Chemosphere 100 (2014) 8–15

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Biological toxicity of groundwater in a seashore area: Causal analysis and its spatial pollutant pattern

Winn-Jung Huang [a], Wei-Yea Chen [b], Yen-Hsun Chuang [b], Yu-Hao Lin [b], Ho-Wen Chen [b,*]

[a] Department of Safety, Health and Environmental Engineering, Hungkuang University, No. 1018, Sec. 6, Taiwan Boulevard, Shalu District, Taichung City 43302, Taiwan, ROC
[b] Department of Environmental Science and Engineering, Tunghai University, No. 1727, Sec. 4, Taiwan Boulevard, Xitun District, Taichung City 40704, Taiwan, ROC

HIGHLIGHTS

- The safety of groundwater usage in seashore area was concerned in Taiwan.
- We proposed a data mining approach on discovering knowledge from groundwater monitoring data.
- Integrating HCA with PCA method could find out the reasons resulting in biological toxicity.
- GIS combined with the Kriging method visualize the spatial patterns of groundwater quality.
- As^{3+} is possibly main contributor to biological toxicity compared to disinfection by-products.

ARTICLE INFO

Article history:
Received 19 June 2013
Received in revised form 9 December 2013
Accepted 31 December 2013
Available online 23 January 2014

Keywords:
Data mining
Spatial statistics
Groundwater safety
Biological toxicity

ABSTRACT

To ensure the safety of groundwater usage in a seashore area where seawater incursion and unexpected leakage are taking place, this paper utilizes the Microtox test to quantify the biological toxicity of groundwater and proposes an integrated data analysis procedure based on hierarchical cluster analysis (HCA) and principal component analysis (PCA) for determining the key environmental factors that may result in the biological toxicity, together with the spatial risk pattern associated with groundwater usage. For these reasons, this study selects the coastal area of Taichung city in Central Taiwan as an example and implements a monitoring program with 40 samples. The results indicate that the concentration of total arsenic in the coastal areas is about 0.23–270.4 $\mu g\,L^{-1}$, which is obviously higher than the interior of Taichung city. Moreover, the seawater incursion and organic pollution in the study area may be the key factors resulting in the incubation of toxic substances. The results also indicate that As^{3+} is the main contributor to biological toxicity compared to other disinfection by-products. With the help of the visualized spatial pollutants pattern of groundwater, an advanced water quality control plan can be made.

1. Introduction

The availability of adequate fresh water is a fundamental requirement for the environmental sustainability of a human community of any size. The dramatic changes of societal complexity due to intensive interactions among stakeholders in agricultural, industrial, and municipal sectors compound water resource allocation and redistribution in many watersheds that are increasingly being managed for multiple objectives. Accelerating demand on water resources for municipal, agricultural and industrial use caused by rapid economic development and improvements in living standards has placed a serious stress on the national water supply.

Like most countries in the word, Taiwan is considered as an area with insufficient water resources due to inherent environmental limitations such as steep terrain and extremely uneven rainfall. According to long-term observations, about 80% of rainfall occurs between April and September. Such extremely uneven stream flows between dry and wet seasons in Taiwan result in the development of reservoirs and weirs for conservation of water resources in almost all river basins. At present, there are about 53 reservoirs and 60 large weirs existing in Taiwan. These existing hydraulic facilities not only play an important role in water allocation but also obstruct the usage of stream flow in the downstream area during the dry season. To satisfy the need of water resources for municipal, agricultural and industrial use, groundwater has become the most important water resource, especially in the coastal area of West Taiwan. Given this fact, the increased usage of groundwater has caused excessive drawdown and land subsidence in many

* Corresponding author. Tel.: +886 4 23590121x33636; fax: +886 4 23595941.
E-mail address: hwchen@thu.edu.tw (H.-W. Chen).

0045-6535/$ - see front matter
http://dx.doi.org/10.1016/j.chemosphere.2013.12.073

coastal areas (Burnett et al., 2001; Stiros, 2001; Carbognin et al., 2004; Phien-wej et al., 2006).

The mixing of groundwater and saline water further complicates the hydro geochemical processes in the coastal aquifer (Moore, 1999; Helena et al., 2000; Kistemann et al., 2008). Such hydrologic conditions increased the anxiety of environmentalists about degraded water quality and groundwater conservation, and also resulted in difficulty in determining the environmental situation by gathering data from monitoring programs without any support by sophisticated data analysis techniques. Being capable of providing meaningful insight from large and chaotic data sets with a large number of parameters, multivariate statistical techniques are broadly utilized to evaluate the spatial patterns of pollutant concentration in groundwater (van den Brink et al., 2007; Belkhiri et al., 2010; Yidana et al., 2010), to assess the quality of groundwater resources by integrating the water quality index (Omo-Irabor et al., 2008; Ahmed Baig et al., 2010), and to resolve hydrological factors such as aquifer boundaries, groundwater flow paths and hydro-chemical parameters (Garcia Pereira et al., 2003; Halim et al., 2010). These applications are valuable for indicating temporal and spatial variations of groundwater quality caused by nature and anthropogenic factors for advanced governing controls.

Principal component analysis (PCA) is one of the common multivariate statistical techniques that are used to achieve great efficiency of data compression from the original data as well as to indicate natural associations between samples and/or variables (Wenning and Erickson, 1994; Astel et al., 2007) by gaining some information useful in the interpretation of the environmental system. PCA consists of diagonalization of the covariance or correlation matrix transforming the original measurements into linear combinations of these measurements, and then the explained variance of each principal component can be maximized. It has been widely used to reveal the relationships among variables as well as to classify them into different latent variables, so that some special features inherent in the measured system can be characterized (Lautre and Fernandez, 2004; Macciotta et al., 2006; Chen et al., 2007; Lucas et al., 2008). For example, the integrated PCA and HCA technologies are used to quantify the health risk of pollutants in drinking water usage as well as to identify the pollution sources for advanced control (Shah and Shaheen, 2008). Recently, many comprehensive approaches consisting of PCA and other data analysis technologies such as geographic information systems (GIS) and hierarchical cluster analysis (HCA) are also proposed to solve environmental problems with spatial characteristics (Zhou et al., 2007; Lima et al., 2010).

Today, groundwater in this area is almost no longer used for drinking or cooking, but is still being used for aquaculture or non-drinking water such as bathing after simple disinfection. Thus, in order to prevent exposing the public to toxic substances, government officials deem it imperative to determine the potential risk map for advanced groundwater quality control (Sekhar et al., 2003). For these reasons, this study presents a monitoring program in the seashore area of Taichung city, together with a toxic test, to measure the toxic potentiality under different physical and chemical environments. The simplest method of risk mapping is simply focusing on the spatial position and environmental characteristics of sites (Peterson, 2006). The spatial distribution of environmental parameters is displayed in the "risk map" and in the preliminary evaluation of the risk of water use. Meanwhile, integrated multivariate analysis consisting of PCA and HCA is used to explain the spatial correlation of groundwater samples among monitoring stations with a large number of environmental variables. To present the spatial patterns of pollutants, the Kriging method (Todini and Ferraresi, 1996; Lark, 2000) was utilized to make the risk map. With this understanding of the spatial variance in groundwater quality, potential effects of changes in environmental pressures can be assessed, and necessary abatement actions to sustain usable water supplies can be identified.

2. Methodology

2.1. Background introduction and sampling design

This paper selects the seashore area of Taichung city, located in central Taiwan, as a case study. The east of the site is connected to the Greater Taichung Metropolitan area with a population of 1.5 million. To the west of the study area, next to the coast line of the Taiwan Strait, is the Taichung coal-fired power plant and Taichung harbor. Within the study area, there is a conventional industry complex named the Guan-Lian industry complex, which has been heavily operated for about 30 years. The study area is facing the threats of saltwater intrusion and unexpected leakages from industrial waste. To evaluate the spatial variance of groundwater quality as well as profiling a risk map, a total of thirty-nine sampling sites, as shown in Fig. 1, were collected all over the area of concern. They are water-wells with depths varying from 20 to 80 m using standard sampling procedures (Eaton et al., 2005) during the dry season (October–May) in 2011; there are ten, sixteen, and thirteen sampling sites located in the Shalu district, Wuchi district, and Longjing district, respectively. In the study area, the high groundwater table and quantity in the wet season may dilute the contributor concentrations of toxic substances, and the seawater intrusion with higher Cl^- concentration is more serious as well, so the sampling data in the dry season was investigated by our integrated multivariate analysis method. From the concentration of chlorite, it is believed that seawater incursion is happening in this area.

2.2. Laboratory analysis and sample characterization

To characterize the groundwater quality in reference to the need for health risk management (Kuo et al., 1998), some characteristic variables of groundwater including non-purgeable dissolved organic carbon (NPDOC), Cl^-, pH, DO, As^{3+}, As^{5+}, turbidity, conductivity, humic substances (H.S.), trihalomethanes (THMs) and haloacetic acids (HAAs) were selected and measured in this study, together with a Microtox test. The measurement of dissolved oxygen (DO) is one of common indicator to evaluate water quality, which is a measure of how much oxygen is dissolved in the water and measured in milligrams per liter (mg L^{-1}). All samples were analyzed for the main chemical descriptors by referring to standard methods (Eaton et al., 2005). DO, temperature, pH and electrical conductivity (EC) were measured in cell sensors, and that sensors were calibrated with standard solutions (all from Merck). The determination of inorganic arsenic species, $As^{3+}$and As^{5+}, was done using the high-performance liquid chromatography (HPLC) with inductively coupled plasma mass spectrometry (ICP-MS) analysis (Perkin–Elmer Sciex-6100 Elan DRC II ICP/MS). The water sample was loaded into the sample loop and then injected on a separation column, and the arsenic species were eluted from the separation column. For organic parameters analysis, water samples were first filtered through 0.45 μm membrane filter. NPDOC was measured by the combustion–infrared method using a total organic carbon analyzer (Model TOC-5000, Shimadzu, Tokyo, Japan). Species of THMs were measured by purge-and-trap injection system (Purge & Trap 4560) coupled with gas chromatography–mass spectrometry (GC/MS) (Agilent 6890N/5973). Species of HAAs measurements involved liquid/liquid extraction with MTBE (Methyl t-butyl ether) and esterification step with diazomethane before gas chromatography/electron capture detection (GC/ECD) (Varian 3800) analysis.

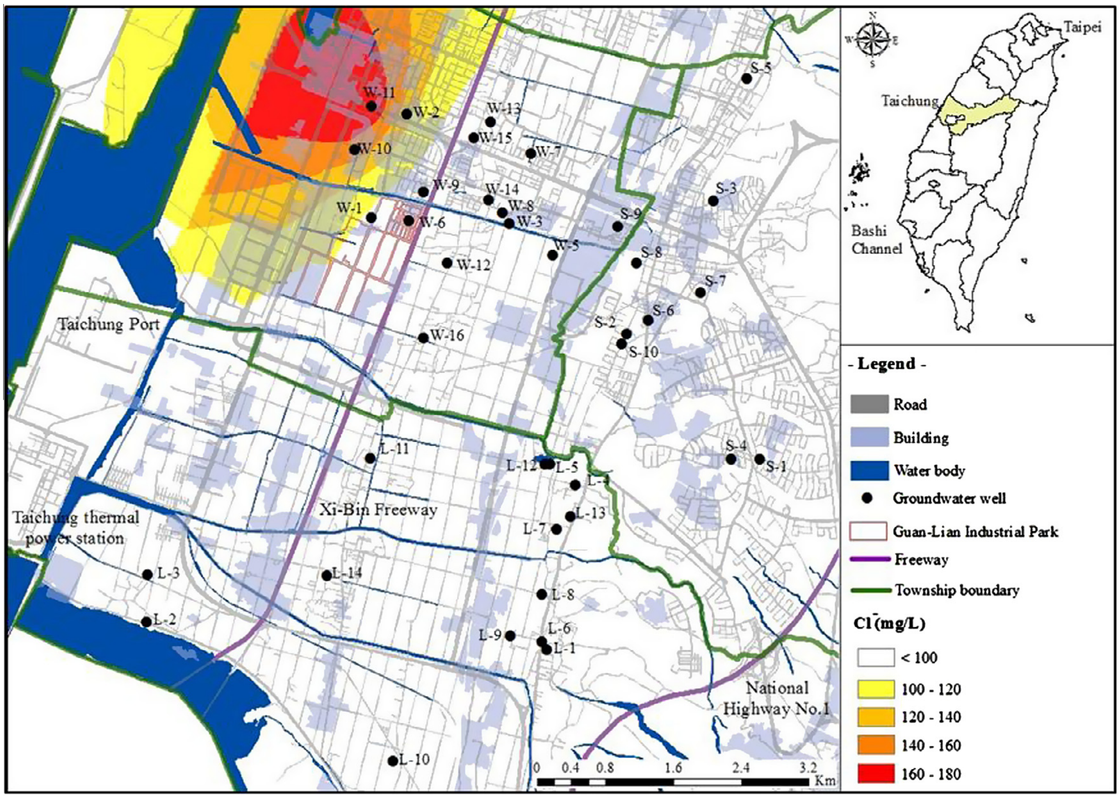


Fig. 1. Systematic environment and spatial distribution of monitoring stations in study area.

Besides the above potential toxic contribution, the biologic toxicity is one of several important indexes to characterize the groundwater quality. However, the biological toxicity is a comprehensive index and its potential contribution could be attributed to the complicated groundwater characteristic variables such as inorganic metal, organic contaminants and their by-products (Marjorie Aelion and Davis, 2007). Because the Microtox acute toxicity sensor can quantify the biological toxicity of a sample by detecting the changes in natural bioluminescence caused by Vibrio fischeri, it is applied to this study to distinguish the biological toxic levels at different sampling sites. The Microtox test utilized in this study involves an SDI Model 500 analyzer and lyophilized cultures of Vibrio fischer NRRL-B-11177 (Bioresource Collection and Research Center, Hsinchu, Taiwan), in which the inhibition of bacterial light emissions was evaluated in duplicate experiments at 15 °C, following exposure for 15, 30 and 60 min. Solutions of the extract at five concentrations in a geometric sequence that had been exchanged with Milli-Q water were analyzed, starting with a 100x concentration factor by diluting 0.5 mL extract with an equal volume of the bacterial suspension. The results of the Microtox test were presented as EC50, which is the effective concentration required to reduce the light output by the organism by 50%. The EC50 values and the corresponding confidence intervals were calculated (95%) using the Log-normal model in the REGTOX software application for Microsoft Excel. In the experiment, three EC50 values were calculated after 15, 30 and 60 min and were denoted as the 15-min EC_{50}, 30-min EC_{50} and 60-min EC_{50}, respectively.

2.3. Causal analysis of biological toxicity and its spatial pollutant pattern

More and more spatial-temporal data for pollutants are becoming available due to increasing monitoring programs. If some useful information inherent in these data can be explored by data mining technologies, it will help decision makers to outline the environmental situations and to plan advantageous strategies for groundwater conservation. In this respect, the PCA approach is utilized in this study to simplify high dimensional variables while retaining most of the primary information, as well as to integrate some individual variables into comprehensive factors which stand for a kind of conceptual environmental characteristic. Followed by HCA, the monitored data with specified similarity can be simplified into several groups for better insight.

The basis of PCA has been well-explained by (Jolliffe, 2002). Briefly, PCA is used for characterizing patterns within large sets of data by re-expressing to a rotated coordinate system in which as much variance as possible is explained by the first few dimensions, in which the eigenvectors of the variance covariance matrix are calculated, so that the principal component score, i.e. the

Table 1
Chemical characteristic of groundwater samples.

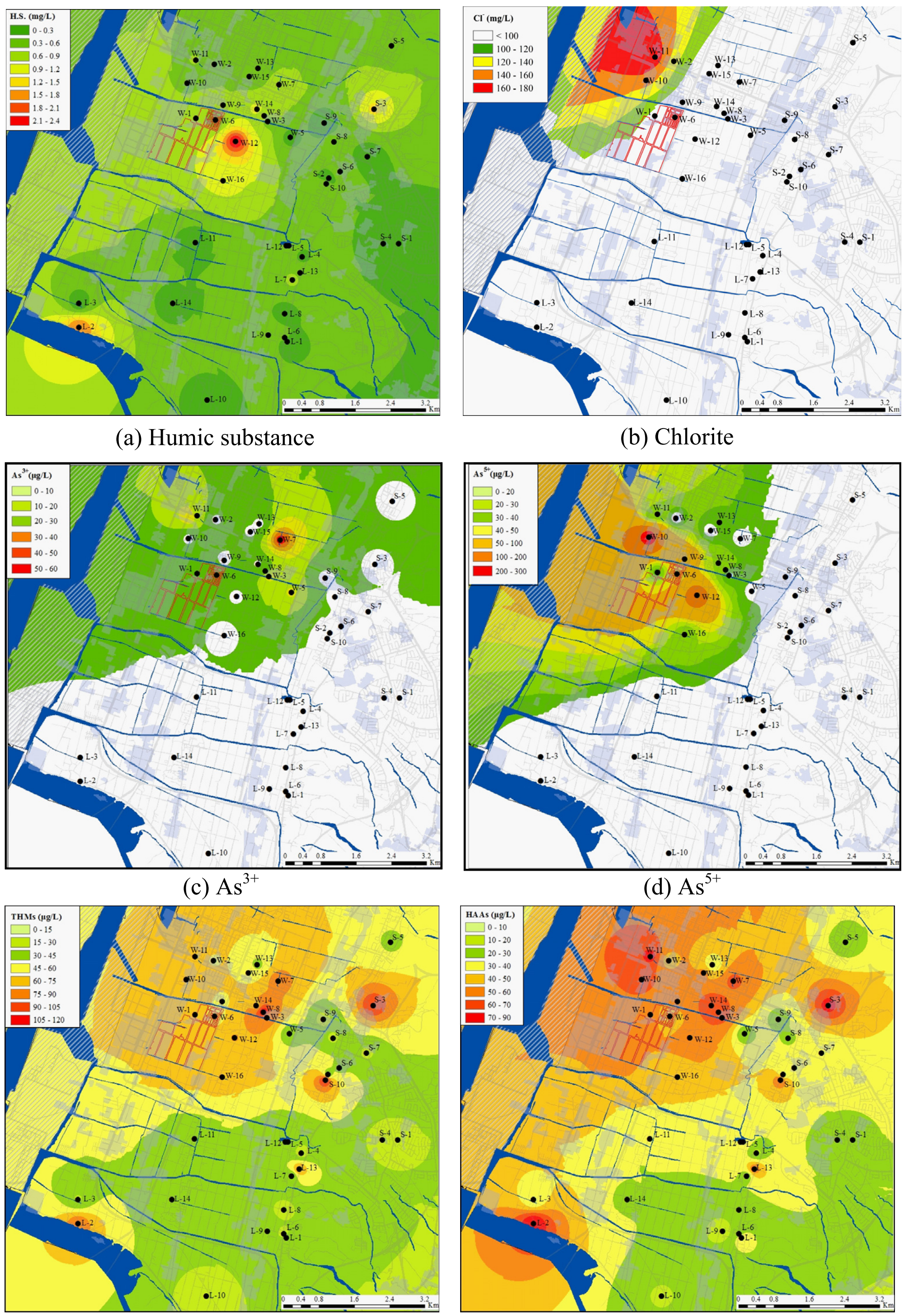
Variable (unit)	Min	Max	Average	Standard deviation
NPDOC (mg L^{-1})	0.26	9.03	1.61	1.61
Cl^- (mg L^{-1})	8.20	500.00	51.71	95.52
DO (mg L^{-1})	1.85	6.73	4.54	1.17
Turbidity (NTU)	0.28	10.00	1.44	2.14
Conductivity (μS cm^{-1})	328.0	8607.0	839.1	1439.8
H.S. (mg L^{-1})	0.09	2.40	0.55	0.49
pH	5.72	8.60	7.08	0.76
As^{3+} (μg L^{-1})	0.00	56.00	4.10	11.11
As^{5+}(μg L^{-1})	0.13	270.40	19.07	49.83
THMs (μg L^{-1})	12.94	3858.98	139.83	595.90
HAAs (μg L^{-1})	21.71	2131.21	102.66	325.91
Biological toxicity (EC50)	0.00	4.57	0.53	1.00

Note: THMs = trihalomethanes; H.S. = humic substances.

weight of the eigenvector can be obtained. The scores of the original variables, also called principal component loadings (PC loadings), can be used to indicate the relationship between the variable and the principal component. In other words, the variable and the principal component will be more strongly related if the PC loading is larger. By doing this, the raw data matrix can be reduced to two or three principal component loadings that account for the majority of the variance. Thus, these factors can be used to account approximately for the required information, just as the original observations do.

HCA is utilized to explore the spatial relationships among the objects by examining their distances, and then a graphic display of how these objects are clustered can be obtained. HCA measures the similarity between every pair of objects with a standardized m-space Euclidian distance that can be shown as Eq. (1) (Davis, 1986).

$$d_{ij} = \sqrt{\frac{\sum_{k=1}^{m}(X_{ik} - X_{jk})^2}{m}} \tag{1}$$



(a) Humic substance (b) Chlorite

(c) As^{3+} (d) As^{5+}

(e) THMs (f) HAAs

Fig. 2. The spatial distribution of measured groundwater parameters.

Table 2
Correlation matrix of water chemistry components.

Variable	Environment parameters							Toxic parameters			
	NPDOC	Cl^-	DO	Turbidity	Conductivity	H.S.	pH	As^{3+}	As^{5+}	THMs	HAAs
NPDOC	1										
Cl^-	.430**	1									
DO	.327*	.261	1								
Turbidity	.151	.585**	−.224	1							
Conductivity	.374*	.948**	.218	.690**	1						
H.S.	.657**	.244	.337*	.022	.235	1					
pH	.405**	.302	.341	.168	.306	.414**	1				
As^{3+}	.320*	.270	.004	.228	.308	.160	.402*	1			
As^{5+}	.172	.310	.016	−.042	.112	.379*	.330*	−.050	1		
THMs	.756**	.512**	.307	.230	.464**	.349*	.337*	.416**	.031	1	
HAAs	.787**	.515**	.314*	.227	.465**	.374*	.353*	.426**	.040	.999**	1

H.S. = humic substances; THMs = trihalomethanes; HAAs = haloacetic acids.
** Correlation significant at $p < 0.001$.
* Correlation significant at $p < 0.05$.

Table 3
Factor loadings.

Variable	Factor	
	F_1 (salination factor)	F_2 (pollution factor)
NPDOC	.261	**.769**
Cl−	**.875**	.317
pH	.216	.648
DO	−.071	**.711**
Turbidity	**.884**	−.145
Conductivity	**.925**	.261
H.S.	.062	**.810**
% Of variance	45.44	24.41
Cumulative %	45.44	69.86

The bold values are the main components in each factor.

in which X_{ik} denotes the kth variable measured on object i and X_{jk} is the kth variable measured on object j. A low distance shows the two objects are similar or "close together", whereas a large distance indicates dissimilarity.

Because the groundwater environmental parameters are not independent and there are a large number of correlated environmental parameters, it is difficult to investigate the dominant factors of toxicity. In this study, the technique of PCA is primarily used to reduce the number of environmental parameters dimensions and obtain the independent principal components to maximally account for the variance of the data. These principal components are then input to the HCA to distinguish the correlation between environmental parameters and toxic substances. Thus, this paper uses the software package SPSS for Windows to

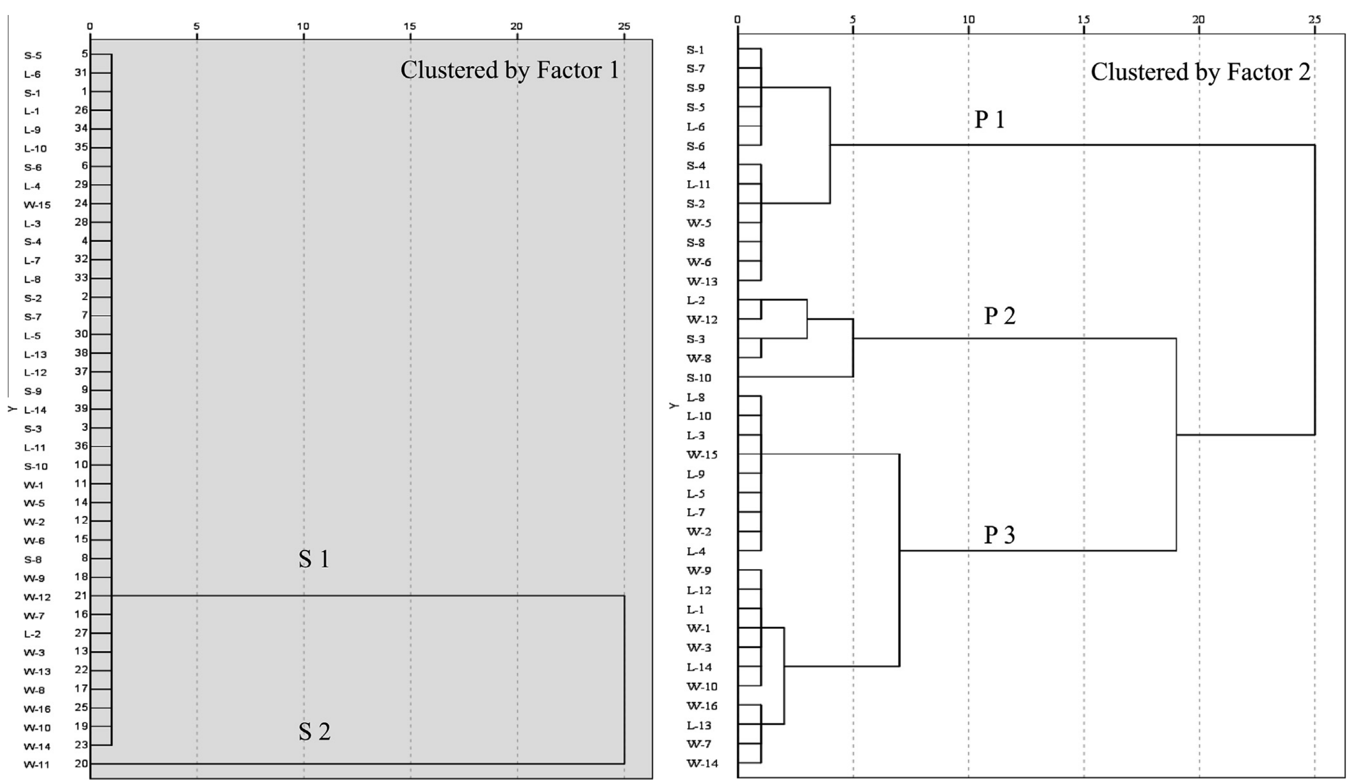


Fig. 3. The results of clustering by using different factor components.

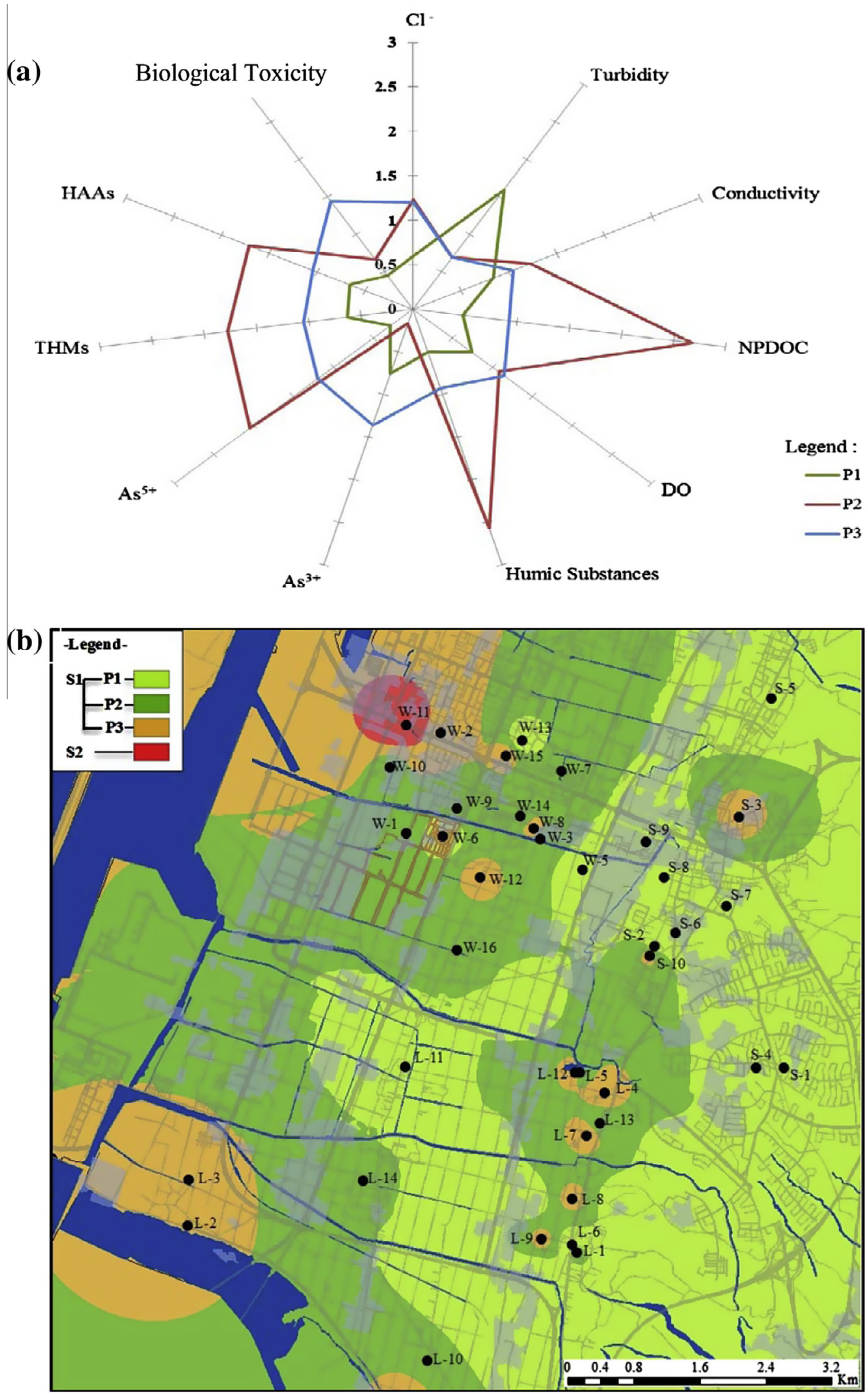


Fig. 4. Distributions of geochemical conditions and toxic pollutants among different clusters. (a) Organic pollution factors; (b) spatial patterns of clustered groups.

determine the structure in the relationships between groundwater quality parameters and identify the most important factor(s) contributing to this structure based on the eigen analysis of the correlation matrix. Following the PCA process for grouping the characteristic parameters of objects, HCA is utilized to explore the spatial relationships among the objects by examining their distances. Finally, GIS integrated with the Kriging method is utilized to visualize the spatial patterns of groundwater quality based on the inverse distance weighting method.

3. Results and discussion

3.1. Chemical characteristics of observations

Chemical data of 39 analyzed groundwater samples are summarized in Table 1. Groundwater pH was predominantly mildly alkaline (pH > 7). Electrical conductivity varied widely from 328 to 8607 $\mu S\ cm^{-1}$ and the highest value occurred at a shallow aquifer next to the Taiwan Strait, especially at sampling sites W-11 and

W-10 (shown in Fig. 2(b)). Such high electrical conductivity and high concentrations of Cl^- have indicated that seawater has intruded into the coastal areas of Taichung City and resulted in complex reactions among chemicals.

According to the results shown in Fig. 2(a), high concentrations of humic substance, which are considered to be the precursors of disinfection by-products, are found in the area surrounding the Guan-Lian industry complex and some sampling sites near the polluted river. It reveals that potential organic pollution may happen in these areas. The concentrations and spatial distributions of As^{3+} and As^{5+} are presented in Table 1 and Fig. 2. The results reveal that sampling sites which are next to the coastal area and industry complex contain higher As concentration (>10 g L^{-1}) (Smedley and Kinniburgh, 2002). Seawater intrusion, geochemical conditions, and pollutant leakage may be the reasons leading to the difference in spatial distribution of As^{3+} and As^{5+}.

Looking at Table 2, we see that conductivity, turbidity, and concentration of Cl^- ions are highly correlated; their correlation coefficients range from 0.585 to 0.690. It reveals that seawater intrusion may increase the turbidity in groundwater. The formation of organic toxic pollutants including THMs and HAAs is not highly correlated to humic substances but to the concentration of NPDOC. On the other hand, As^{3+} has a higher correlation than As^{5+} to THMs and HAAs. This seems to reveal that a high redox potential may stimulate the formation of organic by-products.

3.2. Environmental characteristic analysis

This study uses the environmental parameters mentioned above as the analyzed variables of PCA to define the dominant principal components (i.e. the dominant dimension of the factorial space of data representation). As shown in Table 3, a two-factor model was determined by PCA, in which their percentage of variance and cumulative percentage of variance explained 69.86% of the total variance in the data set. The first factor (F_1), consisting of contained Cl^-, turbidity, and conductivity, accounted for 45.44% of the total variance. Factor 1 is strongly determined by the Cl^- concentration and conductivity, which are significantly correlated to the intrusion of salt water. Then this study named the F_1 factor with a new comprehensive title, "seawater intrusion factor," to make a preliminary identification of the impact area of seawater intrusion in this area. This F_1 would be the input of HCA to classify the site, then distinguish the correlation between environmental conditions and toxic substances.

The second factor (F_2) contains the variables NPDOC, DO, and humic substance, and explained 24.41% of the total variance. The variables contained in F_2 reflect the possible leakage of organic pollutants. The results of correlation analysis shown in Table 2 also reveal the slight positive dependence of the DO on humic substance concentrations in the water samples, as well as showing that sampling sites with high humic substance and DO concentration are located around the Guan-Lian industry complex (shown in Fig. 2(a)), thus this result suggests that the main contribution of H.S. is human activity. Confronted with this situation, factor 2 was denoted as the pollution factor. Then, these two factors are used to stand for the distinctive environmental conditions and to explain the spatial pattern of groundwater pollutants in the study area.

Following the results of PCA, groundwater samples were classified by HCA into some clusters. As shown in Fig. 3(a), the groundwater samples were classified by HCA into two main clusters (S1 and S2) according to their dominant chemical composition (Factor 1 in Table 3). In the cluster S2, only the area near the well W11 can be considered as a seawater injected area with high salinity and conductivity. Though some sampling sites in the cluster S1 have high Cl^- concentration (shown in Fig. 2(b)), they still statistically cannot be assumed to be seawater injected areas.

Excluding the cluster S2, 38 groundwater samples are further segmented into three subgroups (P1, P2, and P3) based on the pollution factors shown in Table 3. The results are shown in Fig. 3(b). To further explore the influence of geochemical conditions upon the formation of toxic pollutants, a radar chart is plotted as Fig. 4(a) in company with a normalization procedure. Surprisingly, it is noted that samples in cluster P3 have lower quantitative values for measured parameters than clusters P1 and P2 except turbidity. Low humic substance and NPDOC concentrations seem to reflect that inorganic matters are the main components of turbidity in cluster P3. Cluster P1 is similar to cluster P2 with regard to Cl^-, DO, and pH parameters. The notable difference between these two clusters for environmental parameters is that there is higher organic substance in cluster P2. These organic matters may stimulate the formation of HAAs and THMs. However, the heavy disinfection by-products are not the main source of biological toxicity. Fig. 4(a) pinpoints that when As^{3+} increases, increasing the biological toxicity, it implies the As^{3+} is a major contributor to biological toxicity.

Finally, Kriging interpolation and GIS technologies are utilized to visualize the spatial pollution patterns of groundwater. Based on the result of Fig. 4(a), the study area is separated into three subareas with different groundwater quality conditions. As shown in Fig. 4(b), the area next to the ocean has higher biological toxicity dominated by As^{3+}, though some organic pollutants also exist in this area. Obviously, the transport of arsenic in groundwater is affected not only by the geochemistry but also by human activities (Karim, 2000; Chakraborti et al., 2010). The area in Fig. 4(b) colored dark green is where higher organic pollutants such as THMs and HAAs can be found. It is believed that the pollutant leakage may happen in this area. However, the biological toxicity in this area is not higher than the seashore area. Strategies for controlling organic leakage in this area are required. Groundwater quality in the light green area is cleaner than other areas, with less organic substances and arsenics. The turbidity seems to be composed of nonorganic substances. How to maintain sustainable usage is the most important issue in this subarea. With the help of GIS and the Kriging approach, government officials can determine the spatial patterns of groundwater quality as well as plan advanced control programs corresponding to different quality conditions for different subareas.

4. Conclusions

To explore the spatial pattern of groundwater quality in the seashore area, this paper presents an integrated analysis procedure based on PCA and HCA methods. By clustering groundwater water quality parameters and segmenting the monitoring data, inherent data complexity and dissimilarity can be mitigated and key influencing factors can serve as a basis for groundwater quality planning and management for a seashore area like Taichung City. The proposed algorithm is not only valuable in dealing with the problem of spatializing the groundwater quality but also in determining the regional pollution sources for a complicated environmental system like the study area in this study. This study concludes that seawater has intruded into the seashore area. In the presence of Cl^-, the amount of toxic substance is obviously increased. High organic substance and continuing high DO occur in the vicinity of the Guan-Lian industry complex; this seems to reveal that unexpected leakage has taken place in this area. Moreover, As^{3+} is one of the most dangerous contributors to biological toxicity.

Acknowledgements

The authors would like to thank the National Science Council of the Republic of China and Tunghai University Global Research and

Education on Environment and Society for financially supporting this research under Contract No. NSC 98-2221-E-241-003-MY3 and 101GREENS006-2, respectively.

References

Ahmed Baig, J., Gul Kazi, T., Qadir Shah, A., Abbas Kandhro, G., Imran Afridi, H., Balal Arain, M., Khan Jamali, M., Jalbani, N., 2010. Speciation and evaluation of Arsenic in surface water and groundwater samples: a multivariate case study. Ecotoxicol. Environ. Saf. 73, 914–923.

Astel, A., Tsakovski, S., Barbieri, P., Simeonov, V., 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Res. 41, 4566–4578.

Belkhiri, L., Boudoukha, A., Mouni, L., Baouz, T., 2010. Application of multivariate statistical methods and inverse geochemical modeling for characterization of groundwater – a case study: Ain Azel plain (Algeria). Geoderma 159, 390–398.

Burnett, W.C., Taniguchi, M., Oberdorfer, J., 2001. Measurement and significance of the direct discharge of groundwater into the coastal zone. J. Sea Res. 46, 109–116.

Carbognin, L., Teatini, P., Tosi, L., 2004. Eustacy and land subsidence in the Venice Lagoon at the beginning of the new millennium. J. Mar. Syst. 51, 345–353.

Chakraborti, D., Rahman, M.M., Das, B., Murrill, M., Dey, S., Chandra Mukherjee, S., Dhar, R.K., Biswas, B.K., Chowdhury, U.K., Roy, S., Sorif, S., Selim, M., Rahman, M., Quamruzzaman, Q., 2010. Status of groundwater arsenic contamination in Bangladesh: a 14-year study report. Water Res. 44, 5789–5802.

Chen, K., Jiao, J.J., Huang, J., Huang, R., 2007. Multivariate statistical evaluation of trace elements in groundwater in a coastal area in Shenzhen, China. Environ. Pollut. 147, 771–780.

Davis, J.C., 1986. Statistics and Data Analysis in Geology. John Wiley & Sons.

Eaton, A.D., Franson, M.A.H., Associati, A.P.H., Association, A.W.W., Federation, W.E., 2005. Standard Methods for Examination of Water & Wastewater: Centennial Edition. Amer Public Health Assn.

Garcia Pereira, H., Renca, S., Saraiva, J., 2003. A case study on geochemical anomaly identification through principal components analysis supplementary projection. Appl. Geochem. 18, 37–44.

Halim, M.A., Majumder, R.K., Nessa, S.A., Hiroshiro, Y., Sasaki, K., Saha, B.B., Saepuloh, A., Jinno, K., 2010. Evaluation of processes controlling the geochemical constituents in deep groundwater in Bangladesh: spatial variability on arsenic and boron enrichment. J. Hazard. Mater. 180, 50–62.

Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J.M., Fernandez, L., 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res. 34, 807–816.

Jolliffe, I.T., 2002. Principal Component Analysis. Springer.

Karim, M.M., 2000. Arsenic in groundwater and health problems in Bangladesh. Water Res. 34, 304–310.

Kistemann, T., Hundhausen, J., Herbst, S., Claßen, T., Färber, H., 2008. Assessment of a groundwater contamination with vinyl chloride (VC) and precursor volatile organic compounds (VOC) by use of a geographical information system (GIS). Int. J. Hyg. Environ. Health 211, 308–317.

Kuo, H.W., Chiang, T.F., Lo, I.I., Lai, J.S., Chan, C.C., Wang, J.D., 1998. Estimates of cancer risk from chloroform exposure during showering in Taiwan. Sci. Total Environ. 218, 1–7.

Lark, R.M., 2000. Designing sampling grids from imprecise information on soil variability, an approach based on the fuzzy kriging variance. Geoderma 98, 35–59.

Lautre, I.G., Fernandez, E.A., 2004. A methodology for measuring latent variables based on multiple factor analysis. Comput. Stat. Data Anal. 45, 505–517.

Lima, D.C., Santos, A.M.P.d., Araujo, R.G.O., Scarminio, I.S., Bruns, R.E., Ferreira, S.L.C., 2010. Principal component analysis and hierarchical cluster analysis for homogeneity evaluation during the preparation of a wheat flour laboratory reference material for inorganic analysis. Microchem. J. 95, 222–226.

Lucas, R., Bunting, P., Paterson, M., Chisholm, L., 2008. Classification of Australian forest communities using aerial photography, CASI and HyMap data. Remote Sens. Environ. 112, 2088–2103.

Macciotta, N.P.P., Dimauro, C., Catillo, G., Coletta, A., Cappio-Borlino, A., 2006. Factors affecting individual lactation curve shape in Italian river buffaloes. Livestock Sci. 104, 33–37.

Marjorie Aelion, C., Davis, H.T., 2007. Use of a general toxicity test to predict heavy metal concentrations in residential soils. Chemosphere 67, 1043–1049.

Moore, W.S., 1999. The subterranean estuary: a reaction zone of ground water and sea water. Mar. Chem. 65, 111–125.

Omo-Irabor, O.O., Olobaniyi, S.B., Oduyemi, K., Akunna, J., 2008. Surface and groundwater water quality assessment using multivariate analytical methods: a case study of the Western Niger Delta, Nigeria. Phys. Chem. Earth, Parts A/B/C 33, 666–673.

Peterson, A.T., 2006. Ecologic niche modeling and spatial patterns of disease transmission. Emerg. Infect. Dis. 12, 1822.

Phien-wej, N., Giao, P.H., Nutalaya, P., 2006. Land subsidence in Bangkok, Thailand. Eng. Geol. 82, 187–201.

Sekhar, K.C., Chary, N.S., Kamala, C.T., Rao, J.V., Balaram, V., Anjaneyulu, Y., 2003. Risk assessment and pathway study of arsenic in industrially contaminated sites of Hyderabad: a case study. Environ. Int. 29, 601–611.

Shah, M.H., Shaheen, N., 2008. Annual and seasonal variations of trace metals in atmospheric suspended particulate matter in Islamabad, Pakistan. Water Air Soil Pollut. 190, 13–25.

Smedley, P.L., Kinniburgh, D.G., 2002. A review of the source, behaviour and distribution of arsenic in natural waters. Appl. Geochem. 17, 517–568.

Stiros, S.C., 2001. Subsidence of the Thessaloniki (northern Greece) coastal plain, 1960–1999. Eng. Geol. 61, 243–256.

Todini, E., Ferraresi, M., 1996. Influence of parameter estimation uncertainty in Kriging. J. Hydrol. 175, 555–566.

van den Brink, C., Frapporti, G., Griffioen, J., Zaadnoordijk, W.J., 2007. Statistical analysis of anthropogenic versus geochemical-controlled differences in groundwater composition in The Netherlands. J. Hydrol. 336, 470–480.

Wenning, R.J., Erickson, G.A., 1994. Interpretation and analysis of complex environmental data using chemometric methods. Trac-Trends Anal. Chem. 13, 446–457.

Yidana, S.M., Banoeng-Yakubo, B., Akabzaa, T.M., 2010. Analysis of groundwater quality using multivariate and spatial analyses in the Keta basin, Ghana. J. Afr. Earth Sc. 58, 220–234.

Zhou, F., Guo, H.C., Liu, Y., Hao, Z.J., 2007. Identification and spatial patterns of coastal water pollution sources based on GIS and chemometric approach. J. Environ. Sci.-China 19, 805–810.