© IWA Publishing 2016
279
Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

Identifying the potential sources of trace metals in water from subsidence area based on positive matrix factorization

Linhua Sun*, Xianghong Liu and Ning Min

School of Resources and Civil Engineering, Suzhou University, Anhui 234000, China

* Corresponding author. E-mail: sunlinh@126.com

Abstract

Water in the subsidence area is a good choice for solving the water shortage in the coal mining area of China. In this study, positive matrix factorization model has been applied for the concentrations of seven kinds of trace metals (Pb, V, Cr, Mn, Co, Ni and Cu) in water from the subsidence area in the Luling coal mine, northern Anhui Province, China for identifying and quantifying their potential sources. Statistical analyses (including coefficients of variation and *P*-value of Anderson–Darling test) of metal concentrations indicate that multi factors (geological weathering/dissolution, filling of coal gauge and anthropogenic discharge) are responsible for the metal concentrations in the water. Based on the variations of Q values, three sources have been determined by US EPA (US Environmental Protection Agency) positive matrix factorization model: coal gauge filling, geological weathering or dissolution and waste discharge. The contribution degrees of these sources for all of the pools are different, and therefore, different strategies (e.g. clean the waste and the coal gauge around and in the pools) should be applied with different pools.

Key words: coal mining, EPA PMF, subsidence area, trace metals, water

INTRODUCTION

Ground subsidence is a common phenomenon in the surface of the earth, which can be induced by either natural or anthropogenic activities, such as dissolution of limestone in the karst terrains, sub-surface mining, natural gas extraction and earthquake, etc. However, it is considered to be the most serious one among the environmental problems induced by underground coal mining in China during the last decades (Li *et al.* 2007; Yao *et al.* 2010; Yang & Liu 2012): the newly formed subsidence area is up to 130 km^2 every year and the subsidence area has raised up to 700,000 hm^2 by the end of 2006 (Meng *et al.* 2009).

Shortage of water in the current world and China is serious because of the variations of natural conditions and the anthropogenic contamination, and this phenomenon is more serious in coal mining areas: more than 71% of the coal mining areas in China were lacking of water and 40% of them were serious (Gui *et al.* 2011). Because of this situation, most of the studies focused on the underground water in the coal mining area (Chen *et al.* 2013; Sun and Gui 2013; Sun *et al.* 2014). However, the water in the subsidence pools of the coal mining area, have not yet been considered before the year of 2000.

Trace elements, especially the toxic ones (e.g. Hg, Cd, Pb, Zn *et al.*), have long been concerned by scientists because of their special characteristics, including multi kinds of sources, toxicities for human health and difficulties of remediation (Begum *et al.* 2009; Peng *et al.* 2009; Alinnor & Obiji 2010). However, although some studies related to the water quality from the subsidence area have been processed (e.g. Liu & Lu 2009; Sun *et al.* 2015), the work related to the trace metals have not been undertaken.

doi: 10.2166/wpt.2016.030

Therefore, in this study, a total of forty-nine water samples have been collected from the subsidence water area in the Luling coal mine, an old coal mine with large area of subsidence area in northern Anhui Province, China, and the concentrations of seven kinds of trace metals (Pb, V, Cr, Mn, Co, Ni and Cu) have been analyzed. Then, their sources have been calculated by EPA positive matrix factorization (PMF) model.

MATERIALS AND METHODS

Study area

There are five coal mines located near the city of Suzhou: the Zhu Xianzhuang and Luling coal mines in the southeast, and Taoyuan-Qinan-Qidong coal mines in the south. The Luling coal mine is located 20 km southeast to the Suzhou City (GPS: E117 °06′30″, N33 °35′59″) (Figure 1). The annual production of coal is 1.5 million tons and the designed service life is 66 years.

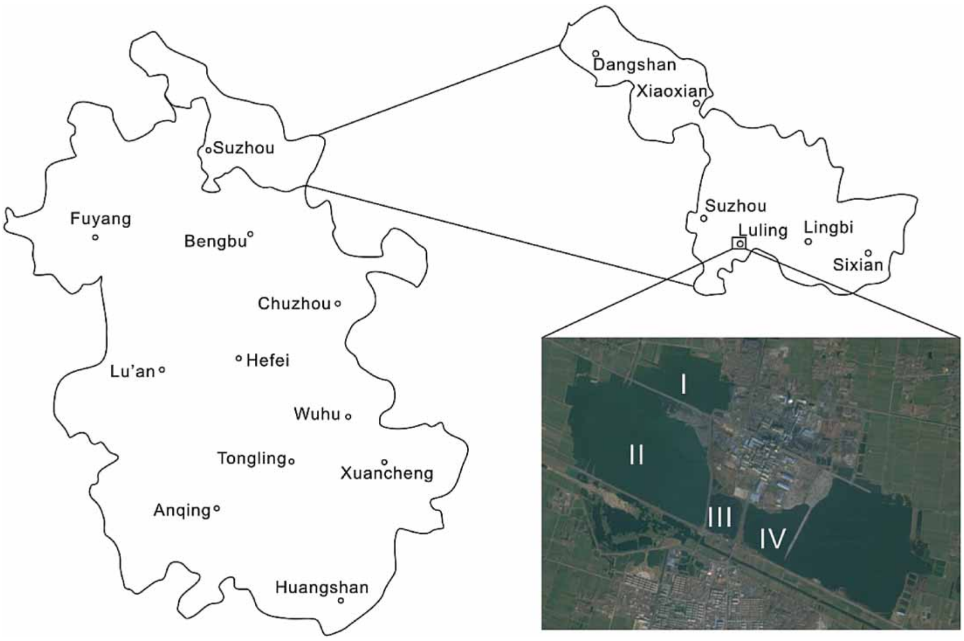

Figure 1 | Location of the study area (I, II, III and IV are sample locations in subsidence water areas around the Luling coal mine).

The main coal seams in the coal mine are 8th and 10th. After coal mining, eight subsidence areas have been formed before 2002, including 81, 82, 83, 84, 86, 88, 101 and 102. And the area has been increased up to 1,000 acres today. Most of the subsidence depth is higher than 7 m, and the water depth is near 3 m (Figure 1). Comparatively, four pools related to this study have different natural conditions: pool I and II have relative high content of water, whereas pool III is a closed system without hydraulic connection with other pools and, surrounded by human residence. Moreover, coal gauge filling is common in all of the pools.

Analytical methods

A total of forty-nine water samples have been collected from the subsidence water areas in the Luling coal mining area (13, 14, 12 and 10 samples from pools I, II, III and IV in the Figure 1, respectively). Water

Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

samples were filtered through 0.45 μm pore-size membrane and collected into a 2.0 L polyethylene bottles that had been cleaned in the laboratory, and immediately acidified to pH < 2 by HNO_3 for prevention of element precipitation and/or adsorption by the bottle. Then the samples were sent to the laboratory for analysis in 24 hours. Analytical processes were taken place in the Engineering and Technology Research Center of Coal Exploration in Anhui Province, China. Concentrations of seven kinds of trace metals (Pb, V, Cr, Mn, Co, Ni and Cu) have been analyzed by Inductively coupled plasma mass spectrometry (ICP-MS) (Element II). The international standard solutions after dilution have been applied for calibration, and most of the samples have relative standard deviations less than 10%.

PMF

PMF is a Principal Component Analysis (PCA)-based receptor model with non-negative constraints that involve solution of quantitative source apportionment equations by oblique solutions in a reduced dimensional space (Paatero & Tapper 1994; Paterson *et al.* 1999). The PMF model defines the concentration matrix of chemical species measured at receptor sites as the product of source composition and contribution factor matrices matrix with a residue matrix. It has long been used for quantifying the source contributions in environmental studies. In this study, the model created by US EPA (PMF 5.0) has been applied, and the detailed information about the software can be found in Norris & Duvall (2014).

RESULTS AND DISCUSSIONS

Descriptive statistics

Analytical results of this study are synthesized in Tables 1 and 2. As can be seen from the Table 1, the mean concentrations of the water samples are Mn > V > Cu > Ni > Cr > Co > Pb. Their mean concentrations are 4.777 μg/l, 3.731 μg/l, 0.885 μg/l, 0.648 μg/l, 0.127 μg/l, 0.092 μg/l and 0.046 μg/l, respectively. However, water samples from the different pools have different concentrations of metals: water samples from the pool IV have the highest mean concentrations of most of the trace metals except for Mn, whereas the water samples from the pool III have the highest mean concentrations of Mn. Moreover, most of the lowest mean concentrations of the trace metals are observed in the pool II. This phenomenon might be, induced by the different natural and anthropogenic conditions.

Table 1 | Descriptive statistics of metal concentrations (μg/l)

	Pb	V	Cr	Mn	Co	Ni	Cu
N	49	49	49	49	49	49	49
Minimum	0.03	1.256	0.043	0.061	0.059	0.271	0.067
Maximum	0.089	8.737	0.239	95.108	0.192	1.242	1.638
Median	0.044	2.585	0.108	0.241	0.076	0.602	0.91
Mean	0.046	3.731	0.127	4.777	0.092	0.648	0.885
Coefficient of Variation	0.278	0.591	0.469	3.93	0.41	0.263	0.371
P-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.052

Previous environmental studies revealed that coefficient of variation is a good indicator for revealing the contribution of chemical constitutes in natural environment from anthropogenic activities: a low coefficient of variation (<0.1) indicates the low degree of anthropogenic contribution, whereas a

Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

Table 2 | Metal concentrations of water from the subsidence areas (μg/l)

	Pb	V	Cr	Mn	Co	Ni	Cu
I ($n = 13$)	0.045	2.48	0.114	0.202	0.077	0.662	0.991
II ($n = 14$)	0.042	2.54	0.061	0.200	0.070	0.570	0.898
III ($n = 12$)	0.045	3.01	0.162	17.7	0.075	0.511	0.430
IV ($n = 10$)	0.052	7.88	0.192	1.61	0.163	0.906	1.274

high coefficient of variation (>0.9) indicates high degrees of anthropogenic contribution. In this study, metal concentrations of the water samples have coefficients of variations range from 0.278 to 3.93, which indicates that all of them are statistically inhomogeneous and cannot be contributed by a single natural source, especially the metal Mn with coefficient of variation (>0.9). This consideration can also be supported by their *P*-values of Anderson–Darling test, as only Cu have *P*-values higher than 0.05 (Table 1), which means that all of the metal concentrations (except for Cu) cannot pass the normality test, and implying that they have multi sources.

Such a phenomenon is consistent with the truth that there are multi factors can affect the environment of the subsidence pools in this study, e.g. geological weathering/dissolution of natural rocks and/or sediments, anthropogenic waste discharge around the pool and the application of coal gauge for filling of subsidence areas.

The number of sources

Before the modeling of EPA PMF model, all of the data were firstly checked by concentration/ Uncertainty model. The results suggest that all of the seven kinds of trace metals have been classified to be strong category, which suggest that all of them can be used for PMF analysis.

Determination of the factor numbers in factor analysis is an important work, and several principles have been used in previous studies, such as the Kaiser criterion (Kaiser 1960), the screen test (Cattell 1966) and the Exner function (Exner 1966), and now, Eigen-value higher than one is the most popular applied criterion. In PMF, the determination of the number of factors is also complicated, because too few factors can combine different sources together, whereas too many factors will make a real factor further dissociate into two or more non-existing sources (Lee *et al.* 1999).

In this study, the variations of Q values (including Qtrue and Qrobust) in response to factor numbers have been applied for solving this issue (Figure 2). As can be seen from the figure, when set the number of factors to be two, the Q_{true} and Q_{robust} have similar values (7.9). However, the Q values are dropping drastically when the number is set to be three (Q value = 0.2), and then stable when the numbers are set to be four (Q value = 0). Moreover, when the number of factors set to be three,

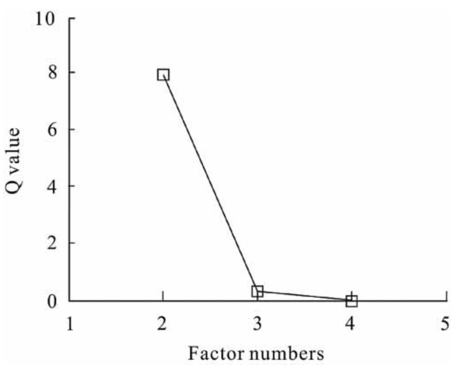

Figure 2 | Q value versus factor numbers.

Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

the coefficients between observed and predicted data of all the trace metals range from 0.190 (for Pb) to 1.00 (for V and Mn), and except for Pb, other six kinds of trace metals have coefficients higher than the critical coefficient ($r_\alpha = 0.365$, $\alpha = 0.01$, $n = 49$, Table 3). When considering of four factors, the coefficient of Pb is also lower than the critical coefficient. Therefore, three sources (factors) are considered to be reasonable, and the following discussion is based on the three factor results.

Table 3 | Coefficients between observed and predicted data

	Pb	V	Cr	Mn	Co	Ni	Cu
r	0.190	1.000	0.597	1.000	0.964	0.911	0.977

Note: $r_\alpha = 0.351$ ($\alpha = 0.01$), $r_\alpha = 0.271$ ($\alpha = 0.05$).

Source profiles

The analytical results of PMF base model run are shown in Figure 3, and the statistical results are synthesized in Table 4. As can be seen from the figure and table, the Factor 1 is dominated by V and

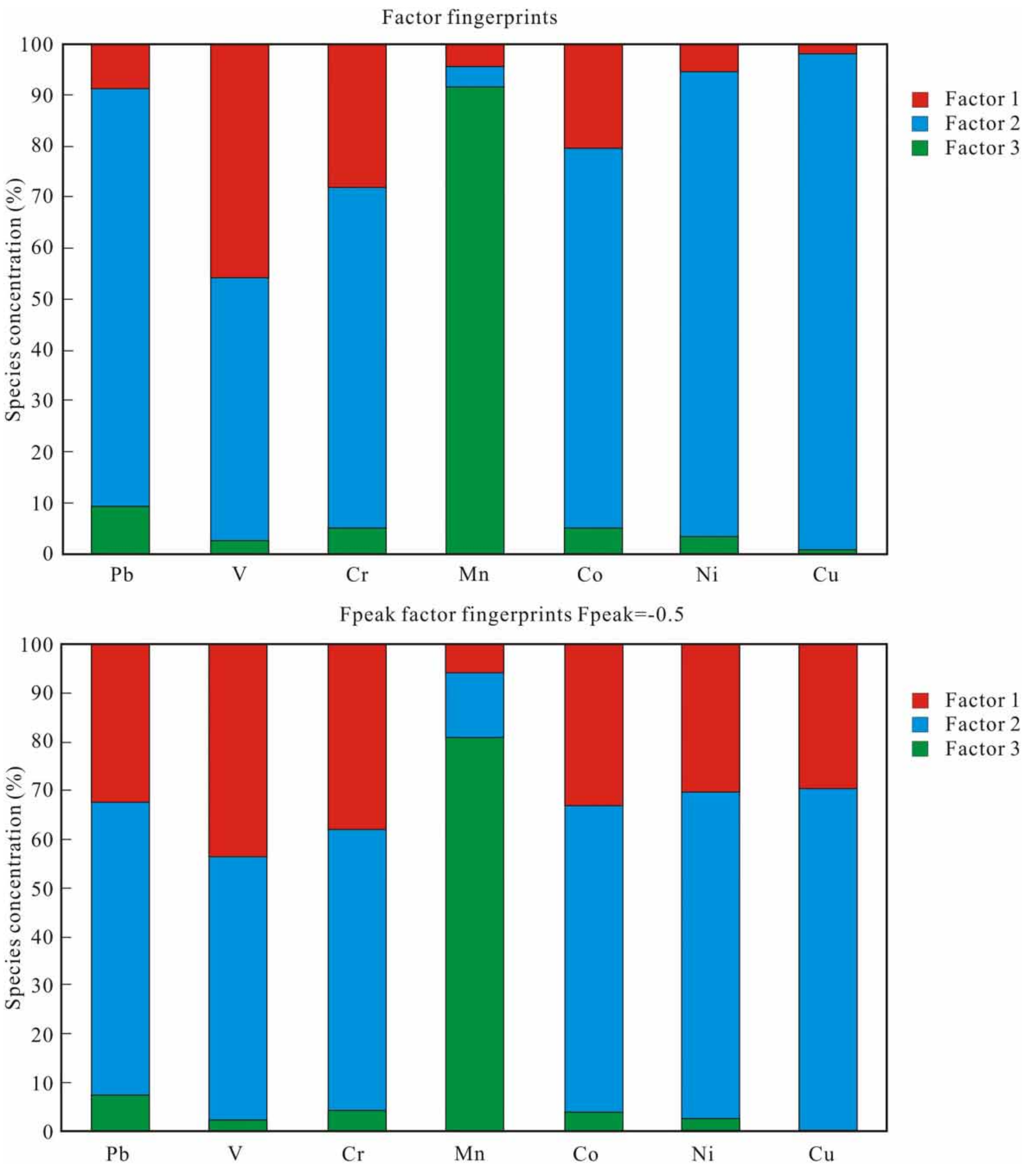

Figure 3 | Factor fingerprints (upper and lower are based on base model and Fpeak rotation, respectively).

Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

Table 4 | Factor contributions (%) for trace metals

	Pb	V	Cr	Mn	Co	Ni	Cu
Factor 1#	8.6	46.2	28.2	4.7	20.4	5.8	1.9
Factor 2#	82.4	51.4	67.0	4.0	75.0	91.2	97.5
Factor 3#	9.0	2.4	4.8	91.3	4.6	3.0	0.6
Factor 1*	32.3	43.6	38.1	5.6	33.4	30.1	29.7
Factor 2*	60.1	54.2	57.5	13.4	62.7	67.3	70.3
Factor 3*	7.5	2.2	4.4	81.0	3.9	2.6	–

Note: # and * mean base model run and after Fpeak (−0.5) rotation run.

followed by Cr, Co, Pb, Ni, Mn and Cu, the contributions of this factor for them are 46.2%, 28.2%, 20.4%, 8.6%, 5.8%, 4.7% and 1.9%, respectively. The factor 2 is dominated by Cu, Ni, Pb and followed by Co, Cr, V and Mn, the contributions for them are 97.5%, 91.2%, 82.4%, 75.0%, 67.0%, 51.4% and 4.0%, respectively. As to the Factor 3, it is dominated by Mn and followed by other trace metals, the contribution for Mn is 91.3%, and others are below 10%.

For better understanding about each factor, the Fpeak rotation has been processed for the analysis with Fpeak strength between −1.0 and 1.0 (step = 0.5). The results suggest that with Fpeak strength equal to −0.5, the Qrobust and Qtrue is the smallest, 49.5 and 2.6, respectively, with coveraged to be 'yes'. The analytical results with Fpeak strength equal to −0.5 are also shown in Figure 3 and synthesized in Table 4 for comparison. As can be seen from the figure and table, the contributions of each factor for each metal have been changed relative to the base model run. Factor 1 is dominated by V and followed by Cr, Co, Pb, Ni, Cu and Mn, Factor 2 is dominated by Cu, Ni, Co, Pb and followed by Cr, V and Mn, whereas Factor 3 is dominated by Mn and followed by Pb, Cr, Co, Ni and V.

The first factor should be considered to be the source related to coal gauge in the study area, because samples with high contributions from this source is located in the pool III and IV (Figure 4), especially after Fpeak rotation. These two pools are all filled by coal gauge with low content of water storage. Moreover, more than 91.8% of the samples have contributions from this source, although with varying contributions as revealed by Figure 4, that's because all of the pools are filled by coal gauge.

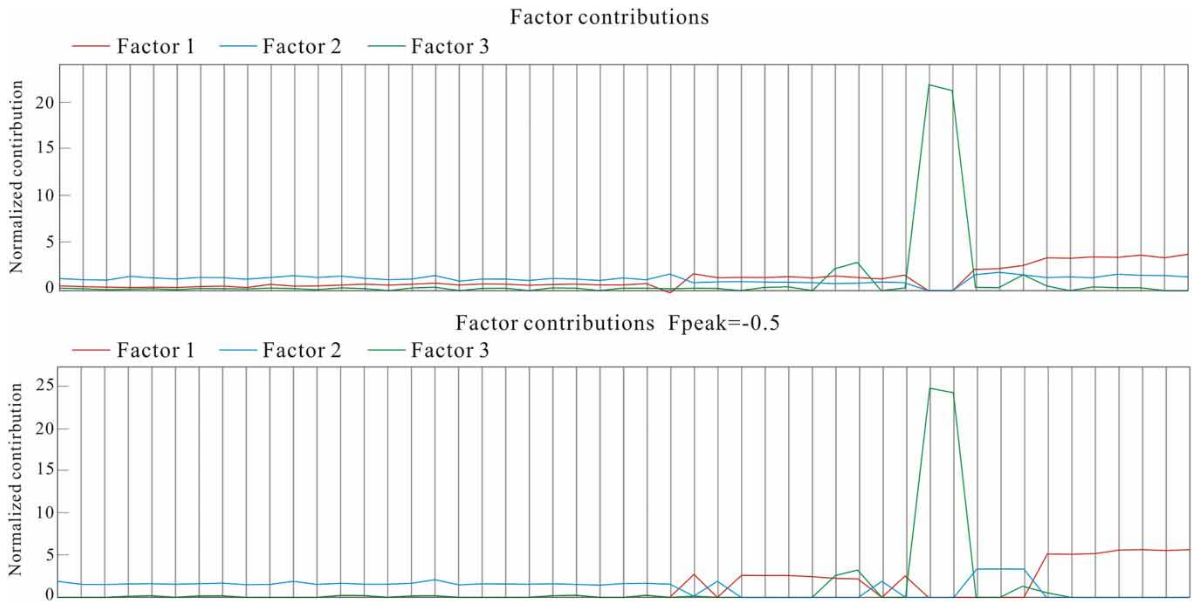

Figure 4 | Factor contributions (upper and lower are based on base model and Fpeak rotation, respectively. Sample 1–13, 14–27, 28–39 and 40–49 are collected from pools I, II, III and IV, respectively).

Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

The second factor can be explained to be the source related to geological weathering or dissolution. That's because it has the highest mean contributions for all of the samples (0–100%, mean = 56.4%), and its contribution is relative stable and high (65.9–100%) in the pool I and II (Figure 4), where have not been dramatically affected by human activities because of their high contents of water storage. Moreover, Cu is one of the dominant metal of this factor, and it is the only one can pass the normality test with *P*-value higher than 0.5 (Table 1).

The third factor, as revealed by Figures 3 and Table 4, is considered to be the source related to waste discharge, because it is dominated by Mn, which is considered to be main pollution related to waste discharge except for N, P and Fe (Yan 2008). It can also be demonstrated by the coefficient of variation, as Mn has coefficient of variation up to 3.93, which indicates that it has high degree of spatial inhomogeneous, which is considered to be an indicator of anthropogenic contribution. It can also be observed in Figure 4 that about 28.6% of the samples are contributed by this factor, but most of them are collected from the pool III and IV, where are rounded by human residence (Figure 1).

Source contributions

The source contributions calculated by EPA PMF are shown in Figure 5. As can be seen from the figure, water samples from different pools have different contributions from the three sources: the contributions from the source 1 (coal gauge), 2 (geological weathering or dissolution) and 3 (waste discharge) for water samples from the pool I are 0–18.9% (mean = 13.3%), 75.5–100% (mean = 85.6%) and 0–10.5% (mean = 1.08%), respectively, whereas their mean contributions for the water samples from the pool II are 26.6%, 72.8% and 0.67%, respectively. The water samples from the pool III are mainly contributed by source 1 (45.8%), and followed by source 2 (27.1%) and 3 (27.1%). As to the water samples from the pool IV, source 1 is the main contributor (65.5%) and followed by source 2 (30.4%).

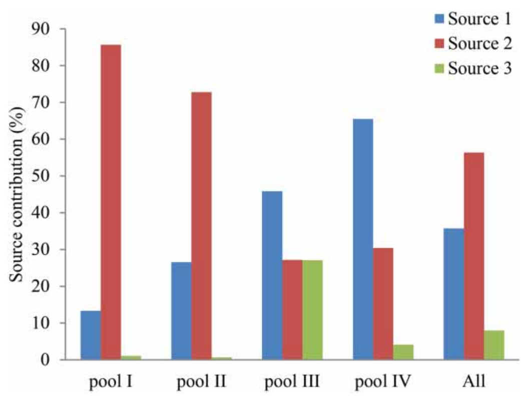

Figure 5 | Source contributions.

Such a phenomenon reflects the different natural and anthropogenic conditions of these four pools. The waste discharge has highest degree of impact on the pool III, that's because it is surrounded by human residence. The coal gauge source has contributions for all of the samples, but the contributions for the pool IV and III are relative obvious, because these two pools are filled by coal gauge, and the water storage in them are relative low. The geological weathering or dissolution source has highest contributions for the water in the pool I and II, high content of water is the main factor, because high content of water can accelerate the weathering and dissolution process, and dilute the pollution from anthropogenic activities. Therefore, before application of the water in these subsidence pools, different strategies (e.g. clean the waste and the coal gauge around and in the pools) should be applied.

CONCLUSIONS

Based on the application of EPA PMF model for the concentrations of seven kinds of trace metals (Pb, V, Cr, Mn, Co, Ni and Cu) in water from the subsidence area in the Luling coal mine, northern Anhui Province, China, the following conclusions have been obtained:

1. The water samples have mean concentrations of Mn > V > Cu > Ni > Cr > Co > Pb, and the metal concentrations among different pools are different with each other, which suggest that they have different natural and anthropogenic conditions.
2. Statistical analysis, including coefficients of variation and Anderson–Darling test of metal concentrations indicate that they are affected by multi factors (geological weathering/dissolution, coal gauge filling and anthropogenic discharge).
3. Based on the variations of Q values, three sources have been determined by EPA PMF model: coal gauge filling, geological weathering or dissolution and waste discharge, and their contributions are different for the water chemistry in different pools.
4. According to the analytical results, different strategies (e.g. clean the waste and the coal gauge around and in the pools) should be applied.

ACKNOWLEDGEMENTS

This work was financially supported by National Natural Science Foundation of China (41302274), the Foundation of Scholarship Leaders in Suzhou University (2014XJXS05).

REFERENCES

Alinnor, I. J. & Obiji, I. A. 2010 Assessment of trace metal composition in fish samples from Nworie River. *Pakistan Journal of Nutrition* **9** (1), 81–85.

Begum, A., Ramaiah, M., Khan, I. & Veena, K. 2009 Heavy metal pollution and chemical profile of Cauvery River Water. *Journal of Chemistry* **6** (1), 47–52.

Cattell, R. B. 1966 The screen test for the number of factors. *Multivariate Behavioural Research* **1**, 245–276.

Chen, S., Fu, X. H., Gui, H. R. & Sun, L. H. 2013 Multivariate statistical analysis of the hydro-geochemical characteristics for Mining groundwater: a case study from Baishan mining, northern Anhui Province, China. *Water Practice & Technology* **8**, 131–141.

Exner, O. 1966 Additive physical properties. I. general relationships and problems of statistical nature. *Collection of Czechoslovak Chemical Communications* **31** (8), 3222–3251.

Gui, H. R., Yao, E. Q., Song, X. M., Yuan, Z. H. & Zhang, G. 2011 *Research on Recycling Technology of Coalmine Water*. China University of Mining and Technology Press, Xuzhou.

Kaiser, H. F. 1960 The application of electronic computers to factor analysis. *Educational and Psychological Measurement* **20**, 141–151.

Lee, E., Chan, C. K. & Paatero, P. 1999 Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. *Atmospheric Environment* **33** (19), 3201–3212.

Li, X. P., Wang, Y. L., Zhang, C. L. & Mi, J. 2007 The study of coal mining induced environmental and geological problems and solutions. *Soil Engineering and Foundation* **21** (2), 70–72.

Liu, F. & Lu, L. 2009 Progress in the study of ecological restoration of coal mining subsidence areas. *Journal of Natural Resources* **24** (4), 612–620.

Meng, Q. J., Feng, Q. Y., Wu, Q. Q., Meng, L. & Cao, Z. Y. 2009 Distribution characteristics of nitrogen and phosphorus in mining induced subsidence wetland in Panbei coal mine, China. *Procedia Earth and Planetary Science* **1**, 1237–1241.

Norris, G. & Duvall, R. 2014 EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide (EPA/600/R-14/108). U.S. Environmental Protection Agency.

Paatero, P. & Tapper, U. 1994 Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **5** (2), 111–126.

Paterson, K. G., Sagady, J. L., Hooper, D. L., Bertman, S. B., Carroll, M. A. & Shepson, P. B. 1999 Analysis of air quality data using positive matrix factorization. *Environmental Science & Technology* **33** (4), 635–641.

Water Practice & Technology Vol 11 No 2
doi: 10.2166/wpt.2016.030

Peng, J. F., Song, Y. H., Yuan, P., Cui, X. Y. & Qiu, G. L. 2009 The remediation of heavy metals contaminated sediment. *Journal of Hazardous Materials* **161** (2), 633–640.

Sun, L. H. & Gui, H. 2013 Groundwater quality and evolution in a deep limestone aquifer, northern Anhui Province, China: evidence from hydrochemistry. *Fresenius Environmental Bulletin* **22** (4 A), 1126–1131.

Sun, L., Gui, H. & Peng, W. 2014 Heavy metals in groundwater from the Wolonghu coal mine, northern Anhui Province, China and their hydrological implications. *Water Practice & Technology* **9** (1), 79–87.

Sun, L. H., Liu, X. H. & Cheng, C. 2015 Quality of water from subsidence area: a case study in the Luling coal mine, northern Anhui Province, China. *Water Practice & Technology* **10** (4), 777–786, doi:10.2166/wpt.2015.096.

Yan, X. S. 2008 Shallow groundwater pollution near a landfill site and water environmental health risk assessment (Doctoral dissertation). Hefei University of Technology, Hefei.

Yang, Q. & Liu, W. S. 2012 The study on land reclamation technology and ecological control for coal mining subsidence area. *Advanced Material Research* **446–449**, 2973–2977.

Yao, D. X., Meng, J. & Zhang, Z. G. 2010 Heavy metal pollution and potential ecological risk in reclaimed soils in Huainan mining area. *Journal of Coal Science and Engineering* **16**, 316–319.