Spatial and temporal variations of nitrogen pollution in Wen-Rui Tang River watershed, Zhejiang, China

Ping Lu · Kun Mei · Yujin Zhang · Lingling Liao · Bibo Long · Randy A. Dahlgren · Minghua Zhang

Abstract Water quality has degraded dramatically in Wen-Rui Tang River watershed, Zhejiang, China, especially due to rapid economic development since 1995. This paper aims to assess spatial and temporal variations of the main pollutants (NH_4^+-N, TN, BOD_5, COD_{Mn}, DO) of water quality in Wen-Rui Tang River watershed, using the geographic information system, cluster analysis (CA) and principal component analysis (PCA). Results showed that concentrations of BOD_5, COD_{Mn}, NH_4^+-N, and TN were significantly higher in tertiary rivers than in primary and secondary rivers. From April 2006 to March 2007, the concentrations of NH_4^+-N (2.25–57.9 mg/L) and TN (3.78–70.4 mg/L) in all samples exceeded Type V national water quality standards (≥ 2 mg/L), while 5.3% of all COD_{Mn} (1.83–27.5 mg/L) and 33.6% of all BOD_5 (0.34–50.4 mg/L) samples exceeded Type V national water quality standards ($COD_{Mn} \geq 15$ mg/L, $BOD_5 \geq 10$ mg/L). Monthly changes of pollutant concentrations did not show a clear pattern, but correlation analysis indicated that NH_4^+-N and TN in tertiary rivers had a significant negative correlation with 5-day cumulative rainfall and monthly rainfall, while there were no significant correlations in primary and secondary rivers. The results of CA and spatial analysis showed that the northern part of Wen-Rui Tang River watershed was the most seriously polluted. This region is characterized by the high population density and industrial and commercial activities. The PCA and spatial analysis indicated that the degraded water quality is caused by anthropogenic activities and poor wastewater management.

Keywords Wen-Rui Tang River · Water quality · CA · PCA · GIS

P. Lu · K. Mei · Y. Zhang · L. Liao · B. Long · R. A. Dahlgren · M. Zhang
The Environmental Geographic Information System (EGIS) Laboratory, School of Environmental Science and Public Health, Wenzhou Medical College, Wenzhou, China

P. Lu
e-mail: plu580@hotmail.com

R. A. Dahlgren · M. Zhang (✉)
Department of Land, Air and Water Resources, University of California, Davis, CA, 95616, USA
e-mail: mhzhang@ucdavis.edu

Introduction

Anthropogenic contamination of waterways constitutes a major problem in both developing and developed countries around the world (Barrett et al. 1998; Gelberg et al. 1999; Tilman et al. 2002; Grande et al. 2003; Fytianos and Christophoridis

2004; Camargo and Alonso 2006; Chen et al. 2006; Nas and Berktay 2006; Chang 2008; Ma et al. 2009). Declining surface water quality directly affects human and ecosystem health, which further impacts economic development and social wellbeing. Degraded water quality may induce various forms of illness, and exhibit reproductive and developmental toxicity (Horrigan et al. 2002; Fewtrell 2004; Avalanja and Bonner 2005). Excess nutrient and pollutant loadings have also been linked to several problems in aquatic ecosystems, including increased frequency and severity of harmful algal blooms, dramatic shifts in trophic relationships, direct toxicity to aquatic organisms, and expansion of coastal hypoxic zones (Boesch et al. 2001; Ocean Commission 2004; Diaz and Rosenberg 2008).

In China, the government has established an environmental monitoring system and carried out numerous water quality monitoring programs in recent years. Based on the monitoring program, the government has established a set of water quality standards to protect waterbodies and guide remediation activities. To fully understand water quality issues, many researchers have successfully investigated the application of multivariate and GIS analyses to address the spatial and temporal variations of water quality in various regions (Grande et al. 2003; Mingoti and Lima 2006; Kallioinen et al. 2006; Holbrook et al. 2006; Singh et al. 2006; Astel et al. 2007; Shreshtha

and Kazama 2007; Zhang et al. 2009). Many of these studies have been able to geographically link water quality deterioration with specific human activities, which can then be used to guide remediation efforts.

This study expands on previous research by applying cluster analysis and principal component analysis in a GIS environment to the urban-dominated Wen-Rui Tang River watershed of eastern China in order to: (1) characterize the water contamination status of the watershed, (2) examine spatio-temporal distributions of water contamination, and (3) investigate potential factors influencing contamination levels. The results provide a valuable tool for water quality agencies that will allow them to more effectively focus limited resources for remediation on severe water pollution problems at the watershed scale.

Materials and methods

Study area

Wen-Rui Tang River watershed in Wenzhou, China was selected as the study site due to its multiple water quality impairments and the availability of appropriate water quality and watershed characterization data (Fig. 1). The upper Wen-Rui Tang River watershed is dominated by
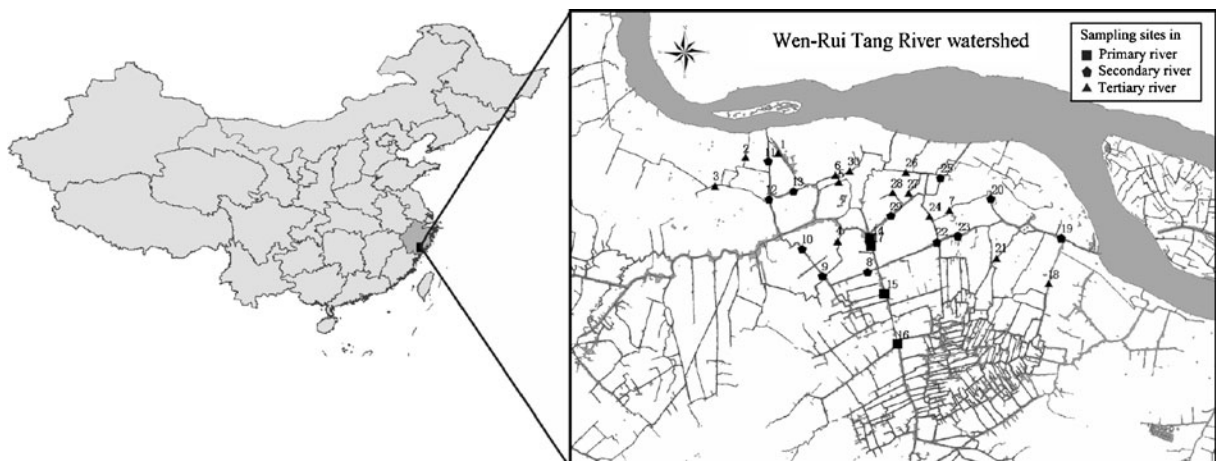

Fig. 1 Study area and monitoring sites

forests and agriculture while the lower watershed is an urban zone with a metropolitan population of about 7 million. It has a subtropical climate with mild, dry winters and hot, humid summers. Average highs are 12°C and 32°C in January and July, respectively. Precipitation is seasonal, with 78.2% of annual precipitation (1,800 mm) occurring from April to October. The Wen-Rui Tang River watershed has a catchments area of 353 km^2 and passing through Wenzhou city in the lower watershed. Twenty years ago, the Wen-Rui Tang River was the "mother" river supporting the city for transportation, aquaculture, agriculture, drinking water, and other aspects of daily life. However, due to rapid industrial growth and urbanization in recent years, the city has generated large volumes of wastewater from untreated sewage and service-oriented enterprises that are discharged directly into the river system. The river water does not currently meet the Type V national water quality standards (SEPBC 2002a), which is the lowest water quality standard that supports aquatic ecosystem health.

Data sources

Water quality data

Water quality data used for this analysis were obtained from the Wenzhou Municipal Plan-

ning Department. These data included six water quality parameters measured monthly at 30 water quality monitoring sites between April 2006 and March 2007 (Fig. 1). The six parameters of pH, dissolved oxygen (DO), 5-day biochemical oxygen demand (BOD$_5$), potassium permanganate—chemical oxygen demand index (COD$_{Mn}$), ammonia-nitrogen (NH$_4^+$-N), and total nitrogen (TN), were selected as important water quality indicators to assess the water quality status of the Wen-Rui Tang River. Nitrate was not analyzed as it is often below detection levels due to the anoxic nature of most waterways. The sampling, preservation, transportation and analysis of the water samples were performed following standard methods (SEPBC 2002b). The basic statistics of the 2,160 observations in the dataset are summarized in Table 1.

Influencing factors data

Watershed characterization data from the Wenzhou Municipal Planning Bureau included a set of GIS shapefiles of Wen-Rui Tang River watershed. Pollution source data included the regional population density, public toilets, sites of livestock farms, and sewage outfalls. The latter consisted mainly of domestic wastewater discharged into the Wen-Rui Tang River, and industrial and services wastewater that is mostly discharged into

Table 1 Statistical description (mean, SE, and N) of water quality parameters

		Primary river	Secondary river	Tertiary river
pH	Mean	6.77	6.74	6.85
	SE	0.029	0.015	0.018
	N	48	144	168
DO(mg/L)	Mean	1.09	1.06	1.00
	SE	0.110	0.093	0.125
	N	48	144	168
NH$_4^+$-N(mg/L)	Mean	6.60	9.14	14.10
	SE	0.246	0.447	0.765
	N	48	144	168
TN(mg/L)	Mean	7.99	11.33	17.22
	SE	0.326	0.575	0.940
	N	48	144	168
BOD$_5$(mg/L)	Mean	7.11	7.73	12.59
	SE	0.638	0.544	0.812
	N	48	144	168
COD$_{Mn}$(mg/L)	Mean	5.99	6.90	8.43
	SE	0.276	0.321	0.332
	N	48	144	168

the nearby Ou River through underground pipes, and not discharged into the Wen-Rui Tang River. Other factors influencing the water quality data included 5-day cumulative rainfall before sampling, monthly rainfall, normal high water mark, and river width.

Analysis methods

Independent t test

Independent *t* test analysis was used to determine pollutant concentration differences among primary, secondary, and tertiary rivers.

Pearson correlation

Pearson's correlation test was applied to determine the relationship between pollutant concentrations and influencing factors, and to test the relationship between a synthesis pollution index and rainfall in the primary, secondary, and tertiary rivers.

Cluster analysis

We used cluster analysis (CA) to analyze the similarities and grouping of samples in temporal and spatial dimensions. CA is an unsupervised pattern detection method that partitions all cases that are dissimilar to different groups (Lattin et al. 2003; McKenna 2003). Hierarchical cluster analysis was applied to the data after log transformation using Ward's method with squared Euclidean distances expressed as follows:

$$d_{ij}(2) = \left[\sum_{k=1}^{m} (x_{ik} - x_{jk})^2 \right]^{1/2}$$

$$(i, j = 1, 2 \cdots n) \tag{1}$$

where d_{ij} is the distance between the *i*th sample and the *j*th sample, x_{ik} is the *k*th parameter of the *i*th sample, x_{jk} is the *kth* parameter of the *j*th sample, and *i, j* = 1,2,3,...,30.

Principal component analysis

We use principal component analysis (PCA) to determine the minimum number of factors that will account for the maximum variance in the data. PCA is a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components (Helena et al. 2000; Zhou et al. 2007). PCA extracts eigenvalues and eigenvectors (a list of loading) from the covariance matrix of original variables to produce new orthogonal variables, known as varifactors (VFs). VFs are linear combinations of the original variables (Pekey et al. 2004; Zhou et al. 2007) and include hypothetical, unobservable, latent variables (Vega et al. 1998; Helena et al. 2000; Zhou et al. 2007).

An integrated evaluation model of PCA was applied following the method of Lin and Zhang (2005) to analyze the synthesis pollution condition of Wen-Rui Tang River watershed. The specific method consisted of three steps.

1. The corresponding coefficients were obtained from the following formula:

$$A1 = B1/SQR(\lambda_1), \ A2 = B2/SQR(\lambda_2) \ldots \tag{2}$$

   where A1/A2 are the corresponding coefficients matrix of parameters; B1/B2 are the loadings matrix of parameters on significant VF1 and VF2; and λ_1/λ_2 are the corresponding eigenvalues of the first and second principal components.

2. The parameters were standardized and multiplied by corresponding coefficients to obtain the expressions of the principle components using the following formula:

$$F1 = A11ZX1 + A12ZX2 + \ldots + A1nZXn \tag{3}$$

$$F2 = A21ZX1 + A22ZX2 + \ldots + A2nZXn \tag{4}$$

where F1/F2 are integrated values of parameters in component1 and component2; ZXn is the z-scale standardized value of the nth parameter; and A1n/A2n are the nth parameter's corresponding coefficients in matrix A1 and matrix A2, respectively.

3. The ratio of eigenvalues to the weights of principal components and integrated value of parameters were calculated by the following formula:

$$F = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \cdots \lambda_n} F_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2 + \cdots \lambda_n} F_2$$
$$+ \cdots + \frac{\lambda_n}{\lambda_1 + \lambda_2 + \cdots \lambda_n} F_n \qquad (5)$$

where F is the integrated value of parameters; F_n is the integrated value of parameters in the nth principal component; and λ_n is the corresponding eigenvalue.

In GIS, overlay analysis and spatial analysis were applied. Spatial analysis of water contamination was performed by interpolation of sampling points by the algorithmic method of "Inverse Distance Weighted" (IDW). Independent t test, CA, PCA, and Pearson's correlation were carried out using SPSS 13.0, and GIS analysis used ArcGIS 9.2 (ESRI, San Diego, USA).

Water quality data preparations

Due to experimental failures, several values were missing from the water quality dataset. Missing data were estimated using temporally interpolated average values. Before multivariate statistical analysis, the normal distribution of each variable was checked by analyzing skewness and kurtosis statistics (Lattin et al. 2003; Papatheodorou et al. 2006). Except for pH, which is a logarithmic function of H+ activity, the distributions of all other parameters were not normal and had skewness and kurtosis values greater than zero. Therefore, all parameters except for pH were transformed as $x' = \log(x)$ (Kowalkowski et al. 2006; Papatheodorou et al. 2006). After log transformation, the skewness and kurtosis values of all of these parameters, except DO, were significantly reduced to ranges

from −0.454 to 0.777 and 0.206 to 0.621, respectively. Because the distribution of DO diverged so far from normality, this parameter was excluded from the analysis. For PCA all selected parameters were z-scale standardized (mean = 0 and variance = 1), which renders the data dimensionless and minimizes the effects of differences in measurement units and variance (Zhou et al. 2007; Zhang et al. 2009). Before the PCA analysis, the Kaiser–Meyer–Olkin (KMO) and Bartlett's Sphericity tests were performed on the parameter correlation matrix in order to examine the validity of the subsequent PCA.

Results

Spatial and temporal variation of the water quality parameters

In the Wen-Rui Tang River, pH values (mean 6.8; range 6.1–7.8) fell within a normal range for natural waters, so pH was not considered further by independent analysis. Results showed that the concentrations of NH_4^+-N, TN, BOD_5, and COD_{Mn}, were consistently higher in tertiary rivers than in primary and secondary rivers (Table 1) and all differences were statistically significant ($p < 0.001$).

NH_4^+-N

Mean NH_4^+-N concentrations in the primary, secondary and tertiary rivers were 6.60, 9.14, and 14.10 mg/L, respectively (Table 1). NH_4^+-N concentrations ranged from 2.25 to 57.9 mg/L, with all sites and sample times exceeding the Type V NH_4^+-N national standard (≥ 2 mg/L; Table 2). About 12.5% of all samples, all of which were from the northern part of Wen-Rui Tang River watershed, exceeded the national health standard (≥ 1 mg/L) by 20-fold, while 38.6% exceeded the national health standard by tenfold. The average value for all sites by month was the lowest in June (4.94 mg/L) and highest in April (15.49 mg/L). There was a decline from April to June, but no regular pattern of changes in NH_4^+-N levels was apparent throughout the remainder of the

Table 2 The percentage of sampling sites that below type V national standards

| Month | The sampling sites below type V national standards | | | |
	NH_4^+-N (≥ 2 mg/L; %)	TN (≥ 2 mg/L; %)	BOD_5 (≥ 10 mg/L; %)	COD_{Mn} (≥ 15 mg/L; %)
April	100	100	40	10
May	100	100	40	3
June	100	100	10	0
July	100	100	17	3
August	100	100	57	0
September	100	100	40	0
October	100	100	23	7
November	100	100	33	3
December	100	100	33	7
January	100	100	40	13
February	100	100	33	3
March	100	100	37	13

year (Fig. 2). Pearson's correlation test showed a significant negative correlation between NH_4^+-N concentration and monthly rainfall, and a very significant negative correlation between NH_4^+-N concentration and 5-day cumulative rainfall in tertiary rivers. However, the correlations between NH_4^+-N and 5-day cumulative rainfall were not significant for the primary and secondary rivers (Table 3). The mean monthly values at each site show that the most seriously polluted regions were in the northern part of the Wen-Rui Tang River watershed (Fig. 3). In contrast, NH_4^+-N concentration was low and its monthly changes were small in the southern region, while the concentration was higher and the monthly changes were quite large in the northeast and northwest (Fig. 4).

TN

Mean TN concentrations in the primary, secondary and tertiary rivers were 7.99, 11.33, and 17.22 mg/L, respectively (Table 1). Based on these mean concentrations, NH_4^+-N comprised from 81% to 83% of total nitrogen for all types of rivers. TN values for all sites and months ranged from 3.78–70.4 mg/L, so all exceeding the Type V national water quality standard of ≥ 2 mg/L (Table 2). A total of 17.5% of all samples, 96.8% of which were from the northern part, exceeded the national health standard (≥ 1 mg/L) by 20-fold, while 51.9% of all samples exceeded the health standard by tenfold. The mean value for all sites

by month was lowest in June (6.33 mg/L) and the highest in April (20.68 mg/L). Similar to NH_4^+-N, there was a decline from April to June, but no regular pattern of change for the remainder of the year (Fig. 2). Pearson's correlation test showed a significant negative correlation between TN concentration and either monthly rainfall or 5-day cumulative rainfall in the tertiary rivers. However, the correlations between these values in the primary and secondary rivers were not significant (Table 3). The mean monthly values at each site show that the most seriously polluted regions were in the northern part of Wen-Rui Tang River watershed (Fig. 3). Concentration changes at each site among months were similar to NH_4^+-N as NH_4^+-N comprises the largest fraction of TN (Fig. 4).

BOD_5

BOD_5 values ranged from 0.34–50.4 mg/L, with 33.61% of all samples exceeding the Type V national standard (≥ 10 mg/L), and 73.33% of all samples exceeding the health standard (≥ 4 mg/L). Mean values for primary, secondary, and tertiary rivers were 7.11, 7.73, and 12.59 mg/L, respectively (Table 1). Changes in mean monthly values for different sites displayed no apparent pattern (Fig. 2). Pearson's correlation test showed a significant negative correlation between the BOD_5 levels and 5-day cumulative rainfall in secondary and tertiary rivers (Table 3). The mean monthly

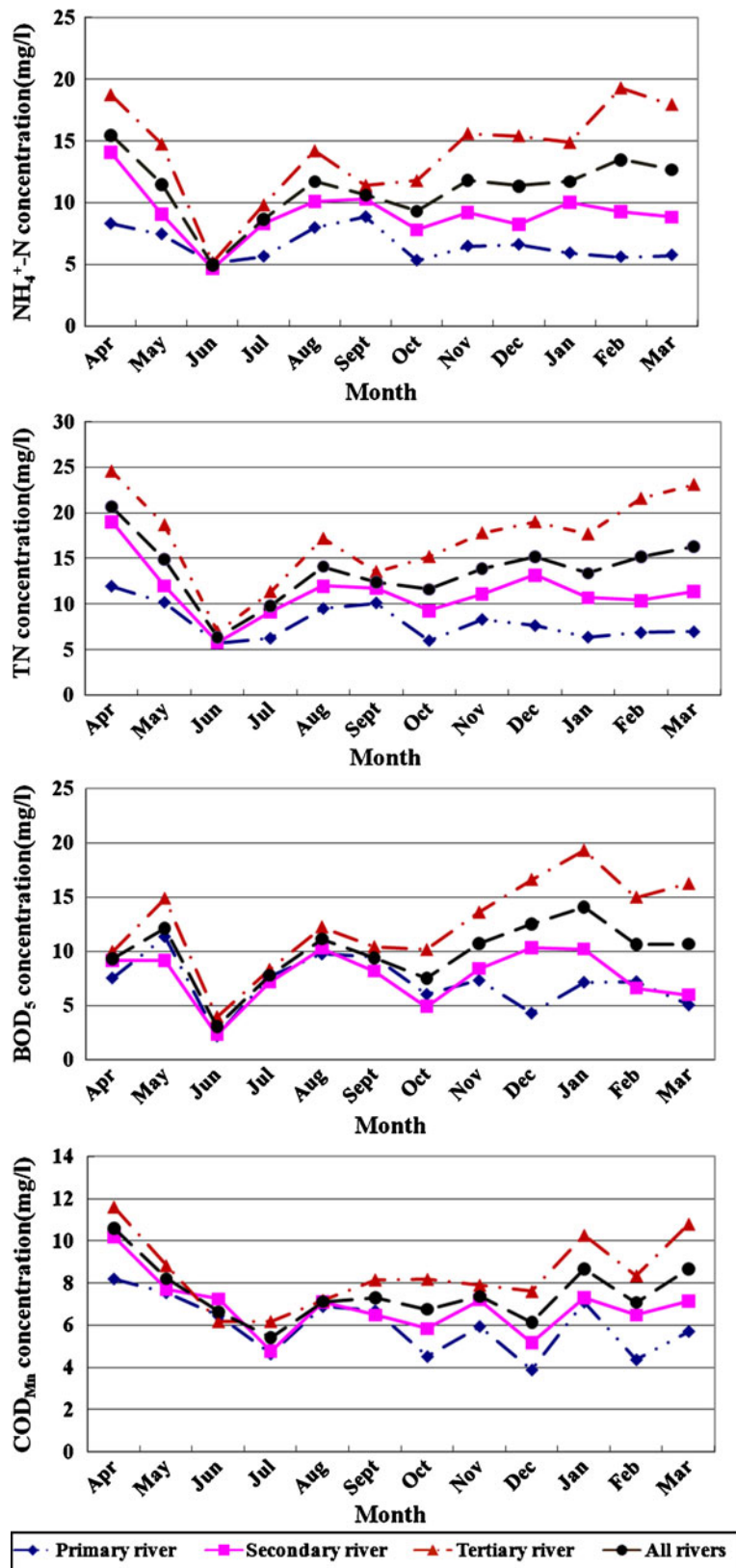**Fig. 2** Temporal variations of BOD$_5$, COD$_{Mn}$, NH$_4^+$-N, TN

Table 3 Correlations of pollutants and rainfall in primary, secondary, and tertiary rivers

Pearson correlation		Primary river	Secondary river	Tertiary river	Monthly rainfall	5-day cumulative rainfall
NH_4^+-N	Primary river	1				
	Secondary river	0.702[a]	1			
	Tertiary river	0.245	0.688[a]	1		
	Monthly rainfall	0.234	−0.209	−0.597[a]	1	
	5-day cumulative rainfall	−0.117	−0.572	−0.718[b]	0.474	1
TN	Primary river	1				
	Secondary river	0.813[b]	1			
	Tertiary river	0.455	0.766[b]	1		
	Monthly rainfall	0.173	−0.218	−0.576[a]	1	
	5-day cumulative rainfall	−0.180	−0.519	−0.674[a]	0.474	1
BOD_5	Primary river	1				
	Secondary river	0.597[a]	1			
	Tertiary river	0.234	0.637[a]	1		
	Monthly rainfall	0.267	−0.131	−0.554	1	
	5-day cumulative rainfall	−0.420	−0.626[a]	−0.631[a]	0.474	1
COD_{Mn}	Primary river	1				
	Secondary river	0.840[b]	1			
	Tertiary river	0.459	0.677[a]	1		
	Monthly rainfall	0.251	−0.107	−0.468	1	
	5-day cumulative rainfall	0.229	0.095	−0.390	0.474	1

[a] Correlation is significant at the 0.05 level (two-tailed)

[b] Correlation is significant at the 0.01 level (two-tailed)

values at each site show that the most seriously polluted regions were in the northern part of Wen-Rui Tang River watershed (Fig. 3). The BOD_5 levels were highest in the northern region and the monthly changes were quite large in almost all regions (Fig. 4).

COD_{Mn}

Mean COD_{Mn} values for primary, secondary and tertiary rivers were 5.99, 6.90, and 8.43 mg/L, respectively (Table 1). The values of COD_{Mn} ranged from 1.83 to 27.5 mg/L, with only 5.3% samples, of which 94.7% were from the northern region, exceeding the Type V national water quality standard (≥ 15 mg/L), and 57.8% of all samples exceeding the national health standard (≥ 6 mg/L). The changes in mean values of different sites among months showed no apparent temporal pattern (Fig. 2). COD_{Mn} concentration and rainfall values did not have any significant correlations (Table 3). The mean monthly values at each site were similar to BOD_5 values (Fig. 3), while fluctuations between sites in different re-

gions and different months did not show a pattern as obvious as that observed for BOD_5 (Fig. 4).

Spatial similarities and grouping

Because the mainstem of the river (primary river) has very few sampling sites (four sites), and two of them are very close, we only analyzed the secondary and tertiary rivers sites using Cluster Analysis (CA).

BOD_5

The CA results for BOD_5 at different sites on the secondary and tertiary rivers are shown in Fig. 5. The sites in the secondary (Fig. 5a) and tertiary (Fig. 5b) rivers have been divided into two broad categories: polluted and seriously polluted clusters. Spatial clustering of BOD_5 levels at all monitoring sites (ungrouped treatment) also indicated two broad types of monitoring sites (Fig. 5c). Similar results from both grouped and ungrouped treatment of sites showed that the seriously polluted sites are located in the northern part of
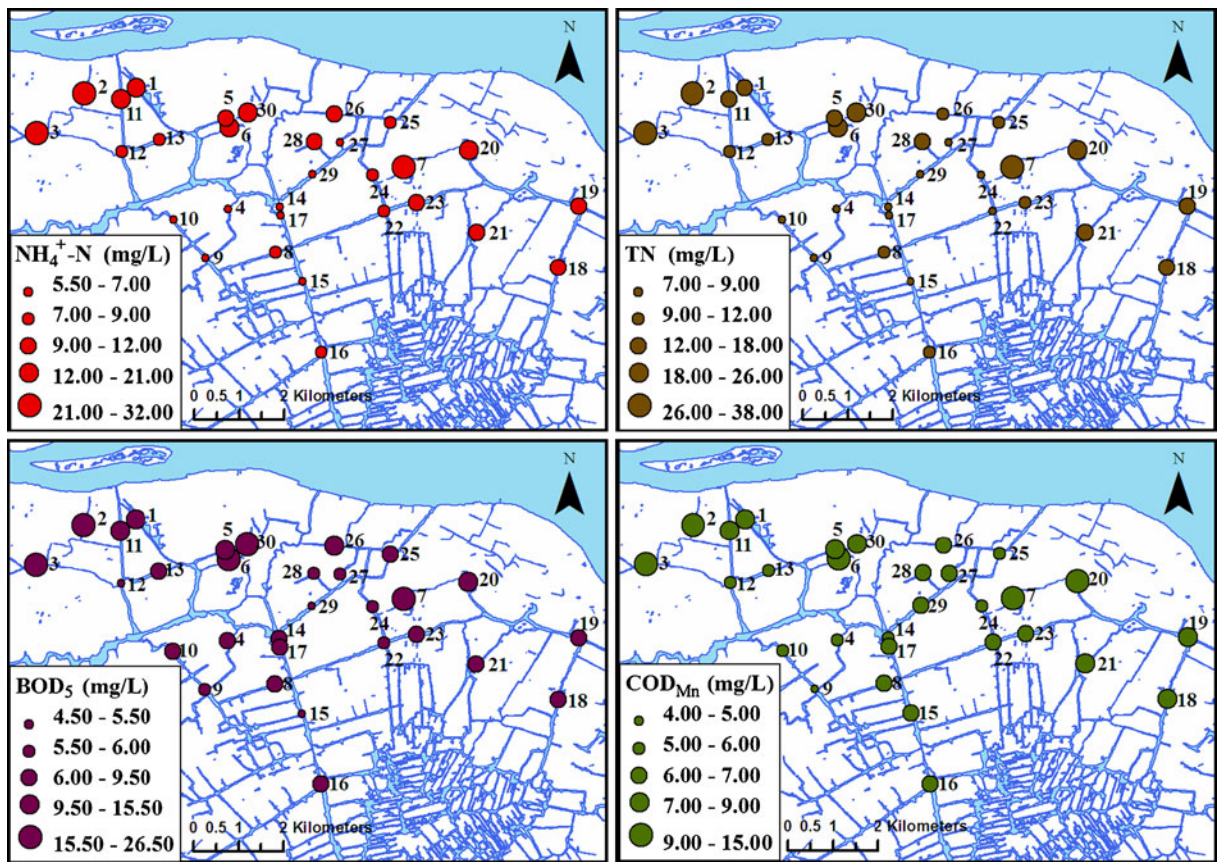
Fig. 3 Spatial variations of BOD$_5$, COD$_{Mn}$, NH$_4^+$-N, TN

the Wen-Rui Tang River watershed (Fig. 6a, b). Temporal clustering of BOD$_5$ from Apr 2006 to Mar 2007 (Fig. 7), shows that the BOD$_5$ pollution condition in June is very different from the other months.

COD$_{Mn}$

The CA results for COD$_{Mn}$ are shown in Fig. 6c, d. As with BOD$_5$, the secondary and tertiary rivers were separated into two categories: polluted and seriously polluted sites (seriously polluted sites indicated by black dots in Fig. 6d). Analysis of all monitoring sites for COD$_{Mn}$ (ungrouped treatment) yielded the same two broad categories of polluted and seriously polluted sites (Fig. 6c). Thus, the grouped and ungrouped treatments yield the same result. The seriously polluted sites are located in the northwest and northeast portions of the Wen-Rui Tang River watershed

(Fig. 6c, d). COD$_{Mn}$ concentration showed no regular temporal pattern over the study period (Fig. 7).

NH$_4^+$-N

The results of CA for NH$_4^+$-N were similar to those obtained for BOD$_5$ and COD$_{Mn}$ with division into polluted and seriously polluted sites (Fig. 6f). Analysis of all sites (ungrouped treatment) yielded the same two broad types, and all of the seriously polluted sites were located in northern part of the Wen-Rui Tang River watershed (Fig. 6e). However, compared to the grouped treatment, the ungrouped treatment yields more seriously polluted sites (Fig. 6e, f). CA of NH$_4^+$-N levels shows more seriously polluted sites compared to the results for other pollutants (Fig. 6e, f). The concentration of NH$_4^+$-N in June is significantly less than in other months (Fig. 7).
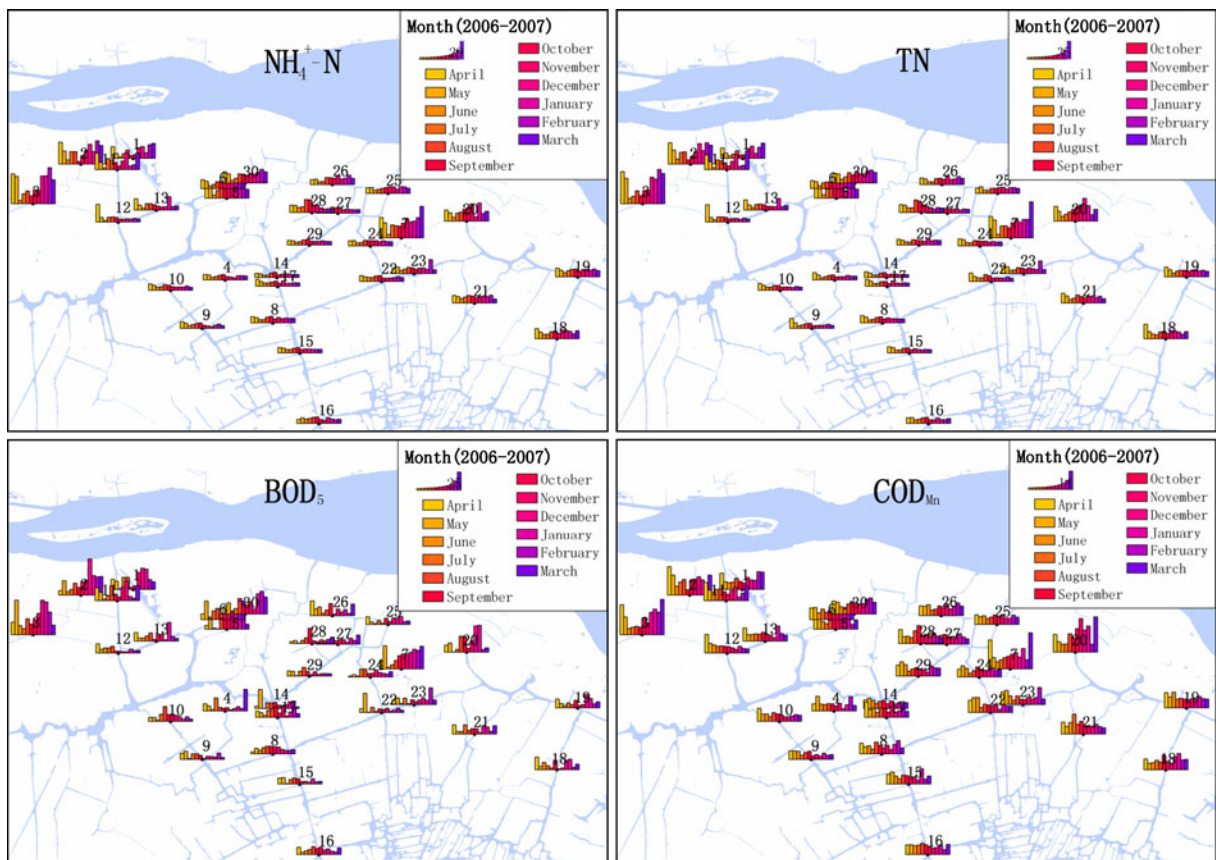
Fig. 4 Temporal and spatial variation of BOD_5, COD_{Mn}, NH_4^+-N, TN. Note: Histogram from left to right is the concentration from 2006.4 to 2007.3, respectively

TN

The CA results for TN were similar to the results for BOD_5, COD_{Mn}, and NH_4^+-N. All secondary and tertiary river sites clustered into polluted and seriously polluted types (Fig. 6g, h), and the seriously polluted sites were in the northern part of the watershed (Fig. 6g, h). The results of grouped and ungrouped treatments of sites showed that the seriously polluted sites were also very similar (Fig. 6g, h). In the temporal clustering, the concentration of TN in June is much lower than that in other months (Fig. 7).

All pollutants

Spatial clustering of all pollutants and sites reveals three heavily polluted regions: the northern region (which includes sampling sites 5, 6, and

30), the northeastern region (which includes sites 7 and 20), and the northwestern region (which includes sites 1, 2, 3, and 11; Fig. 6i).

Influencing factors

The KMO results for Type I (samples drawn from gray sites in Fig. 6i) and Type II (samples drawn from black sites in Fig. 6i) sites were 0.770 and 0.689, respectively. The Bartlett's Sphericity results for Type I and Type II sites were 390 and 639 ($p < 0.01$), respectively. These results confirm that PCA provided significant reductions in dimensionality.

PCA was applied to three different datasets (Type I, Type II, and all sites) to examine differences between the two site types and to identify the latent influencing factors. According to the eigenvalue-one criterion (eigenvalue > 1) PCA of
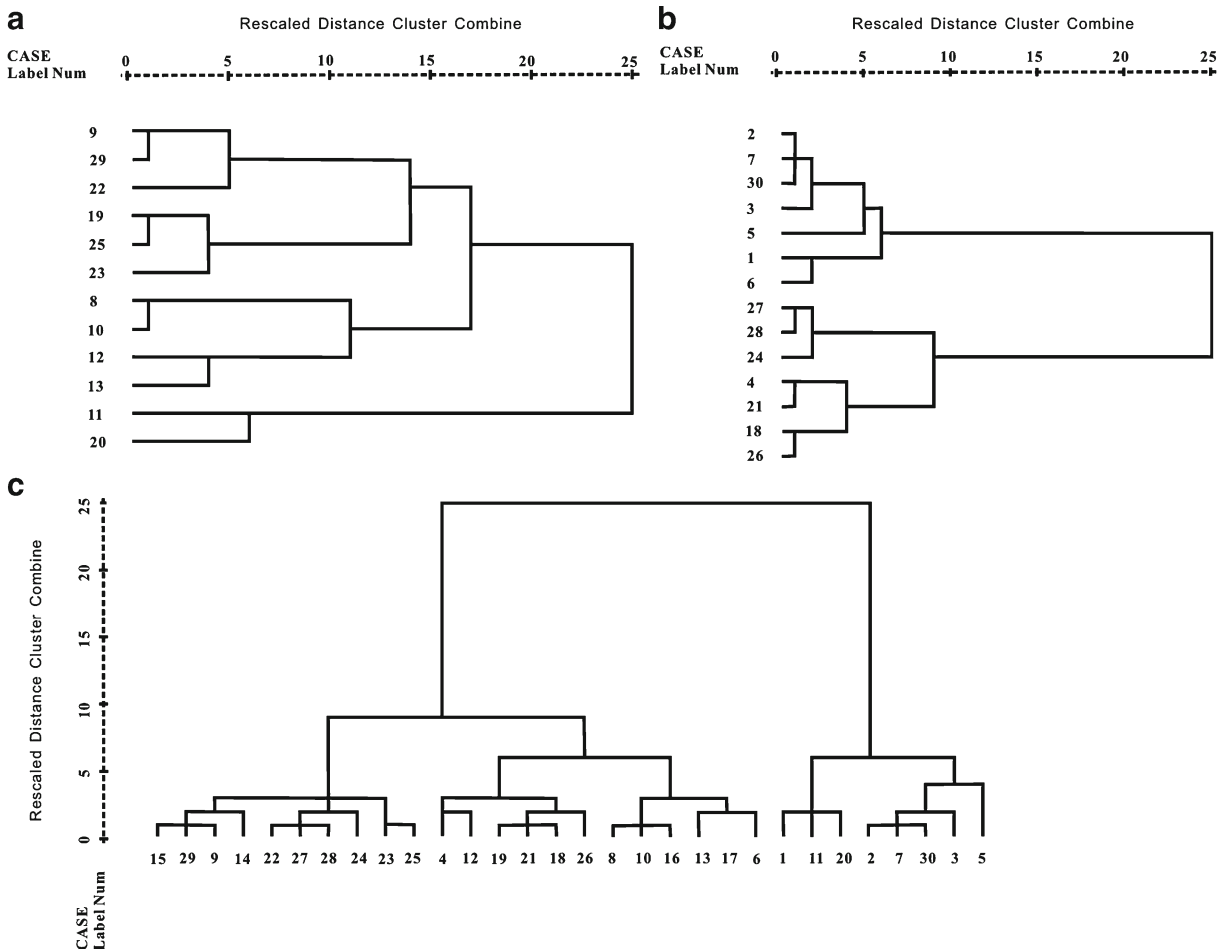
Fig. 5 Dendrogram showing spatial clustering of monitoring sites of BOD_5 to secondary rivers (**a**), tertiary rivers (**b**) and all rivers (**c**)

the three data sets yielded one PC (Table 4) that accounted for 63.8%, 53.3%, and 64.7% of the total variance in Type I, Type II, and all site data sets, respectively. In the VF1, the loading value of TN is highest in Type I (0.956), the loading value of NH_4^+-N is highest in Type II (0.938), and the loading value of NH_4^+-N is highest in all sites (0.949; Table 5).

An integrated evaluation model of the PCA was made in order to obtain the synthetic pollution index for the sampling sites, using the formula:

$$F = 0.26ZX_1 + 0.41ZX_2 + 0.46ZX_3$$
$$+ 0.53ZX_4 + 0.53ZX_5 \qquad (6)$$

in which the coefficients were calculated according to formulas (2), (3), (4), and (5).

Pollution index values were interpolated between sampling points by Arc GIS 9.2 to create month and year pollution index maps for the period April 2006 to March 2007 (Figs. 8, 9). The pollution index maps show that the more seriously polluted region is in northern Wenzhou city, demonstrating consistency between results from PCA and CA. Figure 8 indicates that pollution levels of the river varies by month, with the lowest levels in June, but that the temporal variation is different among Type I and Type II sites. In Type I sites, pollution is heaviest in April, May, August, and September, while it varies slightly in both the
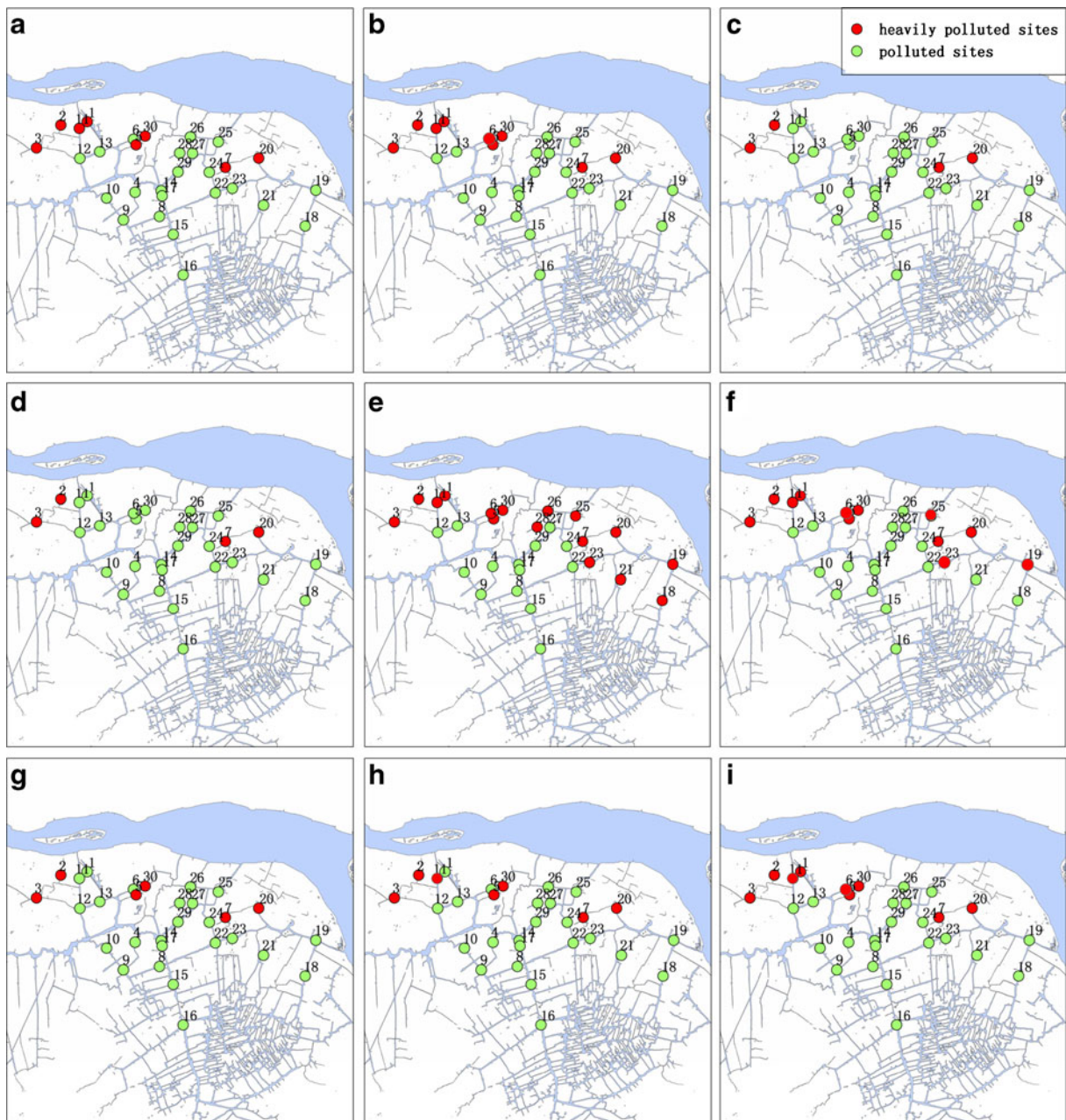
Fig. 6 Comparison of monitoring sites maps of grouped and ungrouped treat of sites of cluster analysis. Note: **a** result of spatial clustering of all monitoring sites of BOD$_5$; **b** result of spatial clustering of monitoring sites of BOD$_5$ to secondary rivers and tertiary rivers; **c** result of spatial clustering of all monitoring sites of COD$_{Mn}$; **d** result of spatial clustering of monitoring sites of COD$_{Mn}$ to secondary rivers and tertiary rivers; **e** result of spatial clustering of all monitoring sites of NH$_4^+$-N; **f** result of spatial clustering of monitoring sites of NH$_4^+$-N to secondary rivers and tertiary rivers; **g** result of spatial clustering of all monitoring sites of TN; **h** result of spatial clustering of monitoring sites of TN to secondary rivers and tertiary rivers; **i** result of spatial clustering of all monitoring sites and all pollutants. *Red spots* are the seriously polluted sites
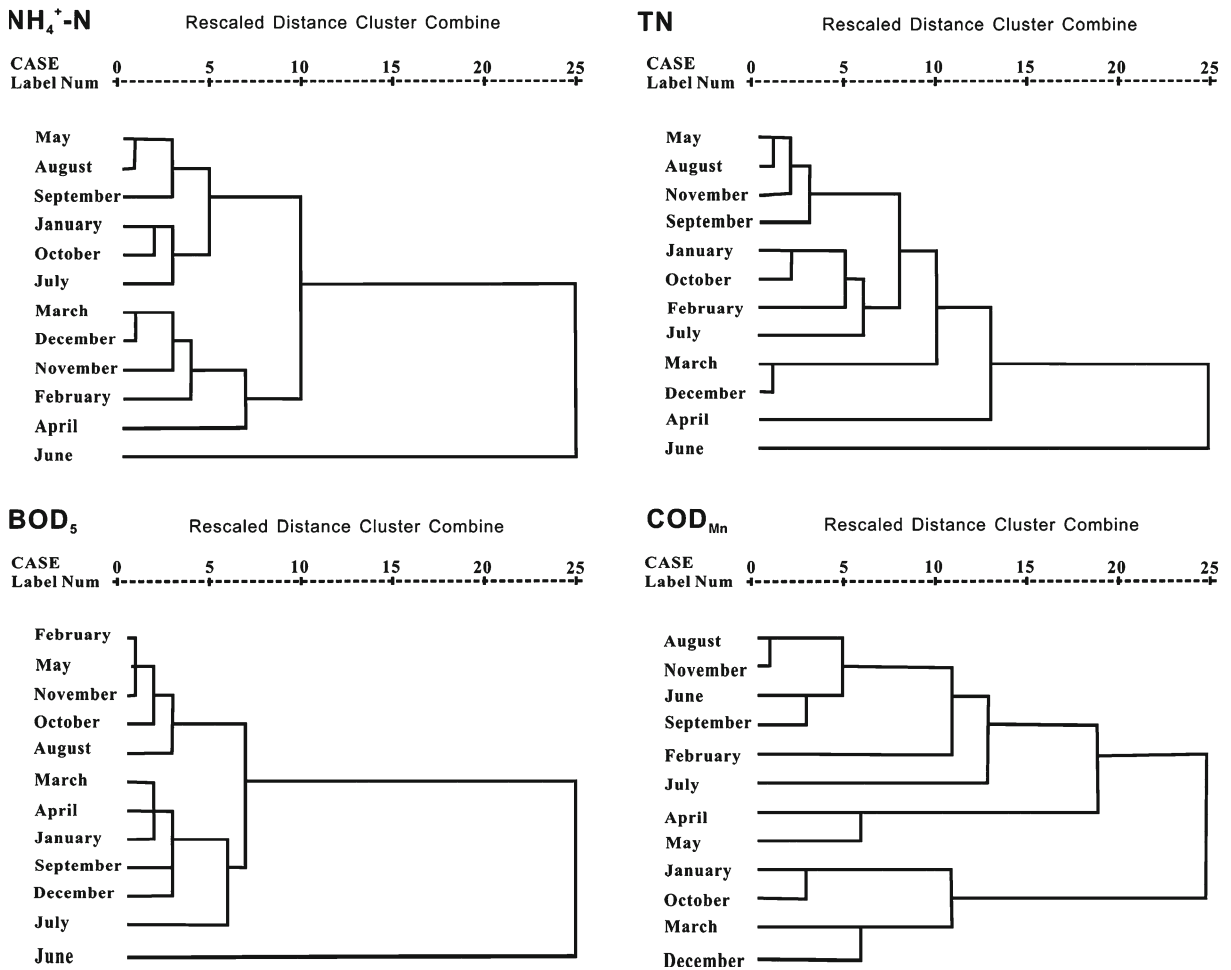
Fig. 7 Dendrogram showing temporal clustering of all months of NH_4^+-N, TN, BOD_5, COD_{Mn}

rainy and dry seasons. In Type II sites, pollution is the most serious in April, decreases in May, reaches its lowest levels in June, and then rises gradually from July to April of the following year. Arc GIS 9.2 analysis shows that pollution is more

serious in blue waterways, followed by green and yellow waterways (Fig. 9).

The levels of individual pollutants showed relationships with each other. Among BOD_5, COD_{Mn}, NH_4^+-N, and TN, the correlational

Table 4 Total variance explained in type I, type II, and all sampling sites

Component	Initial eigenvalues type I		Initial eigenvalues type II		Initial eigenvalues all	
	Total	% of variance	Total	% of variance	Total	% of variance
1	3.191[a]	63.817	2.664[a]	53.286	3.233[a]	64.655
2	0.962	19.231	0.897	17.948	0.841	16.822
3	0.508	10.160	0.837	16.733	0.576	11.513
4	0.286	5.719	0.529	10.589	0.308	6.161
5	0.054	1.073	0.072	1.444	0.042	0.848

[a]The yielded PC according to the eigenvalue-one criterion (the eigenvalue > 1)

Table 5 Loadings of 5 parameters on significant VFs for type I, type II, and all sampling sites

Parameters	Type I	Type II	All
	VF1	VF1	VF1
pH	0.345	0.403	0.474
BOD_5	0.776	0.503	0.734
COD_{Mn}	0.818	0.727	0.820
NH_4^+-N	0.942	0.938	0.949
TN	0.956	0.918	0.947

Type I sample drawn from green sites in Fig. 6i, *Type II* sample drawn from red sites in Fig. 6i

dependences are all quite remarkable, with Pearson's correlation coefficients of: BOD_5 and COD_{Mn}: $r_p = 0.511$, BOD_5 and NH_4^+-N: $r_p = 0.692$, BOD_5 and TN: $r_p = 0.684$, COD_{Mn} and NH_4^+-N: $r_p = 0.753$, COD_{Mn} and TN: $r_p = 0.785$, and NH_4^+-N and TN: $r_p = 0.96$, $P < 0.01$, $n = 360$; Table 6). This study also revealed that the pollution levels show positive correlation with population density, the number of public toilets and sewage outfalls, while they show negative correlation with river width, normal high water mark, monthly rainfall, and 5-day cumulative rainfall before sampling (Table 6).

Correlation analysis of the synthetic pollution index and rainfall in primary, secondary, and tertiary rivers only showed a correlation for tertiary rivers. The correlation between the synthesis pollution index and 5-day cumulative rainfall was quite remarkable ($r_p = -0.789$, $P < 0.01$; Table 7).

Discussion

Pollution of Wen-Rui Tang River

In the Wenzhou urban area, industrial and services wastewaters are primarily discharged into the Ou River through underground pipes, and does not enter the Wen-Rui Tang River. Agricultural pollution sources are mainly located in the upstream rural areas, so they have relatively little impact on the study area. Only a small portion of the study region, in the northeast of the city, was impacted by animal husbandry.

The highly degraded water quality directly affects human health and ecosystem health in the watershed. Although we do not have direct human health impact data, the results indicated that 100% of NH_4^+-N, 100% of TN, 58% of COD_{Mn} and 73% of BOD_5 values exceeded the national health standard. This strongly suggests the need for water managers to pay attention to the environment to reduce the potential impact to human health in the watershed.

Quantitative data on flow velocity were not available, however, it is obvious from visual inspection that the water in the tertiary rivers was almost immobile except during a typhoon, rainstorm, or human-flushing events. The high pollutant concentrations in tertiary rivers are most likely to be associated with the longer hydrological residence times in these river branches. Even though a best management practice of flushing the rivers was adopted in this watershed, it has proven difficult to implement in the tertiary rivers as these streams represent dead-end streams that prevent adequate flushing.

NH_4^+-N pollution was serious throughout the entire region, mainly caused by the untreated excreta in sewage directly discharged into the river, in addition to the excreta from some avian and livestock farms. The TN pollution was also serious throughout the entire region, mainly resulting from these same sources in addition to its generation from the decomposition of the organic matter, such as food residue discarded directly into waterways. The most seriously polluted regions are concentrated in the northern part of the Wen-Rui Tang River watershed, due to high population density, and industrial and commercial activities. The northwestern part of the study area has a large textile industry, while the northeastern part has many avian and livestock farms.

High BOD_5 and COD_{Mn} levels were mostly due to organic matter from excreta and food residue from household sources, hotel sewage, excreta from farms and industrial sewage in the
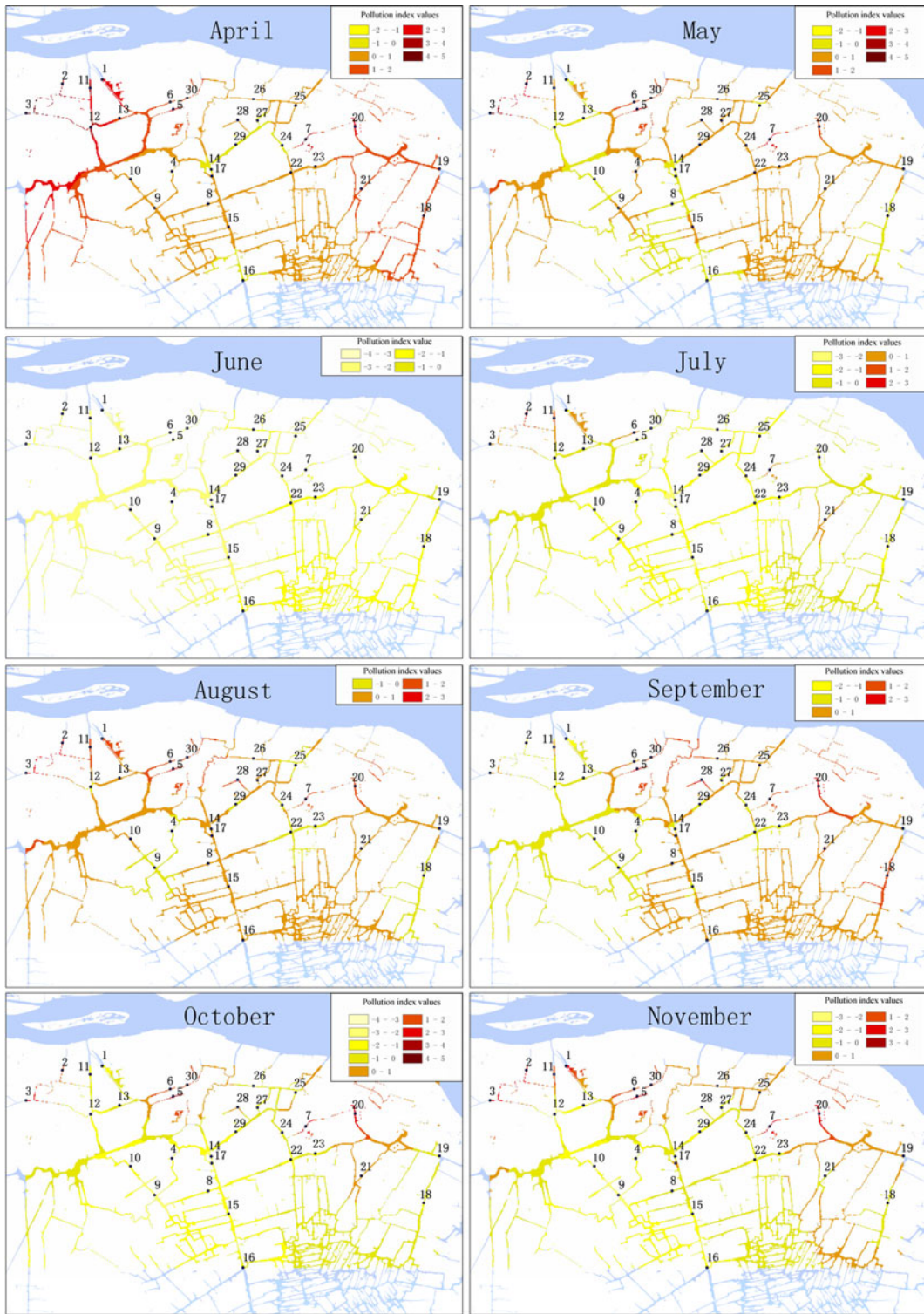
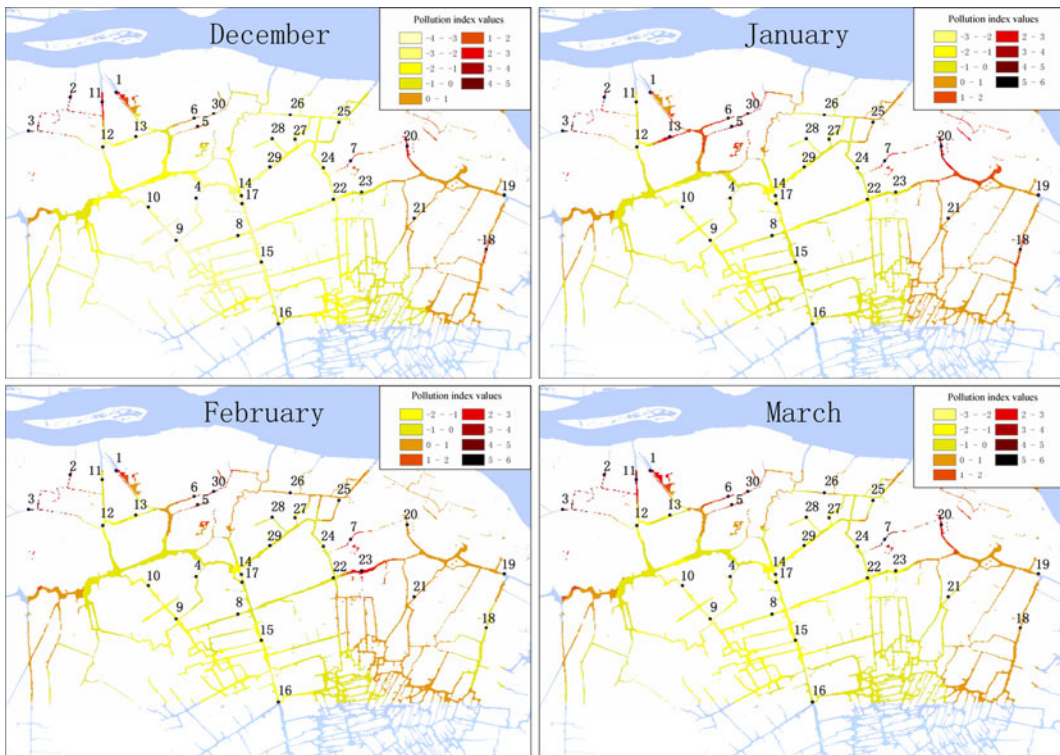**Fig. 8** Synthetic pollution in Wen-Rui Tang River watershed for each month

Fig. 8 (continued)
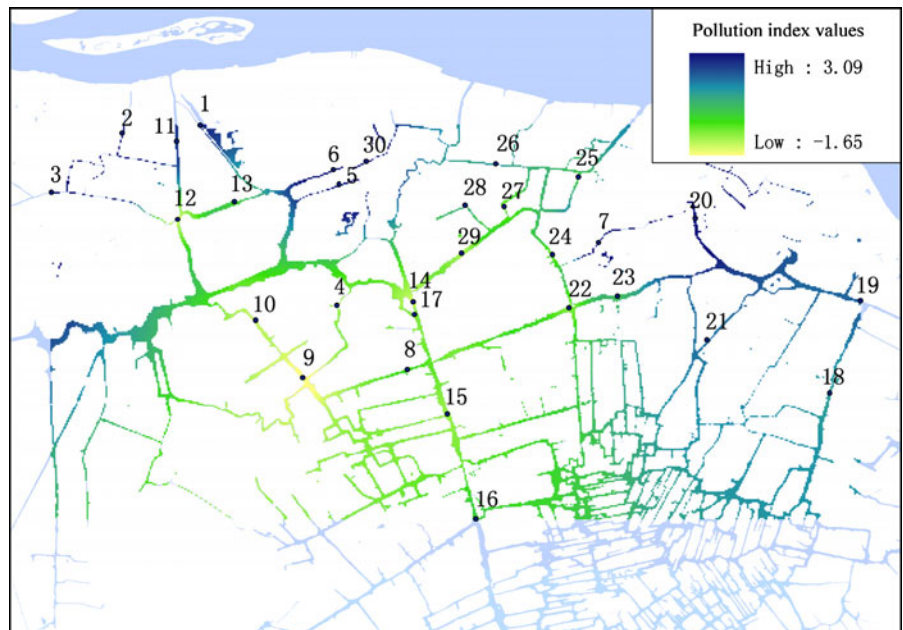

Fig. 9 Synthetic pollution in Wen-Rui Tang waterways

Table 6 Correlations of pollutants, polluter and environment factors

Pearson correlation	River width	Normal high water mark	Population density	Public toilets (500-m circle)	Sewage outfalls (500-m circle)	Monthly rainfall	5-day cumulative rainfall	BOD$_5$	KMnO$_4$	NH$_4^+$-N	TN
River width	1										
Normal high water mark	0.890^a	1									
Population density	−0.391^a	−0.357^a	1								
Public toilets	−0.391^a	−0.271^a	0.898^a	1							
Sewage outfalls	−0.245^a	−0.169^a	0.659^a	0.656^a	1						
Monthly rainfall	0.000	0.000	0.000	0.000	0.000	1					
5-day cumulative rainfall	0.000	0.000	0.000	0.000	0.000	0.474^a	1				
BOD$_5$	−0.298^a	−0.188^a	0.210^a	0.302^a	0.229^a	−0.124^b	−0.218^a	1			
KMnO$_4$	−0.204^a	−0.145^a	0.081	0.155^a	0.056	−0.093	−0.052	0.511^a	1		
NH$_4$-N	−0.383^a	−0.268^a	0.178^a	0.273^a	0.203^a	−0.150^a	−0.220^a	0.692^a	0.753^a	1	
TN	−0.377^a	−0.263^a	0.171^a	0.274^a	0.185^a	−0.148^a	−0.215^a	0.684^a	0.785^a	0.960^a	1

[a] Correlation is significant at the 0.01 level (two-tailed)

[b] Correlation is significant at the 0.05 level (two-tailed)

northwest. The BOD$_5$ method is not influenced strongly by ammonia oxidization due to an initial lag (>5 days) in the oxidation kinetics of ammonia. Organic pollution is serious only in the northern area where levels are above the Type V national water quality standard (Table 2), due to the sources listed above. In this study, the values of BOD$_5$ are higher than COD$_{Mn}$ at many sampling sites. This indicates that only part of the organic matter in the water is oxidized by the potassium permanganate methodology (SEPBC 2002b).

Pollution by BOD$_5$, NH$_4^+$-N and TN were lowest in June throughout the entire region (Fig. 7). This may be due to the continual rainfall before sampling during this month. Monthly fluctuations in pollutant concentrations do not follow obvious patterns, but NH$_4^+$-N, TN, and BOD$_5$ in tertiary rivers have significant correlations with 5-day cumulative rainfall, and NH$_4^+$-N and TN in tertiary rivers have significant correlations with monthly rainfall. However, most pollutants do not have significant correlations with 5-day cumulative rainfall and monthly rainfall in primary and secondary rivers. This is because the government opens sluice gates to flush primary and secondary rivers when pollution is quite serious. As a result, the relations of pollutant concentrations and rainfall are not as distinct in primary and secondary rivers. Opening the sluice gates to flush very narrow branches is less effective, so pollutant concentrations are more significantly influenced by rainfall in tertiary rivers.

In PCA, VF1 of both Type I and Type II sites have very high positive loadings on NH$_4^+$-N and TN (>0.90). The loading value of TN on Type I sites is highest (0.956), while loading of NH$_4^+$-N on Type II sites is highest (0.938; Table 5). This shows that the main source of NH$_4^+$-N pollution is excreta in the northern area (Type II sites), which has a high population density, a temporary population, and avian and livestock farms. The major pollutant in the southern region was TN, and the NH$_4^+$-N in excreta is also a major pollutant.

Comparison to previous research

The study area is primarily a network of urban waterways in the Wen-Rui plain, which has many relatively small tributaries. The flow of water

Table 7 Correlations of synthetic pollution index and rainfall in primary, secondary, and tertiary rivers

Pearson correlation	Primary river	Secondary river	Tertiary river	Monthly rainfall	5-day rainfall	Cumulative
Primary river	1					
Secondary river	0.884[a]	1				
Tertiary river	0.619[b]	0.756[a]	1			
Monthly rainfall	0.070	−0.052	−0.559	1		
5-day cumulative rainfall	−0.248	−0.501	−0.789[a]	0.474	1	

[a]Correlation is significant at the 0.01 level (two-tailed)

[b]Correlation is significant at the 0.05 level (two-tailed)

throughout the entire Wen-Rui Tang River watershed is not natural, but is controlled by reservoirs in the upper watershed and the operation of sluice gates in the lower watershed. The river is largely stagnant during dry periods when the sluice gates are closed, while water flows freely when the sluice gates are opened. So the direction of flow is regulated by the position of the opened sluice gates. These hydrologic characteristics of the watershed are entirely different from other published studies. During most of the time period covered by this study, the sluice gates were closed and water bodies were relatively stagnant, effectively eliminating any effects from upstream flushing and pollution sources. As the opening and closing of floodgates can alter pollutant concentrations, the change of pollutant concentrations between months observed in this study was not as predictable as in other studies unaffected by floodgate operations (Girija et al. 2007; Zhang et al. 2009).

In the present paper, an attempt was made to assess the spatial and temporal water quality changes by multivariate statistics and GIS (Sarkar et al. 2007). Zhang et al. (2009) used CA, DA, and PCA to analyze water quality and identify pollution sources in the Daliao River basin, China. However, comprehensive applications of different multivariable statistical methods, GIS analysis and concrete studies of each pollutant have not been fully explored in river studies in China.

The use for environmental protection

The Wen-Rui Tang River is the principal source of drinking water in this region, though it is polluted very seriously at present. Most drinking water is secured from reservoirs in the upper watershed to avoid the severe pollution associated with the lower watershed. For improving the water quality in any watershed, it is critical to understand the relationships between the environmental quality and economic development. The deteriorating environment can affect the sustainable development of the regional economy. In order to achieve sustainable economic development, the government should seriously consider the environmental costs of economic development, thus integrating social, economic and ecology values.

Results indicated that focusing on the management of domestic sewage would be an effective way of reducing both the nitrogen and organic pollution in this watershed under the current economic and population conditions. The strengthening of water quality regulations in the northern part of the Wen-Rui Tang River watershed, periodic dredging of the rivers in these regions, and water body flushing (especially of tertiary waterways) during the flood season, should be able to alleviate the pollution situation in these areas. Developing efficient water treatment plants and sewage conveyance systems is another critical and effective means to protect the environment and the water quality within the watershed.

Conclusions

Our study is one of the few to analyze the spatial and temporal variations of water quality in urban-dominated waterways of China, aiming to ultimately prevent the adverse health and ecological

effects of water pollution. Results showed that: (1) the entire area in Wen-Rui Tang River watershed is seriously polluted, and the pollution in the northern part is most serious; (2) the values of BOD_5, COD_{Mn}, NH_4^+-N, and TN are significantly higher in tertiary rivers than those in primary and secondary rivers; (3) monthly changes of pollutant concentrations are irregular; and (4) most pollutants are negatively correlated with 5-day cumulative rainfall and monthly rainfall values in tertiary rivers, but not in primary and secondary rivers because of human-induced water flushing events from opening of floodgates.

This study reveals that degraded water quality is related to both anthropogenic activities and poor wastewater management. The lack of proper management policies and wastewater infrastructure during the economic development period may have played a role. The results of this study provide information for the mitigation of further degradation and to improve the existing water quality in the watershed.

References

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 datasets. *Water Research, 41*, 4566–4578.

Avalanja, M. C. R., & Bonner, M. R. (2005). Pesticides and human cancers. *Cancer Investigation, 23*, 700–711.

Barrett, J. H., Parslow, R. C., McKinney, P. A., Law, G. R., & Forman, D. (1998). Nitrate in drinking water and the incidence of gastric, esophageal, and brain cancer in Yorkshire, England. *Cancer Causes and Control, 9*, 153–159.

Boesch, D. F., Brinsfield, R. B., & Magnien, R. E. (2001). Chesapeake Bay eutrophication: Scientific understanding, ecosystem restoration, and challenges for agriculture. *Journal of Environmental Quality, 30*, 303–320.

Camargo, J. A. , & Alonso, A. (2006). Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment. *Environment International, 32*, 831–849.

Chang, H. (2008). Spatial analysis of water quality trends in the Han River basin, South Korea. *Water Research, 42*, 3285–3304.

Chen, J., Wang, Y., & Zhang, H. (2006). Overview on the studies of nitrate pollution in groundwater. *Progress in Geography, 25*, 34–44.

Diaz, R. J., & Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems. *Science, 321*, 926–929.

Fewtrell, L. (2004). Drinking-water nitrate, methemoglobinaemia, and global burden of disease: A discussion. *Environmental Health Perspectives, 112*, 1371–1374.

Fytianos, K., & Christophoridis, C. (2004). Nitrate, arsenic and chloride pollution of drinking water in Northern Greece. Elaboration by applying GIS. *Environmental Monitoring and Assessment, 93*, 55–67.

Gelberg, K. H., Church, L., Casey, G., London, M., Roerig, D. S., Boyd, J., et al. (1999). Nitrate levels in drinking water in rural New York State. *Environmental Research, 80*, 34–40.

Girija, T. R., Mahanta, C., & Chandramouli, V. (2007). Water quality assessment of an untreated effluent impacted urban stream: The Bharalu tributary of the Brahmaputra River, India. *Environmental Monitoring and Assessment, 130*, 221–236.

Grande, J. A., Borrego, J., Morales, J. A., & de la Torre, M. L. (2003). A description of how metal pollution occurs in the Tinto-Odiel rias (Huelva-Spain) through the application of cluster analysis. *Marine Pollution Bulletin, 46*, 475–480.

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 Research, 34*, 807–816.

Holbrook, R. D., Yen, J. H., & Grizzard, T. J. (2006). Characterizing natural organic material from the Occoquan watershed (Northern Virginia, US) using fluorescence spectroscopy and PARAFAC. *Science of the Total Environment, 361*, 249–266.

Horrigan, L., Lawrence, R. S., & Walker, P. (2002). How sustainable agriculture can address the environmental and human health harms of industrial agriculture. *Environmental Health Perspectives, 110*, 445–456.

Kallioinen, M., Huuhilo, T., Reinikainen, S. P., Nuortila-Jokinen, J., & Mänttäri, M. (2006). Examination of membrane performance with multivariate methods: A case study within a pulp and paper mill filtration application. *Chemometrics and Intelligent Laboratory Systems, 84*, 98–105.

Kowalkowski, T., Zbytniewski, R., Szpejna, J., & Buszewski, B. (2006). Application chemometrics in river water classification. *Water Research, 40*, 744–752.

Lattin, J., Carroll, D., & Green, P. (2003). *Analyzing multivariate data*. New York: Duxbury.

Lin, H. M., & Zhang, W. L. (2005). Differences from principal component analysis and component analysis and SPSS. *Statistical Research, 3*, 65–68.

Ma, J. Z., Ding, Z. Y., Wei, G. X., Zhao, H., & Huang, T. M. (2009). Sources of water pollution and evolution of water quality in the Wuwei basin of Shiyang river, Northwest China. *Journal of Environmental Management, 90*, 1168–1177.

McKenna, J. (2003). An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environmental Modelling and Software, 18*, 205–220.

Mingoti, S. A., & Lima, J. O. (2006). Comparing SOM neural network with fuzzy c-means, K-means and traditional hierarchical clustering algorithms. *European Journal of Operational Research, 174*, 1742–1759.

Nas, B., & Berktay, A. (2006). Groundwater contamination by nitrates in the city of Konya, (Turkey): a GIS perspective. *Journal of Environmental Management, 79*, 30–37.

Ocean Commission (2004). *An ocean blueprint for the 21st century*. Final Report. U.S. Commission on ocean policy. Washington, DC.

Papatheodorou, G., Demopoulou, G., & Lambrakis, N. (2006). A long-term study of temporal hydrochemical data in a shallow lake using multivariate statistical techniques. *Ecological Modelling, 193*, 759–776.

Pekey, H., Karakas, D., & Bakog, L. M. (2004). Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses. *Marine Pollution Bulletin, 49*, 809–818.

Sarkar, B. C., Mahanta, B. N., Saikia, K., Paul, P. R., & Singh, G. (2007). Geo-environmental quality assessment in Jharia coalfield, India, using multivariate statistics and geographic information system. *Environmental Geology, 51*, 1177–1196.

SEPBC (2002a). Environmental quality standards for surface water; Prepared and published jointly by: State Environment Protection Bureau of China (SEPBC), State Quality Supervision, Inspection and Quarantine Bureau of China (SQSIQBC), Beijing.

SEPBC (2002b). Water and wastewater analysis method. Prepared and published by: State Environment Protection Bureau of China (SEPBC), China Environmental Science Press, Beijing.

Shreshtha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environmental Modelling and Software, 22*, 464–475.

Singh, K. P., Malik, A., Singh, V. K., & Sinha, S. (2006). Multi-way data analysis of soils irrigated with wastewater—a case study. *Chemometrics and Intelligent Laboratory Systems, 83*, 1–12.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R., & Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature, 418*, 671–677.

Vega, M., Pardo, R., Barrado, E., & Deban, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research, 32*, 3581–3592.

Zhang, Y., Guo, F., Meng, W., & Wang, X. Q. (2009). Water quality assessment and source identification of Daliao river basin using multivariate statistical methods. *Environmental Monitoring and Assessment, 152*, 105–121.

Zhou, F., Guo, H. C., Liu, Y., & Jiang, Y. M. (2007). Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong. *Marine Pollution Bulletin, 54*, 745–756.