundwater: a case study from Baishan mining, northern Anhui Province, China. Water Practice and Technology 8: 131-141

Gui HR, Chen LW (2007) Hydrogeochemistric evolution and discrimination of groundwater in mining district. Beijing: Geological Publishing House

Ramkumar T, Venkatramanan S, Anithamary I, Ibrahim SMS (2013) Evaluation of hydrogeochemical parameters and quality assessment of the groundwater in Kottur blocksTiruvarur district, Tamilnadu, India. Arab J Geosci 6: 101-108

Sun LH, Gui HR, Chen S (2011) Rare earth element geochemistry of groundwater from coal bearing aquifer in Renlou coal mine, northern Anhui Province, China. Journal of Rare Earths 29: 185-192