BAJOPAS Volume 11 Number 2 December, 2018 **http://dx.doi.org/10.4314/bajopas.v11i2.21**

Bayero Journal of Pure and Applied Sciences, 11(2): 169 - 175
Received: March, 2018
Accepted: November, 2018
ISSN 2006 – 6996

SOURCE IDENTIFICATION AND EVALUATION OF SURFACE WATER QUALITY USING FACTOR AND DISCRIMINANT ANALYSIS

Isa Baba Koki[1*], Abubakar Lawal[2] and Syed Noeman Taqui[3]
[1]Department of Chemistry, Yusuf Maitama Sule University Kano, PMB 3220 Kano, Nigeria
[2]Department of Chemistry, Umaru Musa Yar'adua University Katsina, PMB 2218 Katsina, Nigeria
[3]Department of Chemistry, University Malaya, Kuala Lumpur 50603, Malaysia
*Correspondence author: isakoki@yahoo.com, 08039092875

ABSTRACT
Abstract
This study utilizes factor and discriminant analysis to identify the parameter source and evaluate the quality of water from ex-mining ponds and lakes in Selangor. Factor analysis (FA) which explain 83.77% of the surface water quality variation shows that As and Cd that mostly originated from mining activity, and pH are the parameters responsible for the major variation in the surface water quality and were strongly associated with varimax factor 1 (VF1), while Pb, Mn and DO were associated with varimax factor 2 (VF2). The Discriminant analysis (DA) reveals that As, Cd, Mn, Fe and pH are the parameters that significantly differentiate ex-mining ponds from the lake ($p < 0.05$ and F - ratio >> 1), and supported by the correlation study. The elevated metal concentrations in ex-mining ponds compared to lakes were plausibly associated with the past mining operation, thereby indicating heavy metal dominance in ex-mining ponds. DO and BOD are associated with the anthropogenic input from residential sources. Findings of this study therefore show the need and usefulness of multivariate statistical analysis to get information on the quality status of surface water.
Keywords: Discriminant analysis, Ex-mining pond, Factor analysis, Lake, Water quality.

INTRODUCTION

Surface water is the main source of water for domestic and industrial uses in many countries of the world thereby supporting human lives and facilitates economic developments (Gleick 2003). However, the quality of water is of much concern in the recent time due to continuous increase in water contamination as a result of rapid increase in human and industrial activities. Mining is one of the activities that deteriorate the quality of surface water, generating contaminants that are difficult to handle and of health concern in an environment. This renders the water unfit for any beneficial purposes.

The contaminants of concern in mining are heavy metals such as Cd, Pb, As and Mn reported at elevated concentrations in lakes and adjacent rivers (Rojas and Vandecasteele, 2007; Acheampong *et al.,* 2013). These metals originated from oxidation of the sulphide mineral ores such as arsenopyrite (FeAsS) and greenockite (CdS) in the presence of water and oxygen, and are further liberated in higher concentration under acidic pH generated (Koki *et al*., 2017; Low *et al*., 2016). Heavy metal contaminants have the characteristic of high toxicity and mobility, and are difficult to manage in natural environment (Ning *et al*., 2011). The high organic matter associated with mining operation depletes oxygen level in surface water there by endangering aquatic lives (Onichandran *et al*., 2013). Considering the contamination associated with long-years of mining activity in Selangor Malaysia, water quality of ex-mining ponds needs to be studied and evaluated to ascertain the levels and distribution of relevant parameters. This could be helpful in providing the scientific basis for proper management and future uses of the surface waters.

Large water quality data comprising of numerous parameters can accurately be analyzed and evaluated using multivariate statistical analysis. Unlike the use of conventional descriptive analysis to evaluate water quality with many limitations such as, lack of precise source apportionment and long term correlation among parameters. The use of combined chemometric approach in this study is pertinent to evaluate the relationship among the parameters and water pollution source apportionment, and to precisely discriminate the ex-mining ponds from lakes. FA analyzes large number of variables in terms of their common underlying dimensions. This constructs a small number of factors that are linear combinations of the original variables (Rogerson 2006). FA has been used to identify the pollution sources in the study areas. It identifies the latent factor that explains the major variation in the entire data set (Mustapha *et al*., 2013).

DA maximizes the similarities in variances between-groups relative to the within-group (Koklu *et al*., 2010). It is used to sort out the most significant parameters that result in water quality variation among the studied sites (Juahir *et al*., 2010). The objectives of this study were to study the distribution pattern and source of the contaminants, and to identify the dominant parameters that explain the overall water quality differences among the sites using selected chemometric methods.

MATERIALS AND METHODS

Study Area

Selangor (3°20′0″N, 101°30′0″E) is the most populous state in Malaysia with 7.5 million people, it is the industrial and commercial region of Malaysia (Figure 1). The geology of Selangor involves different type of rocks, but mostly dominated by the Kenny hill formation, especially in Puchong district which is an area with abundant ex-mining ponds. There is deposit of limestone basement predominantly around Kuala Lumpur. Selangor state was one of the major producer of tin in Malaysia, reaching up to 22% of the total Malaysia's tin output. There exist abundant ex-mining ponds in Selangor due to the intensive mining activity with about 4909.6 hectares of ex-mining land (Althuwaynee *et al*., 2012; Morgan, 1968). There are also natural lakes that had no mining activity which are mostly utilized for recreational activities. Description of the sampling sites and their coordinates are shown in Table 1.

Figure 1: Map of the study area showing sampling sites in Selangor, Malaysia

Table 1: Sampling Sites in Selangor, Malaysia.

Type	Sites	Code	Location
Ex-mining Ponds	Kelana Jaya	KJA	N 03°05′35.4″ E 101°35′53.2″
	Putra Perdana	PPA	N 02° 57' 27.8" E 101° 36' 52.2"
		PPB	N 02° 57' 46.4" E 101° 36' 21.8"
	Prima Perdana	PRP	N 02° 59' 10.0" E 101° 35' 49.4"
Lakes	Taman Jaya	TJA	N 03° 06' 17.8" E 101° 38' 53.0"
	Perdana	PDL	N 03° 08' 31.9" E 101° 41' 06.5"
	Shah Alam	SAL	N 03° 04' 27.0" E 101° 30' 47.3"

Sampling and Sample Preparation

Water samples were collected from ex-mining ponds and lakes in Selangor. The samples were collected from a 25 cm depth using a Wildco water sampler with nine samples per site for analysis. Conc. HNO_3 (Merck Suprapur) was used to acidify all the water samples to preserve the samples by preventing metals sedimentation in the container (Sadeghi *et al.*, 2012). The samples were stored in acid washed polyethylene bottles and transported to the laboratory in ice boxes conserved at 4°C for metals analysis. Physico-chemical parameeters such as dissolved oxygen (DO) and pH were measured and recorded in the field using portable YSI Pro multi parameter water quality meter, and biological oxygen demand (BOD) was also measured *in-situ* using a portable modern water meter to avoid changes in the bacterial concentration with time. The water sample were filtered using 0.45µm PTFE filters before the metal analysis using inductively coupled plasma-mass spectrophotometry (ICP-MS). The filtered water samples were subsequently analysed for Cd, As, Mn, Pb and Fe, with ICP-MS 7500ce (Agilent Scientific Technology Ltd., USA).

Quality Control

Blank and certified reference materials (CRM) were checked after every ten samples to demonstrate the validity of the previous runs. All analyses were carried out in triplicates and the results were expressed as 95% confidence interval of the mean in µg/L. The R^2 values (coefficients of determination) for ICP-MS calibration curves were all close to 1.0. The CRM shows a good agreement with the certified values with analyte recoveries found to be within the acceptable ranges.

Data Analysis

Multivariate statistical analysis and parameter correlation analyses were carried out using JMP Pro 12 to study the parameter variations among the sampling sites, and to study the relationship among the studied parameters.

RESULTS AND DISCUSSION

The results of statistical summary of selected water quality parameters (Mean and standard deviation) are presented in Table 2. World Health Organisation (WHO) and drinking water quality standard for raw water (DWQSRW) were used as reference standard to evaluate the contamination level. The results show a wide variation in the study areas with mean As concentrations ranging from 1.13 to 116 µg/L. High As concentration in some of the ex-mining ponds above WHO limits (WHO 2011) and DWQSRW (MOH 2004) limit of 10 µg/L may be associated with the past mining operation. Low As concentration in ex-mining pond KJA could be linked to the continuous flow of rain water and domestic effluent (Yap *et al.*, 2007). The mean Cd concentrations were between 0.02 to 12.9 µg/L, low Cd concentrations were observed in the lakes. However, ex-mining ponds record higher concentrations above drinking water Cd recommended value of 3 µg/L except KJA. High Cd concentrations could be much related to the previous mining activity. The concentrations of Pb, Mn and Fe were below the recommended values except Fe in KJA ex-mining pond with concentration of 1166 µg/L which is above the DWQSRW of 1000 µg/L (MOH 2004). High Fe concentration in KJA compared to other studied sites could be related to the flow of domestic effluent.

Table 2: Univariate statistical summary of the selected water quality parameters

	Code	As	Cd	Pb	Mn	Fe	pH	DO	BOD
Ex-mining Ponds	PRP	42.0 ± 0.5	12.19 ± 0.01	3.7 ± 0.3	3.3 ± 0.1	15.1 ± 0.1	8.23 ± 0.09	0.47 ± 0.04	2.77 ± 0.07
	KJA	7.5 ± 0.2	0.03 ± 0.01	0.23 ± 0.03	124 ± 3	1166 ± 11	0.93 ± 0.09	0.91 ± 0.07	5.9 ± 0.4
	PPA	11.3 ± 0.4	12.9 ± 0.2	4.5 ± 0.3	125 ± 2	460 ± 4	6.1 ± 0.2	6.1 ± 0.1	8 ± 1
	PPB	116 ± 2	12.23 ± 0.02	3.6 ± 0.2	1.4 ± 0.3	35.8 ± 0.6	10.00 ± 0.09	0.38 ± 0.03	2.8 ± 0.3
Lake	TJA	1.13 ± 0.06	0.33 ± 0.02	0.52 ± 0.02	186 ± 1	151 ± 4	7.9 ± 0.2	3.3 ± 0.2	7.5 ± 0.2
	PDL	3.8 ± 0.1	0.06 ± 0.01	0.36 ± 0.01	35.7 ± 0.3	228 ± 6	8.32 ± 0.02	3.6 ± 0.1	3.54 ± 0.07
	SAL	3.37 ± 0.2	0.02 ± 0.01	< 0.05	68 ± 2	279 ± 6	8.49 ± 0.03	2.8 ± 0.1	8.85 ± 0.08
Water Standards	WHO	10	3	10	400	-	6.5 - 8.5	4 - 10	< 5
	DWQSRW	10	3	50	200	1000	5.5 - 9.0	5 - 7	6

Metal concentrations are given in µg/L, DO and BOD in mg/L, and pH no unit.
DWQSRW – Malaysia Drinking Water Quality Standard for Raw Water
WHO – World Health Organization Drinking Water Standard

The mean values of pH, BOD and DO are 0.93 to 10.00, 2.77 to 8.85 mg/L, and 0.47 to 6.1 mg/L respectively. These values indicate the influence of mining and other anthropogenic activities around the studied sites some of which are surrounded by residential houses. The pH of lakes are within the acceptable limit of 6.5-8.5, while the ex-mining ponds are not within the acceptable limit except PRP with pH value of 8.23. Depending on the mineralogy of the study area, ex-mining ponds could be acidic or basic (Wolkersdorfer 2008). BOD and DO are parameters indicating quantity of oxygen consumed due to microbial decomposition of organic matter, and the quantity of dissolved oxygen present in surface water respectively; they are indicators of organic pollution. BOD concentrations in KJA, PPA, TJA and SAL are above the recommended values of 5 mg/L and 6 mg/L respectively for WHO and DWQSRW. This could be attributed to the domestic discharge and surface run-off. The levels of DO in all the studied sites are less than acceptable limit of 4 - 10 mg/L and 5 – 7 mg/L respectively for WHO and DWQSRW except an ex-mining pond PPA. This confirms the high level of organic matter influence on the surface water quality of the areas under study.

Factor Analysis

The result of factor analysis employed in this study reveals variation in water quality parameters (Figure 2). Three varimax rotated factors (VF) with eigenvalue > 1 represent 83.77% of the total variation in the data set (Table 3). The sources of pollution were apportioned considering absolute factor loading > 0.75 indicating strong loading, 0.75 – 0.50 moderate loading and 0.50 - 0.30 weak loading (Crowther *et al.*, 2001; Liu *et al.*, 2003). VF1 accounts for 41.68% of the total variance in the entire data set with a strong positive loading on As and pH, and moderate loading on Cd, Mn, Fe and DO. This factor contains variables that are linked to dissolution of mineral ores from mining operation (Navarro *et al.*, 2008).

The high positive loading on pH and DO indicates variation in acidity and alkalinity, and presence of organic matter in the surface water samples. Moderate negative loading on Fe indicates less influence of natural background sources. VF2 consist of strong positive loading on Pb and moderate loading on Mn which are associated with mining and dissolution of rocks. Pb is present in mineral ores as an impurity and gets into the solution during mining operation (Bao *et al*., 2016). Manganese is abundant and distributed on the earth's crust in association with ores of iron (Muthaiah *et al*., 2016). VF3 shows moderate loading on BOD, DO, pH, and negative loading on Fe indicating organic matter influence on the water quality at the studied sites.

Discriminant Analysis

The results of discriminant analysis using multivariate analysis of variance (MANOVA) as shown in Table 4 shows larger F – ratio (much greater than 1), and the corresponding p – value, which represent the significant difference ($p < 0.05$) among the studied parameters. It is clear that pH has the largest F – ratio of 55.601 ($p = 0.0000001$) indicating much differences in acidity between ex-mining ponds and lakes in Selangor. Fe with F – ratio = 41.481 and p = 0.0000012 shows significant difference among the sites indicating contribution of high Fe concentration from mining operation beside the natural source (Madzin *et al*., 2015).

Table 3: Loadings of parameters on significant VFs for ex-mining ponds and lakes

	VF1	VF2	VF 3
As	**0.858193**	0.119333	-0.318400
Cd	**0.658575**	0.271156	-0.104374
Pb	0.213707	**0.931001**	-0.236102
Mn	**0.724760**	**0.622419**	-0.073688
Fe	**-0.678886**	-0.427528	**-0.531342**
DO	**0.601528**	0.463570	**0.505126**
BOD	-0.463923	-0.442354	**0.568237**
pH	**0.765603**	0.204489	**0.596657**
Eigen Value	3.3	1.9	1.3
Variability (%)	41.68	24.71	17.37
Cumulative (%)	41.68	66.40	83.77

Values in bold are significant

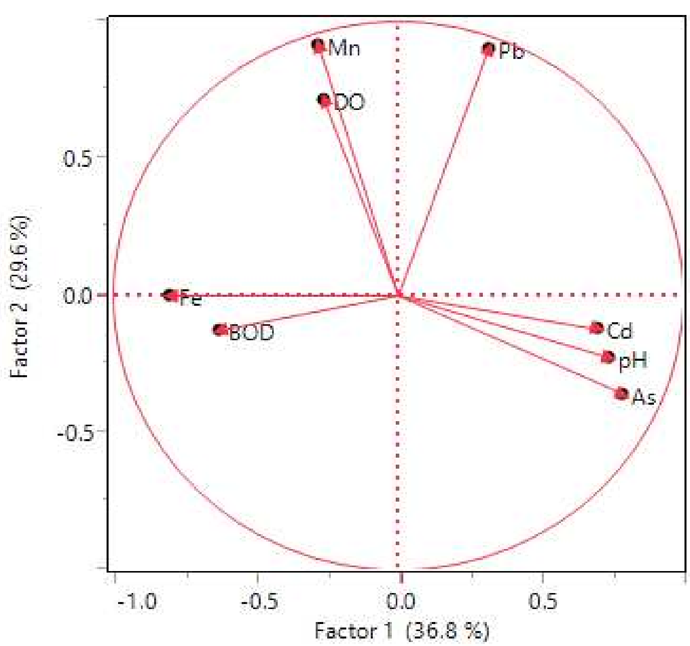

Figure 2: Factor loading plot of parameters under study

Among the toxic heavy metal pollutants, As makes the largest contribution (F – ratio = 31.446, p = 0.0000090) in explaining the difference in water quality between ex-mining ponds and lakes, followed by Cd with (F – ratio = 18.714, p = 0.0000306). Though Pb has an F – ratio slightly greater than 1, no significant difference is observed (F – ratio = 1.832, p = 0.1885145). This is attributed to the low Pb concentrations in the studied sites. This finding support the result of FA suggesting that metal variations are related to the past mining operation, reaffirming that mining activities resulted in the release of toxic heavy metals, thereby significantly affecting the useful nature of the surface water quality (Wolkersdorfer, 2008).

Table 4: Analysis of discriminating variables in ex-mining ponds and lakes in Selangor

Variable	F – Ratio	p - Value
As	31.446	0.0000090
Cd	18.714	0.0000306
Pb	1.832	0.1885145
Mn	8.122	0.0088393
Fe	41.481	0.0000012
DO	0.386	0.5403236
BOD	0.476	0.4969912
pH	55.601	0.0000001

Parameter Relationship

The correlation matrix is presented in Table 5 which shows relationship between parameters under study. A significant positive relationship is observed between As and Cd (r = 0.8689) suggesting a common origin of the metals which are much related to mining activity (Oke and Vermeulen, 2017), and it reflect the influence of heavy metals on the surface water quality. A negative correlation is observed between As and DO (r = -0.5936), high As concentration was reported to lower the DO of the surface water (Buchireddy *et al.*, 2009). The relationship between Pb and Mn (r = 0.7528), DO and Mn (r = 0.5989) are significant. There is strong negative correlation between Fe and pH (r = -0.9429). Low pH liberates metals from their ores into solution especially Fe with very high percent natural abundance (Johnson *et al.*, 2000).

Table 5: Correlation coefficient of parameters under study

	As	Cd	Pb	Mn	Fe	DO	BOD	pH
As	**1.0000**							
Cd	**0.8689**	**1.0000**						
Pb	0.0202	0.1727	**1.0000**					
Mn	-0.4823	-0.4000	**0.7528**	**1.0000**				
Fe	-0.4070	-0.3718	-0.1166	0.2243	**1.0000**			
DO	**-0.5936**	-0.1450	0.4232	**0.5989**	0.0355	**1.0000**		
BOD	0.4997	-0.2531	-0.3838	0.0465	0.2783	0.3796	**1.0000**	
pH	0.4734	0.3404	-0.1169	-0.4292	**-0.9429**	-0.1006	-0.1593	**1.0000**

Bold correlations are significant

CONCLUSION

The results of factor and discriminant analyses revealed a notable influence of mining on the surface water quality. The two-factor models obtained show a significant variation in toxic heavy metals and physicochemical parameters at the studied sites. The water quality variations observed are much related to the anthropogenic input. Heavy metals predominantly As and Cd, and to a lesser extent Pb are associated with ex-mining ponds. Furthermore, the correlation analysis revealed a strong association between As and Cd which proved that metal content in the surface water is significantly influenced by the geochemical properties originating from mining, and are of concern due to their toxicity to humans. Beside the natural sources of Fe and Mn, it is clear that mining activity introduces these metals to the surface water.

The influence of pH originated from mining activity, and residential discharge results in changes in oxygen levels. The chemometric techniques applied in this study gives more specific and objective interpretation of the surface water quality by identifying the source of the pollutants.

Author's contributions: Isa Baba Koki carried out sampling and laboratory analysis; Abubakar Lawal performed multivariate analysis; and Syed Noeman Taqui carried out statistical data analysis.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

Acheampong, M.A., Paksirajan, K., and Lens, P.N. (2013). Assessment of the effluent quality from a gold mining industry in Ghana, Environmental Science and Pollution Research, 20(6): 3799–3811.

Althuwaynee, O. F., Pradhan, B., and Lee, S. (2012). Application of an evidential belief function model in landslide susceptibility mapping. Computers and Geosciences, 44: 120-135.

Bao, K. S.., Shen, J., Wang, G. P., and Tserenpil, S. (2016). Anthropogenic, detritic and atmospheric soil-derived sources of lead in an alpine poor fen in northeast China. Journal of Mountain Science, 13(2): 255-264.

Buchireddy, P. R., Bricka, R. M., and Gent, D. B. (2009). Electrokinetic remediation of wood preservative contaminated soil containing copper, chromium, and arsenic. Journal of Hazardous Materials, 162(1): 490-497.

Crowther, J., Kay, D. and Wyer, M.D. (2001). Relationships between microbial water quality and environmental conditions in coastal recreational waters: The Flyde Coast, UK. Water Research, 35(17): 4029-4038.

Gleick, P. H. (2003). Water use. Annual Review of Environment and Resources, 28(1): 275-314.

Johnson, R.H., Blowes, D.W., Robertson, W.D., and Jambor, J.L. (2000). The hydrogeochemistry of the Nickel Rim mine tailings impoundment, Sudbury, Ontario, Journal of Contaminant Hydrolology, 4: 49–80.

Juahir, H., Zain, S. M., Aris, A. Z., Yusoff, M. K., and Mokhtar, M. B. (2010). Spatial assessment of Langat river water quality using chemometrics. Journal of Environmental Monitoring, 12(1): 287-295.

Koki, I. B., Low, K. H., Juahir, H., Azid, A., and Zain, S. M. (2017). Assessment of water quality of man-made lakes in Klang Valley (Malaysia) using Chemometrics: The impact of mining. Desalination and Water Treatment, 74: 125-136.

Koklu, R., Sengorur, B., and Topal, B. (2010). Water quality assessment using multivariate statistical methods: a case study of Melen River system (Turkey). Water Resources Management 24(5): 959–978

Liu, C.W., Lin, K.H. and Ku, Y.M. (2003). Application of factor analysis in the assessment of groundwater quality in a black foot disease area in Taiwan. Science of the Total Environment, 313: 77–89.

Low, K. H., Isa Baba, K., Hafizan, J., Azman, A., Shima, B., Rabia, I., Hamisu, A. M. and Sharifuddin, M. Z. (2016). Evaluation of water quality variation in lakes, rivers, and ex-mining ponds in Malaysia (review). Desalination and Water Treatment,57(58): 28215- 28239.

Madzin, Z., Shai-in, M. F., and Kusin, F. M. (2015). Comparing heavy metal mobility in active and abandoned mining sites at Bestari Jaya, Selangor. Procedia Environmental Sciences, 30: 232-237.

MOH (2004). Ministry of Health Malaysia, National Standard for Drinking Water Quality, Document D1, Malaysia.

Morgan, R. (1968). An Analysis of Basin Asymmetry in the Klang Basin, Selangor. Bulletin of the Geological Society of Malaysia, (3-5): 17.

Mustapha, A., Aris, A. Z., Juahir, H., Ramli, M. F., and Kura, N. U. (2013). River water quality assessment using environmentric techniques: case study of Jakara River Basin. Environmental Science and Pollution Research, 20(8): 5630-5644.

Muthaiah, V. P. K., Chen, G.D., Ding, D., Salvi, R., and Roth, J. A. (2016). Effect of manganese and manganese plus noise on auditory function and cochlear structures. Neurotoxicology, 55: 65-73.

Navarro, M.C., Pérez-Sirvent, C., Martínez-Sánchez, M.J., Vidal, J., Tovar, P.J., Bech, J., (2008). Abandoned mine sites as a source of contamination by heavy metals: a case study in a semi-arid zone, Journal of Geochemical Exploration, 96: 183–193.

Ning, L., Liyuan, Y., Jirui, D., & Xugui, P. (2011). Heavy metal pollution in surface water of Linglong gold mining area, China. Procedia Environmental Sciences, 10, 914-917.

Oke, S., and Vermeulen, D. (2017). Geochemical Modeling and Remediation of Heavy Metals and Trace Elements from Artisanal Mines Discharge. Soil and Sediment Contamination: An International Journal, 26(1): 84-95.

Onichandran, S., Kumar, T., Lim, Y.A., Sawangjaroen, N., Andiappan, H., Salibay, C.C., Chye, T.T., Ithoi, I., Dungca, J.Z., Sulaiman, W.Y.W., Ling, L.Y. and Nissapatorn, V. (2013). Waterborne parasites and physicochemical assessment of selected lakes in Malaysia, Parasitology Research, 112(12): 4185–4191.

Rogerson P.A. (2006). Statistical methods for geography: a student's guide. Sage, London

Rojas, J.C. and Vandecasteele, C. (2007). Influence of Mining Activities in the North of Potosi, Bolivia on the Water Quality of the Chayanta River, and its Consequences. Environmental Monitoring and Assessment, 132(1): 321-330.

Sadeghi, S., Harchegani, M. K., and Younesi, H. (2012). Suspended sediment concentration and particle size distribution, and their relationship with heavy metal content. Journal of Earth System Science, 121(1): 63-71.

WHO (2011). Guidelines for drinking-water quality: Geneva: World Health Organization.

Wolkersdorfer, C. (2008). Water management at abandoned flooded underground mines: fundamentals, tracer tests, modelling, water treatment. Berlin: Springer Science & Business Media.

Yap, C.K., Ismail, A., and Chiu, P.K. (2007). Water quality and dissolved heavy metal concentrations in surface water collected from Kelana Jaya Lakes. Asian Journal of Water Environment and Pollution, 4: 187–190.