Spatio-temporal variation analysis of hydrochemical characteristics in the Luanhe River Basin, China

Ying Xie, Xuyong Li, Huiliang Wang and Wenzan Li

ABSTRACT

The analysis of river pollution and assessment of spatial and temporal variation in hydrochemistry are essential to river water pollution control in the context of rapid economic growth and growing pollution threats in China. In this study, we focused on hydrochemical characteristics of the Luanhe River Basin (China) and evaluation of 12 hydrochemical variables obtained from 32 monitoring stations during 2001–2010. In each study year, the streams were monitored in the three hydrological periods (April, August, and October) to observe differences in the impacts of agricultural activity and rainfall pattern. Multivariate statistical methods were applied to the data set, and the river water hydrochemical characteristics were assessed using the water quality identification index (WQIIM). The results showed that parameters had variable contribution to water quality status in different months except for ammonia nitrogen ($\text{NH}_4\text{-N}$) and total nitrogen (TN), which were the most important parameters in contributing to water quality variations for all three periods. Results of WQIIM revealed that 18 sites were classified as 'meeting standard' while the other 14 sites were classified as 'not meeting standard', with most of the seriously polluted sites located in urban area, mainly due to discharge of wastewater from domestic and industrial sources. Sites with low pollution level were located primarily in smaller tributaries, whereas sites of medium and high pollution levels were in the main river channel and the larger tributaries. Our findings provide valuable information and guidance for water pollution control and water resource management in the Luanhe River Basin.

Key words | factor analysis, hydrochemical parameters, Luanhe River Basin (China), principal component analysis, water quality identification index method

Ying Xie

Xuyong Li (corresponding author)

Huiliang Wang

Wenzan Li

State Key Laboratory of Urban and Regional

Ecology,

Research Center for Eco-Environmental Sciences,

Chinese Academy of Sciences,

Beijing 100085,

China

E-mail: xyli@rcees.ac.cn

Ying Xie

Huiliang Wang

Wenzan Li

Graduate School,

Chinese Academy of Sciences,

Beijing 100049,

China

INTRODUCTION

It is imperative to prevent and control river pollution and to have reliable information on the water quality for effective management (Singh *et al.* 2005). Also it is necessary to interpret a large number of water quality data, but in a huge and complex data matrix comprising a large number of physico-chemical parameters, which are often difficult to interpret and draw meaningful conclusions from (Dixon & Chiswell 1996; Singh *et al.* 2005). Today, multivariate statistical methods such as factor analysis (FA) or principal component analysis (PCA) have been widely used to characterize and evaluate surface and freshwater quality, and they are useful for evidencing temporal and spatial variations caused by natural and anthropogenic factors (Singh *et al.* 2004; Muñoz-Carpena *et al.* 2005).

Studies investigating the spatial variability of water quality have reported that water quality issues are highly

dependent on land use patterns and influence from watershed runoff discharge (Caccia & Boyer 2005; Bu *et al.* 2010). Also spatial variations of the pollutants mainly include the concentration changes between different rivers and between the different sites of the same river. So an analysis should focus on seasonal and spatial variations.

One critical step to control river pollution effectively is to analyze the characterization of the water quality (Kolovos *et al.* 2002; Huang *et al.* 2009). To understand the pollution situation and characteristics accurately, various mathematical techniques based on hydrochemical analysis have been developed. For example, the water quality identification index method (WQIIM) can express comprehensive water quality information. This method determines the water quality status (class) in line with the Chinese national water quality standards and it could be used to

obtain a continual description of the most polluted waters (Liu et al. 2010).

In recent years, with rapid economic growth, the river water quality has been seriously degraded and most sectors of northern China do not meet the Chinese national water quality standards. The water pollution of the Luanhe River Basin (LRB) represents the common situation of most rivers in the north of China, including various types and spatial patterns of pollutants, various processes and multiple sources. In order to develop and improve pollution control strategies, assessing the kind and extension of water pollution accurately and identifying important water quality variables are needed (Wang et al. 2011). The objectives of this study were to: (1) extract the most important parameters in assessing seasonal variations of the LRB; (2) assess the hydrochemistry of surface water quality based on WQIIM, revealing the pollution levels of the rivers.

METHODS AND MATERIALS

Study area

The LRB is located 115°30' to 119°45' E, and 39°10' to 42°40' N. The Luan River is 888 km long with a drainage area of 44,070 km², of which 98% is mountainous. The watershed receives an average annual precipitation of 560 mm, mostly

in summer (70–80%), especially in July and August, and an average annual runoff of $46.94 \times 10^8 \text{ m}^3$ (Li & Feng 2007). The LRB is divided into upstream and downstream parts by the Panjiakou and Daheiting reservoirs. There is almost no water in rivers in the downstream region of the LRB, so this paper focused on the upstream region.

The water quality monitoring center of the Chengde Branch of the Hebei Provincial Survey Bureau of Hydrology and Water Resources (SBHWR) monitored hydrochemical parameters over 11 years (2000–2010) at 32 monitoring sites. These sites were distributed within the study area and covered the main river channel (Luan River) and eight associated tributaries (Figure 1). The tributaries are Yixun River, Liu River, Xingzhou River, Laoniu River, Sa River, Bao River, Wulie River and Panjiakou-Daheiting Reservoir.

Sampling and monitoring data

From 2000–2010, we sampled once in each site in early April, August and October. The hydrochemical parameters included dissolved oxygen (DO), ammonia nitrogen (NH₄-N), nitrate nitrogen (NO₃-N), chemical oxygen demand (COD), total phosphorus (TP), total nitrogen (TN), fluoride (F⁻), arsenic (As), cadmium (Cd), chromium (Cr⁶⁺), lead (Pb) and mercury (Hg). The parameters were measured according to *Standard Methods* (APHA 1998).

Figure 1 | Study area and water quality monitoring sites.

Statistical technique

PCA is a statistical method that is used to reduce variable dimension (Thurston & Spengler 1985). PCA is used to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ranked with their importance in explaining the contributions to the independent variable so that the first few retain most of the variation present in all of the original variables (Jolliffe 2002; Wang et al. 2011). The computational procedure of PCA can be divided into three steps: (1) singular value decomposition; (2) dimension determination; (3) factor rotation. The varimax rotated method is commonly used in PCA (Statheropoulos et al. 1998; Sindosi et al. 2003).

Mathematical details of PCA and FA can be found in Thurston & Spengler (1985). Statistical analysis was carried out using SPSS 16.0 for Windows™ (SPSS Inc. 2007). Plots or graphs were completed using Microsoft Office Excel 2007.

Assessment of hydrochemical characteristics

The WQIIM is a new tool for general assessment of surface water. The calculation details as follows (Liu et al. 2010).

Single-factor water quality identification index (P_i)

The P_i consists of an integer and a decimal fraction. Pollution grade can be judged by the integer, and the difference of pollution degree in same grade can be judged by the decimal fraction. The P_i can be expressed by the following formula:

$$P_i = X_1.X_2.X_3$$

where X_1 is the integer, and shows the grade of water quality; X_2 is the decimal fraction, and shows the degree of monitoring data in interval of X_1 class water quality changing; X_3 is the comparison result of water quality grade and function goal grade.

(1) $X_1.X_2$ calculation

When the water quality grade is between I and V:

- (a) For the general indicators (except DO, pH, temperature and so on),

$$X_1.X_2 = a + \frac{C_i - C_{ls}}{C_{us} - C_{ls}}$$

- (b) For DO:

$$X_1.X_2 = a + 1 - \frac{C_i - C_{ls}}{C_{us} - C_{ls}}$$

where C_i is the monitoring value of i target; C_{us} is the upper limit of i target in water quality standard interval of class a; C_{ls} is the lower limit of i target in water quality standard interval of class a; $a = 1, 2, 3, 4, 5$, based on monitoring data and national standards.

When the water quality is worse than or equal to grade V:

- (a) For general indicators (except DO, pH, temperature and so on),

$$X_1.X_2 = 6 + \frac{C_i - C_{vus}}{C_{vus}}$$

where C_{vus} is the upper limit of i target in water quality standard interval of class V.

- (b) For DO:

$$X_1.X_2 = 6 + \frac{C_{vls} - C_i}{C_{vls}} m$$

where C_{vls} is the lower limit of i target in water quality standard interval of class V; m is the correction coefficient, $m = 5$ in this study.

- (2) X_3 calculation

$$X_3 = X_1 - f_1$$

where f_1 is the goal grade of water environment functional area. Note: when $X_3 > 9$, $X_3 = 9$.

Comprehensive water quality identification index (I_{wq})

The I_{wq} , a river water quality assessment index based on the single factor water quality identification index, can be calculated by the following formula:

$$I_{wq} = C_1.C_2.X_3.X_4 = \left[\frac{1}{m+1} \left(\sum_{i=1}^m P_i + \frac{1}{n} \sum_{j=1}^n P_j \right) \right] X_3.X_4$$

where $C_1.C_2$ shows comprehensive water quality index; P_i is the single factor water quality index of the main pollution indicator (that is $X_1.X_2$ in the single factor water quality identification index), and each indicator takes up one weight; m is the number of the main pollution indicators; P_j is the single factor water quality index of other indicators, and all the non-main pollution indicators take up one weight; n is the number of non-major pollution

indicators; X_3 is the number of indicators which cannot reach the water quality standards; X_4 represents the comparison results of water quality categories and function zoning category. $\text{NH}_4\text{-N}$, $\text{NO}_3\text{-N}$, COD, TP and TN were selected as main pollution indicators in this study.

Water quality grade can be determined based on the value of I_{wq} ; the classification basis and the corresponding comprehensive water quality grades were: $1.0 \leq I_{\text{wq}} \leq 2.0$, class I; $2.0 < I_{\text{wq}} \leq 3.0$, class II; $3.0 < I_{\text{wq}} \leq 4.0$, class III; $4.0 < I_{\text{wq}} \leq 5.0$, class IV; $5.0 < I_{\text{wq}} \leq 6.0$, class V; $6.0 < I_{\text{wq}} \leq 7.0$, class inferior V and not malodorous black; and $I_{\text{wq}} \leq 7.0$, class inferior V and malodorous black.

RESULTS AND DISCUSSION

Temporal variance

General water quality composition

For each year, 32 stations were monitored for April, August and October, which was done to investigate the impacts of agricultural practices and rainfall patterns on the stream water quality. The results in Table 1 indicated that the water quality in August was the best among the three periods except for COD, As, and Pb. The main reason for this is that, during the August flood season, great rainfall events increase non-point source pollutant runoff to the stream, but since August is the 'golden age' of plant growth, plant roots retain a portion of pollutants. Great rainfall events also increase stream flow rate and dilute contaminants, resulting in lower concentrations on average. In contrast, when all the agricultural activities are finished in October, plant root interception is not working; rainfall is also reduced and stream flow rate slowed down, so that the dilution effect disappears, and the water quality gets worse. In April contaminants accumulated over winter could start running off during the spring season due to soil thawing, thus resulting in high contaminant concentrations in stream water.

Temporal variations of water quality parameters

In PCA, projections of the original variables on the subspace of PCs are called component loadings and coincide with the correlation coefficients between PCs and variables. Component loadings of the first two retained PCs for each month are presented in Figure 2. In April, for example, the principal component 1 (PC1)

Table 1 | Mean, minimum (min), maximum (max), and standard deviation (SD) of the physicochemical parameters for the three hydrological periods; units: mg/L

Parameter	April					August					October				
	Sample size	Mean	Min	Max	SD	Sample size	Mean	Min	Max	SD	Sample size	Mean	Min	Max	SD
DO	314	8.63	< DL ^a	19.00	3.41	333	6.39	< DL	16.60	2.99	322	7.39	< DL	17.90	2.93
$\text{NH}_4\text{-N}$	319	5.44	< DL	169.00	17.47	329	1.29	< DL	60.20	4.94	237	2.38	< DL	83.50	9.23
$\text{NO}_3\text{-N}$	322	3.23	< DL	39.70	3.24	335	3.09	0.03	22.70	2.55	325	3.38	< DL	15.60	2.49
COD	322	9.13	0.60	58.10	11.25	335	17.42	1.00	840.00	65.04	333	6.64	0.60	109.10	11.95
TP	187	0.23	< DL	3.56	0.41	204	0.11	< DL	0.78	0.14	202	0.13	< DL	1.82	0.22
TN	128	8.78	1.67	59.30	9.43	157	6.30	0.47	36.10	4.61	149	7.00	0.91	56.80	6.28
F^-	297	0.68	< DL	3.94	0.60	309	0.62	0.09	4.67	0.49	297	0.72	< DL	3.82	0.53
As	282	0.007	< DL	0.26	0.0207	305	0.009	< DL	0.76	0.052	311	0.005	< DL	0.471	0.033
Cd	311	0.001	< DL	0.02	0.0021	319	0.001	< DL	0.04	0.004	333	0.001	< DL	0.20	0.011
Cr	311	0.021	< DL	0.14	0.0208	323	0.015	< DL	0.11	0.017	333	0.015	< DL	0.214	0.021
Pb	311	0.004	< DL	0.09	0.0109	325	0.006	< DL	0.22	0.018	333	0.003	< DL	0.07	0.010
Hg	260	0.00004	< DL	0.0024	0.0002	275	0.00002	< DL	0.00049	0.00006	287	0.00002	< DL	0.00047	0.00006

^a < DL: Lower than detection limit value. For the calculation purposes, values < DL were replaced with 0.

Figure 2 | Component loadings for the first component (PC1) and second component (PC2) in April, August and October ((a): PC1 in April; (b): PC2 in April; (c): PC1 in August; (d): PC2 in August; (e): PC1 in October; (f): PC2 in October).

was positively influenced by nutrient-related parameters (i.e. NH₄-N, TP, TN) and was negatively affected by DO and NO₃-N. Therefore, this component seems to measure the preponderance of physical and nutrient-related water quality parameters over the other parameters. This component also reveals that the Hg and Cd were less important in accounting for river water quality variations in April. PC2 was positively influenced by As, Cd and Pb while negatively by NO₃-N, TN, TP, and NH₄-N (Figure 2(b)). Component loading patterns obtained for

PC1 and PC2 in August (Figures 2(c) and 2(d)) and October (Figures 2(e) and 2(f)) were quite different from those in April, e.g. PC2 in August was positively influenced by anthropogenic inputs (i.e. NO₃-N and TN) and was negatively and largely impacted by COD. Such different results further reveal a highly seasonal variation of water quality parameters in this dynamic river system. Figure 2 also revealed that NH₄-N and TP were always the most important variables contributing to water quality variations in the LRB for all three periods.

Spatial variance

Hydrochemical evaluation of surface water quality in the LRB

Using WQIIM, we evaluated mean values of 12 hydrochemical parameters for all monitoring sites (Table 2). According to the national surface water quality standards of China, the LRB was divided into two categories: meeting standard (MS), which included the sites that were classified as I, II and III, and not meeting standard (NS) including the sites that were classified as IV, V and inferior V. There were 18 MS monitoring sites, and 14 NS monitoring sites (Table 2, Figure 1).

The results of WQIIM showed that M5 (1.900) had the best water quality in the LRB, and N6 (6.835), N9 (8.555) and N10 (13.494) were seriously polluted. Most of the seriously polluted sites are located in urban areas, mainly due to discharge of wastewater from domestic and industrial sources. Sites with low pollution level were located primarily in smaller tributaries, whereas sites of medium and high pollution levels were in the main river channel and the larger tributaries. In addition, according to the survey of SBHWR in 2010, the average content of *Escherichia coli* was 4,500 per litre, which met Chinese national water quality standards. In the NS sites, the

Escherichia coli content was more than 20,000 per litre in urban areas, which did not meet the standards. The pollutants were mainly from municipal sewage outlets.

Pollution level of rivers in the LRB

The average I_{wq} values of Luan River, Xingzhou River, Yixun River, Wulie River, Laoniu River, Liu River, Bao River, Sa River and Panjiakou-Daheiting Reservoir were calculated (Figure 3).

If $I_{wq} \leq 4$, it means that the river water quality was relatively good. So the results of WQIIM indicated that the Yixun and Wulie rivers as well as the Liu River were badly polluted. The Luan and Bao rivers were lightly polluted. Pollution in the Liu River was more serious than other rivers. The factors which caused Liu River pollution were complex, including direct discharges to the river from non-point sources (such as agricultural and urban runoff), and from point sources (such as industrial wastewater and municipal sewage). Panjiakou and Daheiting Reservoirs are located downstream on Luan River, and their water quality was superior to the incoming Luan River waters. The main reason was that the reservoirs play a role of dilution and purification. The results of WQIIM were in line with the actual situation of the study area. This showed the appropriateness and the effectiveness of this method.

Table 2 | Results of hydrochemical evaluation and their classification of water quality^a

Site	Result	Grade	Site	Result	Grade
M1	3.401	III	M17	3.601	III
M2	3.902	III	M18	3.502	III
M3	3.101	III	N1	4.012	IV
M4	3.301	III	N2	4.914	IV
M5	1.900	I	N3	5.524	V
M6	2.702	II	N4	5.724	V
M7	3.501	III	N5	5.622	V
M8	3.601	III	N6	6.835	Inferior V, not malodorous black
M9	3.802	III	N7	5.424	V
M10	3.803	III	N8	4.612	IV
M11	3.203	III	N9	8.555	Inferior V, and malodorous black
M12	3.001	III	N10	13.494	Inferior V, and malodorous black
M13	3.701	III	N11	4.612	IV
M14	3.301	III	N12	5.322	V
M15	3.702	III	N13	4.711	IV
M16	3.702	III	N14	4.213	IV

^aM1–M18, meeting standard (MS) sites; N1–N14, not meeting standard (NS) sites.

CONCLUSIONS

Our results in the LRB demonstrated that there are temporal and spatial variations of hydrochemistry in the LRB. It can be seen from the results of seasonal analysis that the concentrations of nutrients and some heavy metal indicators in August were lower than those in April and October. This situation is closely related to August being the golden age of plant growth and having high rainfall; thus the streams are probably affected by root absorption and river dilution. The results of the FA showed that in addition to $\text{NH}_4\text{-N}$ and TP the importance of parameters in each month is not the same. According to the obtained score of WQIIM, 18 sites were classified as 'meeting standard (MS)' while the other 14 sites were classified as 'not meeting standard (NS)'. Sites of low pollution level were located primarily in small tributaries, whereas sites of medium high pollution levels were in the main river channel and the large tributaries, Liu River being the most seriously polluted. So when river management measures are carried out, the government must firstly control the area where the worst water quality is detected.

Figure 3 | Variation in comprehensive water quality identification index (I_{wq}) of rivers. 1 – Luan River, 2 – Xingzhou River, 3 – Yixun River, 4 – Wulie River, 5 – Laoniu River, 6 – Liu River, 7 – Bao River, 8 – Sa River, 9 – Panjiakou–Dahelting Reservoirs.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 41071323) and the National Major Science and Technology Projects for Water Pollution Control and Management (No. 2012 ZX07029-002, 2012ZX07203-003). The authors thank the water quality monitoring center of Chengde Branch of Hebei Provincial Survey Bureau of Hydrology and Water Resources for providing monitoring data.

REFERENCES

- APHA 1998 *Standard Methods for the Examination of Water and Wastewater*, 20th ed. American Public Health Association/American Water Works Association/Water Environment Federation, Washington, DC, USA.
- Bu, H. M., Tan, X., Li, S. Y. & Zhang, Q. F. 2010 Temporal and spatial variations of water quality in the Jinshui River of the South Qinling Mts., China. *Ecotoxicol. Environ. Saf.* **73** (5), 907–913.
- Caccia, V. G. & Boyer, J. N. 2005 Spatial patterning of water quality in Biscayne Bay, Florida as a function of land use and water management. *Mar. Pollut. Bull.* **50** (11), 1416–1429.
- Dixon, W. & Chiswell, B. 1996 Review of aquatic monitoring program design. *Water Res.* **30** (9), 1935–1948.
- Huang, F., Wang, X. Q. & Lou, L. P. 2009 Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques. *Water Res.* **44** (5), 1562–1572.
- Jolliffe, I. T. 2002 *Principal Component Analysis*, 2nd ed. Springer-Verlag, New York.
- Kolovos, A., Christakos, G., Serre, M. L. & Miller, C. T. 2002 Computational Bayesian maximum entropy solution of a stochastic advectionreaction equation in the light of site-specific information. *Water Resour. Res.* **38** (18), 1318–1334.
- Li, J. & Feng, P. 2007 Runoff variations in the Luanhe River Basin during 1956–2002. *J. Geograph. Sci.* **17** (3), 339–350.
- Liu, X., Li, G., Liu, Z., Guo, W. & Gao, N. 2010 Water pollution characteristics and assessment of lower reaches in Haihe River Basin. *Pro. Environ. Sci.* **2**, 199–206.
- Muñoz-Carpena, R., Ritter, A. & Li, Y. C. 2005 Dynamic factor analysis of groundwater quality trends in an agricultural area adjacent to Everglades National Park. *J. Contamin. Hydro.* **80** (1–2), 49–70.
- Sindosi, O. A., Katsoulis, B. D. & Bartzokas, A. 2003 An objective definition of air mass types affecting Athens, Greece: the corresponding atmospheric pressure patterns and air pollution levels. *Environ. Technol.* **24** (8), 947–962.
- Singh, K. P., Malik, A., Mohan, D. & Sinha, S. 2004 Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – a case study. *Water Res.* **38** (18), 3980–3992.
- Singh, K. P., Malik, A. & Sinha, S. 2005 Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques – a case study. *Analytica Chim. Acta.* **538** (1–2), 355–374.
- SPSS Inc. 2007 *SPSS 16.0 Base User's Guide*. SPSS Inc., Chicago, IL, USA.
- Statheropoulos, A. M., Vassiliadis, N. & Pappa, A. 1998 Principal component and canonical analysis for examining air pollution and meteorological data. *Atmos. Environ.* **32** (6), 1087–1095.
- Thurston, G. D. & Spengler, J. D. 1985 A qualitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmos. Environ.* **19** (1), 9–25.
- Wang, H., Li, X. & Xie, Y. 2011 Hydrochemical evaluation of surface water quality and pollution source apportionment in the Luan River basin, China. *Water Sci. Technol.* **64** (10), 2119–2125.