ORIGINAL PAPER

Correlations between algae and water quality: factors driving eutrophication in Lake Taihu, China

Y. P. Li · C. Y. Tang · Z. B. Yu · K. Acharya

Abstract Rapid population increase and economic growth in eastern China has lead to the degradation of many water bodies in the region, such as Lake Taihu, the third largest freshwater lake in China. Using data from recent investigations, the correlations between algae (measured as chlorophyll-a) and water quality indices in Lake Taihu were described by multivariate statistical analyses, and the key driving factors for the lake eutrophication were identified by principal component analysis. Results revealed strong spatiotemporal variation in the correlations between algae and water quality indices, suggesting that the limiting factor for the dominant algae growth depends on seasonality and location and it is necessary to reduce both nitrogen and phosphorus inputs for a long-term eutrophication control in this hyper-eutrophic system. Water temperature was another important controlling factor for algal growth in the lake. Using principal component analysis, nutrient contaminations from anthropogenic and natural inputs were identified as the key driving factor for the water quality problems of the lake. Moreover, five principal components were extracted and characterized with high spatial and seasonal variations in Lake Taihu. The key driving factors were believed to influence spatial variations including heavily polluted areas located in the northern and northwestern parts of the lake, where many manufacturing factories were built and wastewater from domestic and industrial plants was discharged. Based on this analysis, attention should be paid to effective land management, industrial wastewater treatment, and macrophytic vegetation restoration to reduce the pollutant loads and improve water quality. Principal component analysis was found to be a useful and effective method to reduce the number of analytical parameters without notably impairing the quality of information in this study.

Keywords Cyanobacteria bloom · Large shallow lakes · Nutrient limitation · Principal component analysis

Y. P. Li · C. Y. Tang
Key Laboratory of Integrated Regulation and Resource
Development on Shallow Lakes, Ministry of Education,
Hohai University, Nanjing 210098, China

Y. P. Li (✉) · C. Y. Tang
College of Environment, Hohai University,
Nanjing 210098, China
e-mail: liyiping@hhu.edu.cn

Z. B. Yu
Department of Geoscience, University of Nevada, Las Vegas,
NV 89119, USA

Z. B. Yu
College of Hydrology and Water Resource, Hohai University,
Nanjing 210098, China

K. Acharya
Division of Hydrologic Sciences, Desert Research Institute,
Las Vegas, NV 89119, USA

Introduction

Taihu Basin has become the most industrialized and urbanized area in China (Qin et al. 2007). By 2000, 10 % of the gross domestic product (GDP) of China was generated in this basin, while it accounted for only 0.4 % of the land surface and 3 % of the population (NSB 2000). Lake Taihu, located in the Taihu Basin, is the third largest lake in China with a water surface area of 2,338 km², an annual average water depth of 1.9 m,

and a mean hydraulic retention time of 300 days (Mao et al. 2008). It provides critical services to China's economy. However, Lake Taihu has been subject to pollutant loading from an array of nonpoint and point sources, and experienced severe water quality degradation over the past several decades. The lake water quality is continuing to deteriorate; eventually, the water quality could decline to a level where the lake will no longer provide potable drinking water or serve as a habitat for fish and birds (Dong et al. 2008; Hu et al. 2008). For example, the algal bloom events that occurred during the summer of 2007 led to the restriction of water supplies for approximately five million residents in Wuxi city (CD 2007).

The worsening condition of Lake Taihu has prompted researchers and government officials to seek mitigation strategies. The large regions of the lake and its watershed have limited previous research investigations to local scales not encompassing the entire watershed (Dong et al. 2008; James et al. 2009), to limited water quality indices using data from a small number of monitoring stations (Chen et al. 2003; Zhu 2008), or to remote sensing analyses over short time periods (Wang et al. 2008; Wang and Shi 2008). The spatiotemporal distributions of and interactions between algal concentrations and water quality indices based on long-term monitoring data that encompass the entire watershed area and the key driving factors over season need to be further studied (Paerl et al. 2011a).

Evaluating multiple pollutants at different monitoring stations in a watershed concurrently is a complex task that can help in understanding the mechanisms behind the spatial variations of those pollutants (Razmkhah et al. 2010). Various external factors impact water quality; alternatively, these factors may have annual trends and fluctuations depending upon local conditions, seasonal variations, and hydrodynamic circulation, as well as chemical and biological processes. The volume of monitoring data and the complexity of relationships among variables poses a barrier to comprehensive evaluation of pollutants (Fan et al. 2010). Multivariate statistical techniques, such as principal component analysis (PCA), are effective tools to reduce the complexity of large-scale data sets to enable their evaluation, spatial analysis, and temporal analysis (Unmesh et al. 2006). The PCA technique is widely used in environmental management studies and is useful to identify characteristics of the major factors that affect water quality (Xu et al. 2009).

The objectives of this study were to (1) investigate the spatiotemporal distributions of correlations between chlorophyll-*a* (Chl-*a*) and water quality indices, (2)

clarify the main principal components from all of the water quality parameters and characterize the spatio-temporal distribution of those indices, (3) identify the key driving factors for the lake eutrophication due to natural or anthropogenic influences of point or non-point sources of pollution via the results of PCA.

Materials and methods

To address the objectives listed above, lake and watershed data were collected from a wide range of sources including lake monitoring efforts, publications, and local agencies (Tables 1, 2). Hydrological and water quality data were collected from 58 monitoring stations with 32 in tributaries and 26 in the lake (Fig. 1). The runoff from the 32 tributaries accounted for approximately 90 % of the total runoff input to and output from the lake (Qin et al. 2007), while the 26 stations in the lake covered most of the lake regions from the littoral area to open water, including Zhushan Bay, Meiliang Bay, Gonghu Bay, Central lake, Western lake, South lake and East Taihu Bay. Hydrologic data for tributaries included average monthly flow discharge, water level, flow velocity and direction. Water quality samples were collected monthly for the tributary and in-lake monitoring stations from 2000 to 2005 which included dissolved oxygen (DO), pH, water temperature (WT), Secchi depth (SD), total nitrogen (TN), total phosphorus (TP), phytoplankton biomass and Chl-*a*. In-lake water quality samples were collected at 0.5 m depth, which is considered to be representative of the water column due to the shallow and well-mixed nature of the lake (Li et al. 2011b). A YSI-6600 multi probe was used to measure DO and WT. TN, TP, and Chl-*a* were measured using the Chinese standard methodology for lake eutrophication surveys (Jin and Tu 1990). This method is similar to the American standard methods (APHA 1998) for those parameters (James et al. 2009).

To address the roles of water quality parameters and limiting factors in algal growth in different seasons and lake regions, temporal trends between monthly phytoplankton biomass and TN, TP, air temperature were analyzed. In addition, the spatial distributions of correlations between Chl-*a* and water quality were analyzed for the warm (March to September) and cold seasons (October to February) using Microsoft Excel 2003. Correlation coefficients between parameters for each monitoring station in the lake were calculated using multivariate statistical analyses. Each correlation coefficient was calculated based on 72 data sets of monthly observed values from 2000 to 2005 (12×6). The equation for the correlation coefficient is:

Int. J. Environ. Sci. Technol. (2014) 11:169–182

171

Table 1 Water quality parameter annual means with standard deviations for Lake Taihu from 2000 to 2008

Year	TN (mg/L)	TP (mg/L)	Chl-a (μg/L)	Phytoplankton biomass (mg/L)	SD (cm)	pH	DO (mg/L)	WT (C)
2000	2.68 ± 2.23	0.102 ± 0.080	21.87 ± 29.7	6.80 ± 7.34	42.3 ± 24.7	7.88 ± 0.43	9.96 ± 2.23	16.7 ± 8.5
2001	2.49 ± 2.17	0.096 ± 0.095	12.13 ± 14.6	10.02 ± 8.57	41.9 ± 25.1	8.02 ± 0.73	9.09 ± 2.83	18.3 ± 8.7
2002	2.91 ± 2.55	0.100 ± 0.093	17.16 ± 22.1	6.45 ± 4.47	42.9 ± 24.2	8.00 ± 0.38	8.42 ± 2.75	19.0 ± 7.5
2003	3.39 ± 2.76	0.104 ± 0.090	21.31 ± 24.2	7.16 ± 4.26	37.3 ± 23.6	7.94 ± 0.39	8.65 ± 2.76	17.3 ± 9.0
2004	3.57 ± 2.87	0.121 ± 0.118	29.32 ± 41.7	7.77 ± 6.30	33.5 ± 21.6	8.10 ± 0.45	8.82 ± 2.58	17.0 ± 8.3
2005	3.19 ± 2.32	0.100 ± 0.077	31.49 ± 40.4	11.50 ± 6.51	33.1 ± 23.8	8.20 ± 0.37	9.50 ± 2.79	17.3 ± 9.7
2006	2.85 ± 0.93	0.097 ± 0.020	48.6 ± 59.8	10.83 ± 5.51	–	–	–	–
2007	2.35 ± 0.88	0.082 ± 0.035	23.1 ± 10.7	8.59 ± 5.21	–	–	–	–
2008	2.42 ± 0.89	0.072 ± 0.019	25.6 ± 16.5	–	–	–	–	–

All data come from the database discussed in Table 2

Chl-a Chlorophyll-a, *TN* total nitrogen, *TP* total phosphorus, *DO* dissolved oxygen, *SD* Secchi depth, *WT* water temperature "–" means no data

Table 2 Information about databases for Lake Taihu, China

Data type	Year(s)	Data coverage	From	Publications
Hydrology	2000–2005	Lake and Tributaries	TBAMWSC	Water quality bulletins
Water quality	2000–2005	Lake and Tributaries	TBAMWSC	Water quality bulletins
GIS map	2005	Taihu Basin	TBAMWSC	–
Meteorology	1980–2000	Taihu basin	DSCCWR	http://www.waterdata.cn/
Pollutant source	2000	Taihu basin	DSCCWR	http://www.waterdata.cn/
Socioeconomic	2000–2008	Taihu basin in Jiangsu	JPBS	Statistical Yearbook of Jiangsu
	2004–2007	Taihu basin in Zhejiang	ZPBS	Statistical Yearbook of Zhejiang
	2004–2007	Taihu basin in Anhui	APBS	Statistical Yearbook of Anhui
	2000–2008	Taihu basin in Shanghai	SBS	Statistical Yearbook of Shanghai

TBAMWSC is Taihu Basin Authority of Ministry of Water Resources; DSCCWR is Data Sharing Center of China Water Resources; JPBS, ZPBS, APBS, SBS are Jiangsu Provincial Bureau of Statistics, Zhejiang Provincial Bureau of Statistics, Anhui Provincial Bureau of Statistics and Shanghai Bureau of Statistics, respectively

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2} \sqrt{\sum (y - \bar{y})^2}} \quad (1)$$

where R is correlation coefficient, and $\bar{x}$ and $\bar{y}$ are the mean values of variables x and y, respectively.

In order to clarify the main principal components from all of the water quality parameters and identify the key driving factors for the lake eutrophication, the PCA method was performed using the statistics package SPSS. The PCA method is used as a dimension reducing technique by extracting a number of principal components accounting for the intercorrelation of the variables involved. The method is most valuable when high correlation is observed among the variables, an observation often made in lake eutrophication studies. Moreover, PCA does not impose requirements for normality and homoscedasticity and

intends to have a better interpretation of variables (Kazi et al. 2009; Noori et al. 2010). In mathematical terms, PCA and its scores involve the following five major steps: (1) start by coding the variables $X_1, X_2, \ldots, X_k$, to have means of zero and unit variance (i.e., standardize the variables to ensure they have equal weight in the further analysis); (2) calculate the covariance matrix C and correlation matrix R; (3) calculate the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ and the corresponding eigenvectors $A_1, A_2, \ldots, A_k$ based on the correlation matrix R; (4) rank order eigenvalues and corresponding eigenvectors, then discard any components accounting for a very small proportion of the total variation in data sets; and (5) develop the factor loading matrix to infer the principal parameters (Bengraine and Marhaba 2003; Fan et al. 2010). The principal components scores can be expressed as follows:
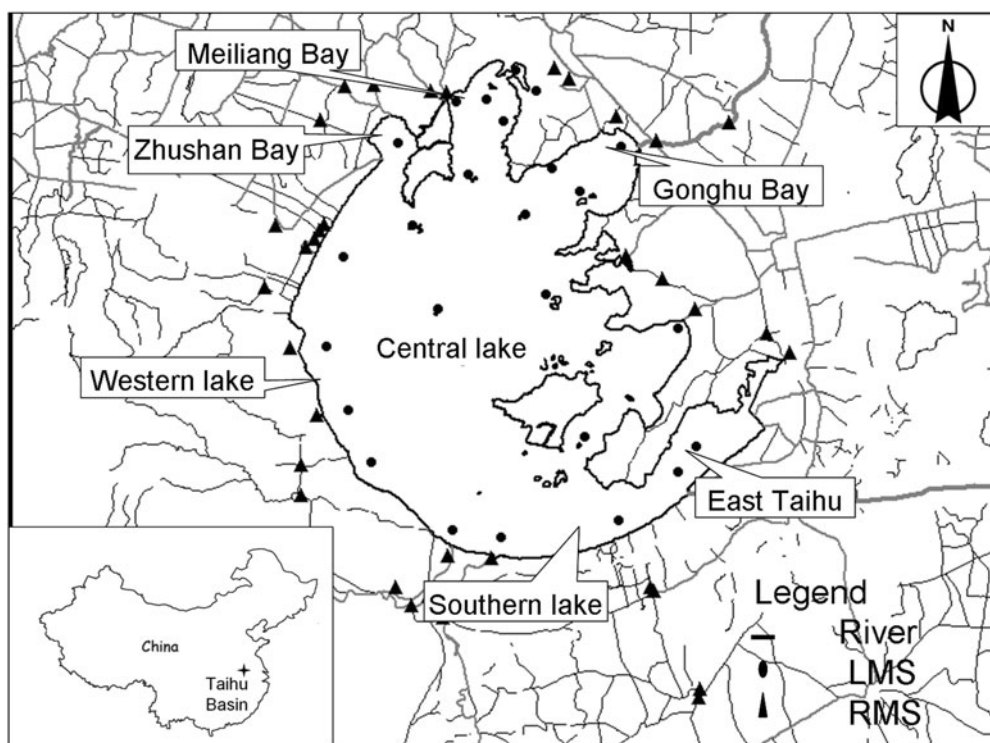
172

Int. J. Environ. Sci. Technol. (2014) 11:169–182

Fig. 1 Sampling stations in Lake Taihu and surrounding tributaries, China. LMS refers to lake monitoring stations; RMS refers to river monitoring stations

$$
\begin{bmatrix} F1_j \\ F2_j \\ F3_j \\ \cdots \\ Fi_j \end{bmatrix} = \begin{bmatrix} A_{11} \cdot X_{1j} + A_{12} \cdot X_{2j} + \cdots + A_{1k} \cdot X_{14j} \\ A_{21} \cdot X_{1j} + A_{22} \cdot X_{2j} + \cdots + A_{2k} \cdot X_{14j} \\ A_{31} \cdot X_{1j} + A_{32} \cdot X_{2j} + \cdots + A_{3k} \cdot X_{14j} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A_{i1} \cdot X_{1j} + A_{i2} \cdot X_{2j} + \cdots + A_{ik} \cdot X_{14j} \end{bmatrix} \quad (2)
$$
$$
k = 1, 2, \ldots 14
$$

where Fi_j is the component score, A_{ik} is the component loading, X is the measured value of variable, i is the component number, j is the sample number, and k is the number of variables. In this study, PCA included TN, TP, 5-day biochemical oxygen demand (BOD_5), potassium permanganate indices (COD_{Mn}), ammonia nitrogen ($NH_4^+ - N$), nitrate nitrogen ($NO_3^- - N$), nitrite nitrogen ($NO_2^- - N$), Chl-a, orthophosphate ($PO_4^{3-} - P$), DO, WT, pH, SD, and suspended solids (SS). Eigenvalues are used to determine the number of PCs for further study and describing the underlying data set (Ouyang et al. 2006).

The principal components scores were calculated from the variables' means for an independent data set of 26 in-lake stations as shown in Fig. 1. The spatial variations of each principal component score in the lake (Fig. 6) were obtained by interpolating their calculated score from 26 in-lake stations across the entire water body surface (Surfer, version 8.0, Golden Software) using an inverse distance weighting method where the weight decreased as distance increased from the interpolated points.

Results and discussion

Temporal variation of correlation between algal and water quality indices

Analysis of annual average values of the main water quality parameters TN, TP, and algae concentration (indicated by Chl-a) from 2000 to 2008 (Table 1) suggested that the water quality in Lake Taihu has deteriorated, but with differences among the parameters. TN and TP presented very similar temporal patterns. The lowest TN (2.35 mg/L) and TP (0.082 mg/L) values were observed in 2007 and 2008, respectively; alternatively, their values peaked in 2004 at 3.57 and 0.121 mg/L, respectively. The Chl-a concentrations fluctuated drastically over the period of 2000–2008. More specifically, they peaked at 48.6 μg/L in 2006, bottoming out at 12.13 μg/L in 2001. The SD value showed an overall downward trend maintaining a depth of close to 42 cm for the first 3 years and then declined gradually to 33.1 cm in 2005. During this period, the pH value was approximately 8.0 with slight fluctuations
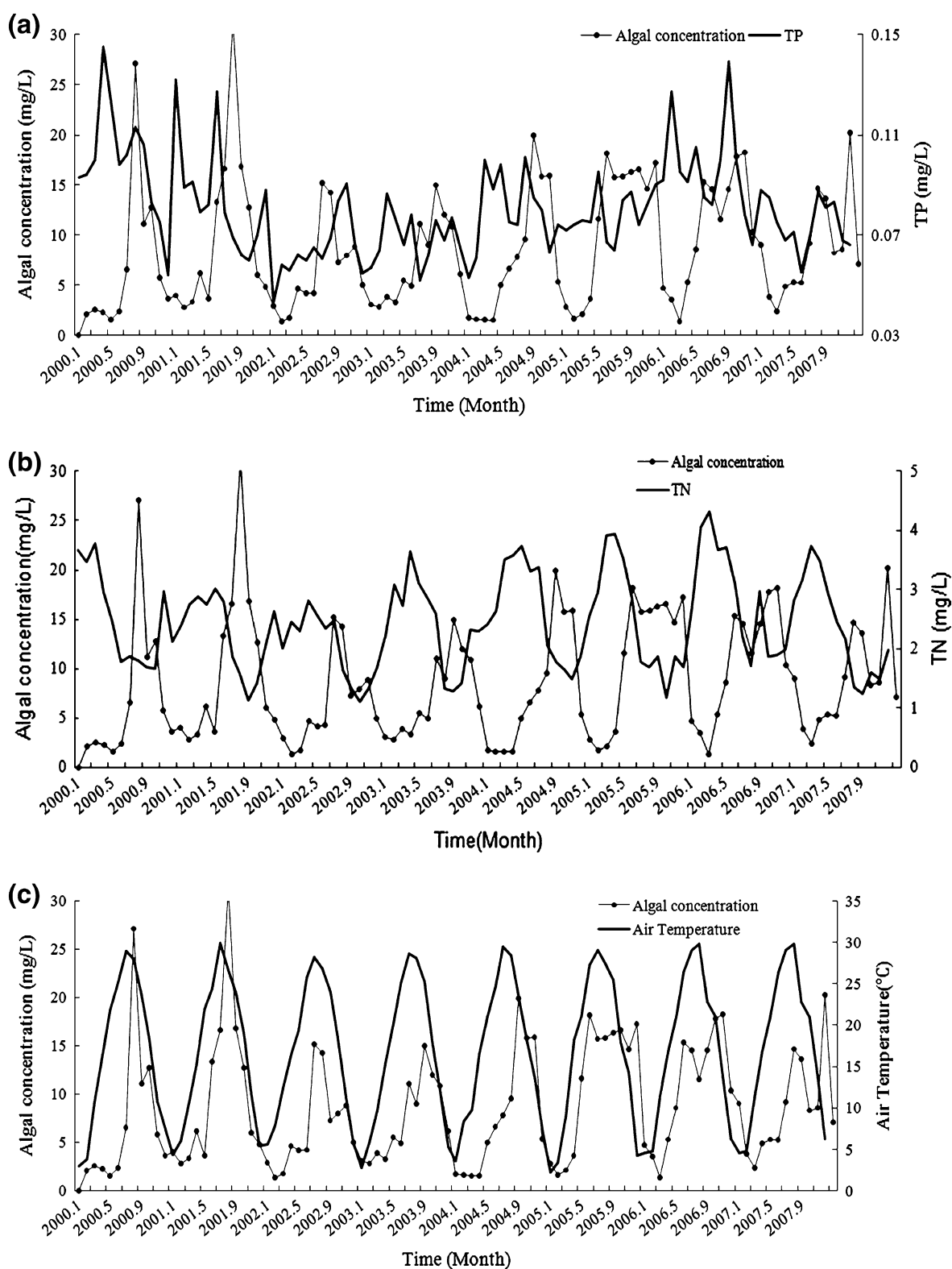
Fig. 2 Comparison of TN, TP, and air temperature with algal concentration (measured as phytoplankton biomass) monthly values in Lake Taihu from 2000 to 2007
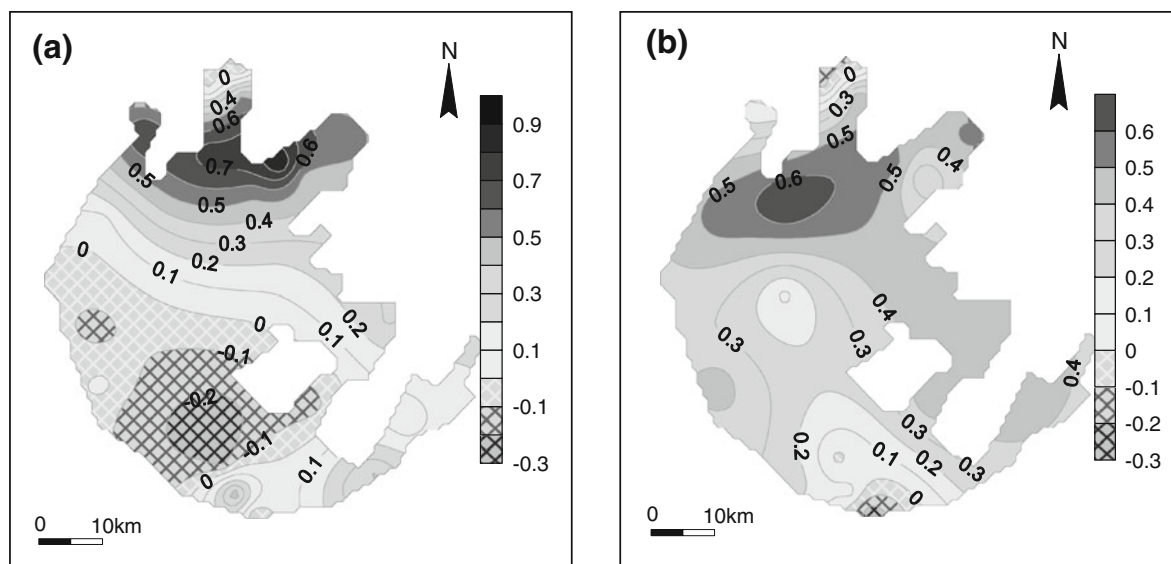
Fig. 3 Spatial distribution of correlation coefficients between total phosphorus (TP) and chlorophyll-*a* (Chl-*a*) by season in Lake Taihu, China. **a** Warm season (March to September), **b** cold season (October to February)
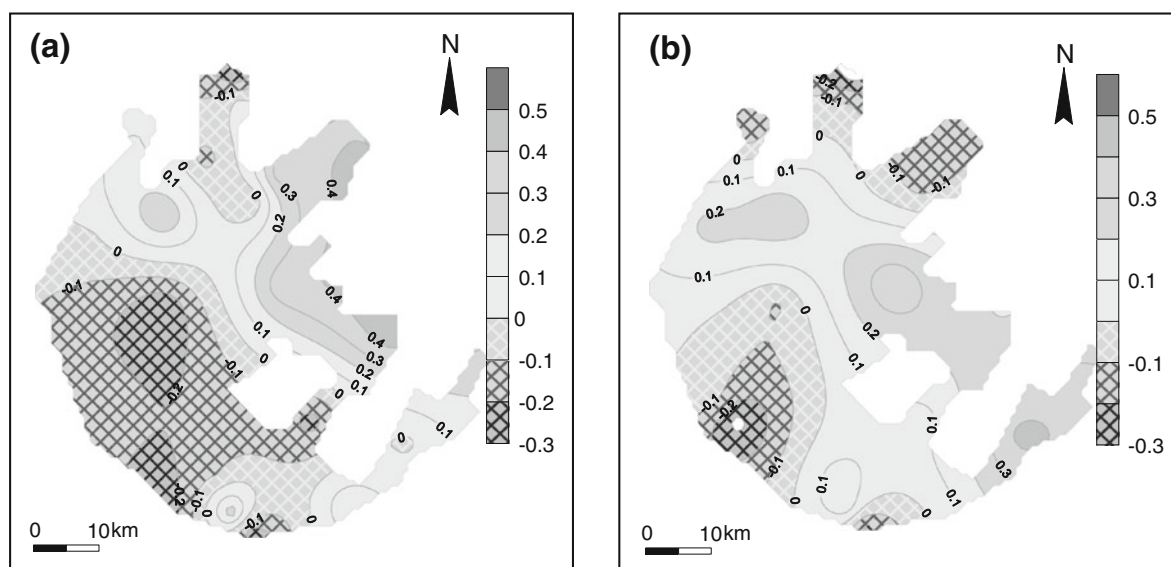

Fig. 4 Spatial distribution of correlation coefficients between total phosphorus (TN) and chlorophyll-*a* (Chl-*a*) by season in Lake Taihu, China. **a** Warm season (March to September), **b** cold season (October to February)

never exceeding 0.2. The DO value started at 9.96 mg/L in 2000, fell to 8.42 mg/L in 2002, and ended at 9.5 mg/L in 2005. The WT value peaked at 19 °C in 2002.

The monthly correlations between algae concentration (measured as phytoplankton biomass in this section due to the data availability) and air temperature, TN, TP based on the monthly observed data from 2002 to 2007 (Fig. 2) showed that both water quality parameters and

their correlations changed over time. There was a similar fluctuation between TP and algae concentration (Fig. 2a), with a relatively low level in spring and winter while at much higher level in summer. The TN and algae concentration in this work (Fig. 2b) displayed a remarkable negative correlation. Maximum TN was in the winter and early spring, and minimum TN was in the summer, contrary to the monthly algae concentration pattern. Algae concentration displayed nearly the
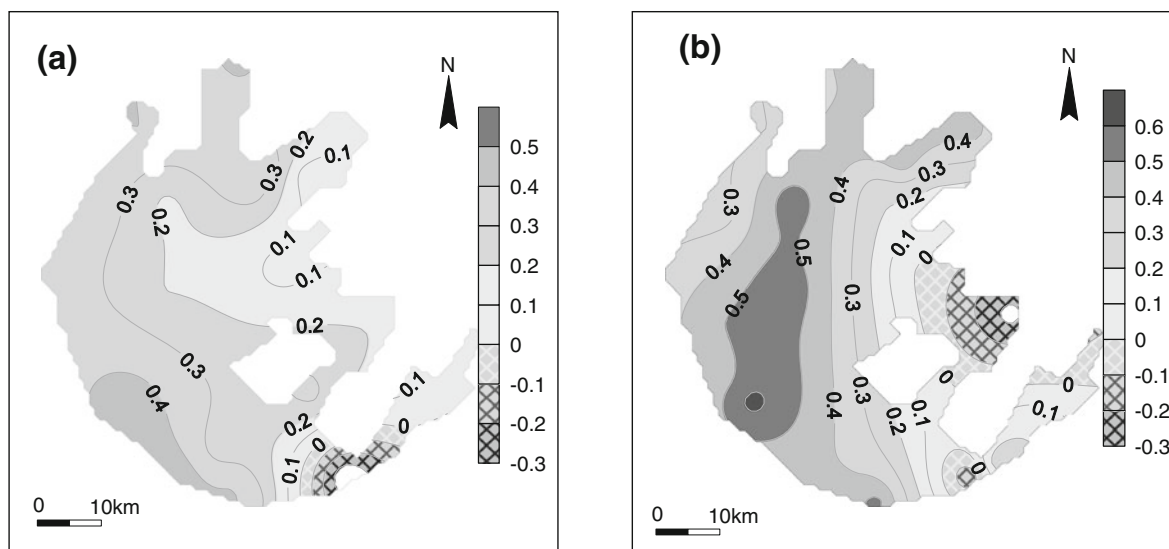
Int. J. Environ. Sci. Technol. (2014) 11:169–182

175

Fig. 5 Spatial distribution of correlation coefficients between water temperature (WT) and chlorophyll-*a* (Chl-*a*) by season in Lake Taihu, China. **a** Warm season (March to September), **b** cold season (October to February)

Table 3 Loadings of water quality variables on the first five PCs in Lake Taihu

Variables[a)]	Principal components				
	F1	F2	F3	F4	F5
NH_4^+-N	0.907	0.167		−0.162	
TN	0.816	0.417	−0.175		0.181
TP	0.781	0.238	0.115	0.402	0.108
PO_4^{3-}-P	0.737		0.235	0.106	0.193
BOD_5	0.706	0.532			
COD_{Mn}	0.588	0.559		0.263	
Chl-a	0.115	0.770			
NO_2^--N	0.342	0.704			
WT	−0.177	0.268	0.843		0.208
DO	−0.367		−0.736	0.209	
NO_3^--N		0.378	−0.540		0.475
SS		−0.104		0.837	
SD		−0.174	0.290	−0.680	
pH	−0.214				−0.894
Eigenvalue	5.04	1.99	1.29	1.15	1.01
PV (%)	36.03	14.23	9.18	8.22	7.05
CPV (%)	36.03	50.26	59.44	67.66	74.71

Results less than the absolute value of 0.1 were omitted, while values greater than 0.6 were considered correlated factors

NH_4^+-N = ammonia, TN = total nitrogen, TP = total phosphorus, PO_4^{3-}-P = orthophosphate, BOD_5 = 5-day biochemical oxygen demand, COD_{Mn} = chemical oxygen demand, Chl-*a* = chlorophyll-*a*, NO_2^--N = nitrite nitrogen, WT = water temperature, DO = dissolved oxygen, NO_3^--N = nitrate nitrogen, SS = suspended solids, SD = Secchi depth, P V, CPV = percentage or cumulated percentage of components' contributions to total variations
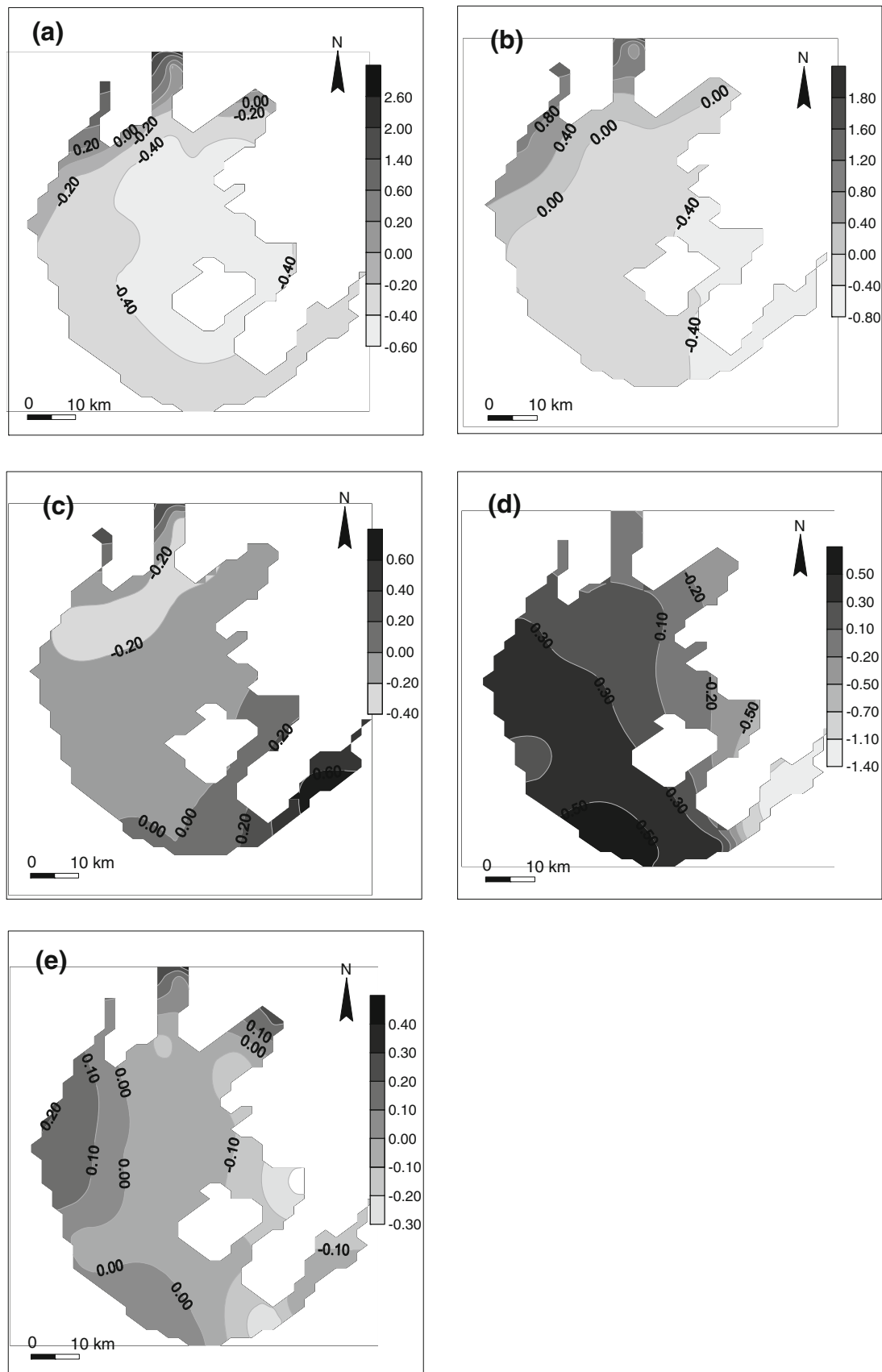
Fig. 6 Spatial distributions of PCA scores. **a** F_1, **b** F_2, **c** F_3, **d** F_4, **e** F_5

same trend as air temperature based on the monitoring data from 2002 to 2007 (Fig. 2c). It seemed that the increase in air temperature facilitated the abundance of algae.

Spatial variation of correlation between algae concentration and water quality indices

To evaluate algae and water quality indices and determine limiting factors for algae growth in different seasons and regions of the lake, the spatial distributions of correlation coefficients between algae and water quality indices (e.g., TN, TP, WT) were analyzed for the warm season (March to September) and cold season (October to February).

TP and Chl-*a* were positively correlated during the warm season with coefficient values (R) ranging from 0.5 to 0.9 in regions with higher TP concentrations such as Zhushan, Gonghu, and Meiliang bays in the northern lake (Fig. 3a). In general, correlations were stronger when the TP concentration was between 0.09 and 0.16 mg/L, but when the TP concentration increased to above 0.16 mg/L, correlations declined as seen in the inner ($R = 0.9$–0.6) and outer ($R = 0.2$–0.4) regions of Meiliang Bay (Fig. 3a). In the cold season, stronger correlations (R ranging from 0.4 to 0.6) between Chl-*a* and TP were observed in Gonghu Bay and in parts of Meiliang Bay, and weaker correlations were observed in Zhushan and northern Meiliang bays (R value 0–0.3, Fig. 3b). Generally, no strong correlations were found between Chl-*a* and TN concentrations in most lake regions ($R < 0.2$), with an exception of a small area of the East Epigeal Zone during the warm season (R ranging from 0.3 to 0.6) (Fig. 4). Additionally, the N:P ratio in Lake Taihu ranged from 50 to 100.

WT was found to be another important factor for algae growth based on its correlation with Chl-*a* concentration (Fig. 5). Positive correlations existed in the northern and western parts of the lake during both warm and cold seasons. No clear correlations between Chl-*a* and WT were observed in the eastern part of the lake. Correlations between WT and Chl-*a* during the cold season ($R = 0.3$–0.7) were higher than the warm season ($R = 0.1$–0.3) in most regions of the lake. No significant correlations were observed between Chl-*a* and other indices.

Identification of the principle components in Lake Taihu

Principal component analysis was used to identify the key principal components for eutrophication in Lake Taihu. The sorted rotated factor analyses results along with eigenvalues and percentages of variance for Lake Taihu

were carried out. The eigenvalue results less than the absolute value of 0.1 were omitted, whereas values greater than 0.6 were considered correlated factors (Table 3). Rotation of the axis defined by factor analysis produced a new set of factors, each one involving primarily a subset of the original variables with as little overlap as possible, so that the original variables are divided into groups. The first five principal components explained 74.71 % of the total variance (Table 3) for Lake Taihu. Therefore, five PCs were extracted based on the eigenvalue greater than one rule (Ruggieri et al. 2011).

The first factor accounted for 36.03 % of the total variance, which was strong positively loaded with $NH_4^+ - N$, TN, TP, $PO_4^{3-} - P$ and BOD_5 (Table 3), with strong negative loading of pH and DO. This component distinguishes the importance of anthropogenic inputs (e.g., $NH_4^+ - N$, TN, TP, $PO_4^{3-} - P$ and BOD_5) over the natural inputs (e.g., pH). The second factor, explaining 14.23 % of the total variance, was found to be strongly associated with Chl-a, which was used to describe the eutrophication status in the lake. The third factor was accounting for only 9.18 % of the total variance, with strongly positive WT loaded with negative DO. This component reflected the physical environment in the lake. The fourth factor, accounting for only 8.22 % of the total variance, was strongly loaded with SS and with negative water transparency. SS may be responsible for decreasing water transparency in the lake, explaining the water color of the lake. pH in the fifth factor loaded independently. This factor contributed 7.05 % of the total variance. According to the loadings of water quality variables in each factor, we named the first five PCs as nutrient index (F_1), eutrophication index (F_2), water temperature and dissolved oxygen index (F_3), water color index (F_4), and alkalinity acidity index (F_5).

The spatial variability of principle components scores in Lake Taihu

The spatial variations of each principal component score in the entire lake (Fig. 6) showed high spatial heterogeneity. The spatial patterns of the nutrient index (F_1) score presented a general gradient of higher values in the north and northwest and lower values toward the east and south in the lake (Fig. 6a), highlighting the heavily polluted area in Lake Taihu located in the northern and northwestern lake regions, such as Meiliang Bay, Zhushan Bay, and Gonghu Bay, where many manufacturing factories located and wastewater from domestic and industrial plants was discharged. The eastern lake regions had lower scores than

other regions, especially East Taihu, which suggested those areas had better water quality. The scores of F_2 and F_3 (Fig. 6b, c) showed similar spatial patterns to F_1, explaining that high nutrient concentrations in the lake increased the Chl-*a* concentrations and decreased the DO concentrations in the lake. The scores of water color (F_4) had higher values at the central lake and those areas where large tributaries flow in, and had lower values in eastern lake regions, especially East Taihu (Fig. 6d). The surrounding inflow tributaries carried a large number of pollutants and sediments into the lake, increasing the turbidity and decreasing the water transparency in the littoral areas. On the other hand, due to the shallowness and largeness of Lake Taihu, wind wave-induced sediment resuspension also can increase the turbidity in the open water, thus decreasing the water transparency. Moreover, aquatic plants in East Taihu can dampen the wave effect, stabilize the sediments, and attenuate sediments resuspension, which could increase the water transparency. Therefore, the lower score of F_4 appeared in East Taihu. Significant spatial variation of the F_5 score (Fig. 6e) showed that alkalinity was relatively larger in those lake regions connecting to the inflow rivers, including Meiliang Bay, Gonghu Bay, and Zhushan Bay, suggesting that the pollutants from the surrounding rivers have a great impact on the water quality of the nearby lake regions. In addition, the low value in East Taihu was a consequence of the low pH of water due to the presence of aquatic vegetation. Such observations underline the fact that the environment had a strong feedback effect on the pH value.

Spatio-temporal distribution of water quality parameters in Lake Taihu

The results suggest that water quality in the lake has been continuing to deteriorate in recent years. The population increase and economic growth have been the key driving forces to environmental pollution in Lake Taihu. It has been reported that population growth was increased by 4.8 %, and economic growth rates of Taihu basin regions were greatly accelerated from 2000 to 2007 (Jin et al. 2006). TP concentration greater than 0.015 mg/L is thought to be sufficient to cause algal blooms (Richardson et al. 2007), and TP concentrations in Lake Taihu exceeds this threshold throughout the year. TN concentrations were greater in the spring and winter than the rest of the year. This may be due to the fact that large amounts of chemical fertilizer were used in the spring as it was the beginning of the primary agricultural season and farm land constituted the majority of the land use from 1999 to 2007 (Li et al. 2008). Zhu (2009) found that dissolved nitrogen mostly

in the form of nitrate accounted for 79 % of the TN in Lake Taihu suggesting that chemical fertilizer as a non-point pollution source played an important role in the water quality degradation of the lake. Furthermore, TN in Lake Taihu was comparable to other eutrophic lakes around the world, e.g. Lake Apopka (~ 4 mg/L) in the USA and Lake Lago Trasimeno in Italy (~ 2.5 mg/L) (Havens and Elia 2009). Our data suggested that the excessive nutrients in Lake Taihu accelerated the process of eutrophication and degraded the water quality in the lake, similar to other eutrophic lakes in the world, such as Lake Norrviken in Sweden (Routh et al. 2009) and Lake Washington in the USA (Arhonditsis et al. 2003). A brief decline in nutrient concentration in Lake Taihu in 2001 was probably due to the water transfer project from Yangtze River to Lake Taihu. Fresh water from the Yangtze River was transferred into Lake Taihu via Wangyuhe River (Fig. 1) and taken out via Taipuhe River (Fig. 1). The test run of this project started in 2001. The average volume of water diverted from Yangtze River was approximately 1.9 billion m^3 annually, of which about 0.9 billion m^3 (approximately one-fifth of the lake volume ~ 4.6 billion m^3) flowed into Lake Taihu directly (Li et al. 2011a).

The algae concentrations were higher in summer and lower in winter and spring, in accordance with the change trend of air temperature. Their positive and significant correlation (Fig. 2) suggested that algae concentration was triggered mainly by temperature and light availability and maintained by the pool of nutrients, which were supplied regularly by the high nutrient concentrations in Lake Taihu. Since the average high air temperature in the summer was around 30 °C and matches the optimum growth temperature of phytoplankton (around 25–30 °C, Zhou and Yu 2004) in Taihu Basin, the elevated levels of phytoplankton biomass measured in summer in the lake can be explained as a direct consequence of the summer solar radiation peak plus the over-enrichment of nutrients (Rapala et al. 1997; Vézie et al. 2002). During the cold season, reduced water temperature and light intensity are less favorable for algae growth. As a consequence, phytoplankton biomass is diminished.

The high degree of correlation between Chl-*a* and TP during the warm months seems to indicate that Lake Taihu was P limited in parts of the lake (especially in late spring and early summer) where the TP concentrations range from 0.07 to 0.15 mg/L, such as the northern Centre Zone, southern Meiliang Bay and Gonghu Bay (Fig. 3a). This was similar to what was determined by Kotak et al. (1995), who found a positive correlation between abundance of biomass and TP. However, the degrees of correlation between Chl-*a* and

TP were very weak when P concentrations >0.16 mg/L, such as in northern Meiliang Bay. It is suggested that the limiting factor for dominant algae growth depends on seasonality and location. Additionally, it should be noted that it is difficult to draw conclusion on nutrient limitation based solely on TP or TN relationships with Chl-a in this shallow lake, since nutrient exchange between sediments and the water column is highly efficient and sediment resuspension events are frequent. These events re-inject nutrients into the water column, which will affect both availability of and limitations by nutrients. Furthermore, Xu et al. (2010) and Paerl et al. (2011a) recently found that phytoplankton is N limited in northern Meiliang Bay and parts of central lake in late summer and P limited in winter and spring. In addition, Xu et al. (2010) as well as McCarthy et al. (2009) also found free orthophosphate to actually be present in the water column during summer Microcystis cyanobactyerial blooms in Lake Taihu. This is further evidence that N, rather than P, is most likely the limiting nutrient during the critically important summer period and that the externally loaded N that is supplied during winter–spring is likely an important factor determining the magnitude and extent of summer cyanobacterial blooms in this lake. These evidences have also been found in other similar freshwater systems. For example, Elser et al. (2007) emphasized that N and P limitations are equivalent within both terrestrial and freshwater systems by comparing a large-scale meta-analysis of experimental enrichments. Lewis and Wurtsbaugh (2008) also reviewed and found that more and more evidences obtained by multiple investigators at numerous sites over the last 30 years show that N limitation in lakes is common rather than exceptional, suggesting that N is at least as likely as P to be limiting to phytoplankton growth. Therefore, as far as Lake Taihu is concerned, although P load reduction is important, N load reduction is essential for controlling the magnitude and duration of algal blooms in Taihu (Paerl 2009). It is essential to reduce both N and P inputs for long-term eutrophication and cyanobacterial bloom control in this hyper-eutrophic system (Paerl et al. 2011a, b).

Identification of the key driving factor for eutrophication in Lake Taihu by PCA

Principal component analysis has been shown to be a useful technique in the evaluation of environmental data. The main advantage of this method is that the new variables are uncorrelated among themselves and their number is reduced without significant loss of the original information (Ouyang 2005). In the present work, the 14 original variables were reduced to five key independent factors, e.g., nutrient index (F_1), eutrophication index (F_2), water temperature and dissolved oxygen index (F_3), water color index (F_4), and alkalinity acidity index (F_5). The first five components accounted for more than 74 % of the total variance of the data, well representing the physical and chemical characteristics of water quality in Lake Taihu. The contribution of each component on the total variance of the data (36.03, 14.23, 9.18, 8.22, and 7.05 % for F_1 to F_5) explained that nutrient contaminations from anthropogenic and natural inputs are the principal environmental problems of the lake, and eutrophication is another key water quality problem for the lake. The first two components could reflect the pollution status and water quality characteristics of the lake. The final three component indices could describe the physical and chemical environment in the lake, suggesting that water temperature and dissolved oxygen, water color, and alkalinity acidity were closely associated with the degradation of water quality in Lake Taihu.

Additionally, the primary water quality variables in each component and their relationships could indicate the mechanisms of eutrophication and water pollution in the lake. For example, the strong positive loading of $NH_4^+ - N$, TN, TP, $PO_4^{3-} - P$ and BOD_5 (representing anthropogenic pollution) with negative loading of pH and DO for the first component indicate that high levels of dissolved organic matter consume large amounts of oxygen. Decomposition processes then lead to the formation of organic acids, CO_2, and ammonia. Hydrolysis of these acids, dissolution of CO_2 in the water column, and/or oxidation of NH_4^+ ions under oxic conditions by the nitrification processes cause a decrease of water pH values. The latter process demonstrates the lowering of pH in the presence of excess ammonia, which originates from fertilizer industries, manufacturing area, and commercial composites (Kim et al. 2005; Singh et al. 2004). Furthermore, higher SS load and negative water transparency in the fourth factor may be correlated with resuspended sediments through churning action by wind-induced waves and currents due to the large size and shallowness of Lake Taihu. In general, PCA is a very effective method to reduce complex variables, and these new factors are easier to quantify and analyze than the original variables for water resources management.

The principal component scores could be used to assess the water quality and identify the key driving factors for the eutrophication in the lake. The spatial

heterogeneity of each principal component score in the entire lake (Fig. 6) can identify the heavily polluted lake regions and the primary causes for them. Based on the results, Meiliang Bay, Zhushan Bay, and Gonghu Bay located in the northern lake regions suffered from severe water pollution followed by northwestern and eastern lake regions having relatively lower nutrient and eutrophication levels compared to other lake areas. In general, the water quality and eutrophication decreased from north to the center and from west to east, suggesting that algal blooms events occurred only in sections of the lake rather than the entire lake. In addition, the spatial trends in nutrient and eutrophication scores strongly corresponded with observed nutrient fluxes from the adjoining rivers in densely populated areas, demonstrating the influence of watershed pollutant loads on lake water quality. For example, the main inflow tributaries with high discharge (20–50 m^3/s) were Chendonghe, Taigeyunhe, and Wujingang located in the northern and northwestern lake regions, carrying 67–75 % of the total nutrient load to Lake Taihu for each season, resulting in the high nutrient concentrations in the nearby lake regions (Li et al. 2011b). These major tributaries are surrounded by many manufacturing factories, densely populated residential areas and agricultural areas, receiving wastewater from them. It is suggested that anthropogenic inputs are of main concern for managers, especially in the northern and northwestern lake regions. Therefore, controlling pollutant loads from sewage and industrial plants in those areas should be regarded as the top priority for Chinese regulators for improving the watershed and the lake. The low nutrient levels in East Taihu were associated with the presence of substantial aquatic plants, which can absorb nutrients from the lake, and compete with algae for light and nutrients. Therefore, an additional management alternative is to establish macrophytic vegetation that serves to trap sediments and retard resuspension. The vegetation could also be harvested to remove nutrients from the system. Macrophytes were once prevalent throughout much of the lake, but have since receded due to a loss in lake clarity (Horppila and Nurminen 2005; Stefanidis and Papastergiadou 2007). The feasibility of reestablishing macrophyte beds should also be a research priority.

Conclusion

Principal component analysis was found to be an effective technique in the evaluation of environmental data in Lake Taihu without significant loss of original information. The results of correlation analysis and PCA suggest that nutrient contaminations from anthropogenic and natural inputs were the principal environmental problems of the lake, and eutrophication is another key water quality problem for the lake. Moreover, Lake Taihu is characterized with high spatial and seasonal variations in water quality and eutrophication, decreasing from north to the center and from west to east, suggesting that algal blooms events occurred only in sections of the lake rather than the entire lake. Heavily polluted areas located in northern and northwest sections of the lake, such as Meiliang Bay, Zhushan Bay, and Gonghu Bay, strongly correspond with the major adjoining rivers in densely populated and industrialized areas, demonstrating the influence of watershed pollutant loads on lake water quality. Therefore, controlling pollutant loads from sewage and industrial plants and effective land management in those areas should be regarded as the top priority for Chinese regulators for reducing the pollutant loads and improving the water quality. The results of this study also have implications for the thousands of degraded water bodies in the eastern China and the rest of the world.

References

APHA (1998) Standard methods for the examination of water and wastewater. American Public Health Association/Water Environment Federation

Arhonditsis G, Brett MT, Frodge J (2003) Environmental control and limnological impacts of a large recurrent spring bloom in Lake Washington, USA. Environ Manag 31(5):603–618

Bengraine K, Marhaba TF (2003) Using principal component analysis to monitor spatial and temporal changes in water quality. J Hazard Mater 100(1–3):179–195

Chen YW, Fan CX, Teubner K, Dokulil M (2003) Changes of nutrients and phytoplankton chlorophyll-a in a large shallow lake, Taihu, China: an 8-year investigation. Hydrobiologia 506(1–3):273–279

China Daily (CD) (2007) China Announces $14B Lake Cleanup. http://www.mwr.gov.cn/english/20071029/ 87506.asp. Accessed 29 Oct 2007

Dong XH, Bennion H, Battarbee R, Yang XD, Liu EF (2008) Tracking eutrophication in Taihu Lake using the diatom record: potential and problems. J Paleolimnol 40(1):413–429

Elser JJ, Bracken MES, Cleland EE, Gruner DS, Harpole WS, Hillebrand H, Ngai JT, Seabloom EW, Shurin JB, Smith JE (2007) Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecol Lett 10(12):1135–1142

Fan XY, Cui BS, Zhao H, Zhang ZM, Zhang HG (2010) Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environ Sci 2:1220–1234

Havens KE, Elia AC (2009) Zooplankton–phytoplankton relationships in shallow subtropical versus temperate lakes Apopka (Florida, USA) and Trasimeno (Umbria, Italy). Hydrobiologia 628:165–175

Horppila J, Nurminen L (2005) Effects of different macrophyte growth forms on sediment and P resuspension in a shallow lake. Hydrobiologia 545(1):167–175

Hu WP, Zhai SJ, Zhu ZC, Han HJ (2008) Impacts of the Yangtze River water transfer on the restoration of Lake Taihu. Ecol Eng 34(1):30–49

James RT, Havens K, Zhu GW, Qin BQ (2009) Comparative analysis of nutrients, chlorophyll and transparency in two large shallow lakes (Lake Taihu, PR China and Lake Okeechobee, USA). Hydrobiologia 627(1):211–231

Jin X, Tu Q (eds) (1990) The standard methods for observation and analysis in lake eutrophication, 2nd edn. Chinese Environmental Science Press, Beijing, p 240

Jin XL, Gao JF, Zhao GJ (2006) Impacts of 20-year socio-economic development on the trend of aquatic environment of the Taihu basin. Resour Environ Yangtze Basin 15(3):298–302

Kazi TG, Arain MB, Jamali MK, Jalbani N, Afridi HI, Sarfraz RA, Baig JA, Shah AQ (2009) Assessment of water quality of polluted lake using multivariate statistical techniques: a case study. Ecotox Environ Safe 72(2):301–309

Kim JH, Kim RH, Lee J, Cheong TJ, Yum BW, Chang HW (2005) Multivariate statistical analysis to identify the major factors governing groundwater quality in the coastal area of Kimje, South Korea. Hydrol Process 19(6):1261–1276

Kotak BG, Lam AKY, Prepas EE, Kenefick SL, Hrudey SE (1995) Variability of the hepatotoxin microcystin-LR in hypereutrophic drinking water lakes. J Phycology 31(2):248–263

Lewis WM Jr, Wurtsbaugh WA (2008) Control of lacustrine phytoplankton by nutrients: erosion of the phosphorus paradigm. Int Rev Hydrobiol 93(4–5):446–465

Li M, Sun ZH, Yang SJ, Qiao WF (2008) The analysis of land use change and driving forces of Taihu Lake rim. Geo-Inf Sci China 10(2):136–141 Chinese

Li YP, Achary K, Yu ZB (2011a) Modeling impacts of Yangtze River water transfer on water ages in Lake Taihu. China. Ecol. Eng. 37(2):325–334

Li YP, Acharya K, Stone MC, Yu ZB, Young MH, Shafer DS, Zhu JT, Gray K, Stone A, Fan LL, Tang CY, Warwick J (2011b) Spatiotemporal patterns in nutrient loads, nutrient concentrations, and algal biomass in Lake Taihu, China. Lake Reserv Manag 27(4):298–309

Mao JQ, Chen QW, Chen YC (2008) Three-dimensional eutrophication model and application to Taihu Lake, China. J Environ Sci 20(3):278–284

McCarthy MJ, James RT, Chen Y, East TL, Gardner WS (2009) Nutrient ratios and phytoplankton community structure in the large, shallow, eutrophic, subtropical Lakes Okeechobee (Florida, USA) and Taihu (China). Limnology 10(3):215–227

National Statistics Bureau (NSB) (2000) Statistics Almanac of China. China Statistics Press, Beijing, China

Noori R, Karbassi A, Sabahi MS (2010) Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste predicting. J Environ Manag 91(3):767–771

Ouyang Y (2005) Evaluation of river water quality monitoring stations by principal component analysis. Water Res 39(12):2621–2635

Ouyang Y, Nkedi-Kizza P, Wu QT, Shinde D, Huang CH (2006) Assessment of seasonal variations in surface water quality. Water Res 40(20):3800–3810

Paerl HW (2009) Controlling eutrophication along the freshwater-marine continuum: dual nutrient (N and P) reductions are essential. Estuar Coast 32(4):593–601

Paerl HW, Xu H, McCarthy MJ, Zhu GW, Qin BQ, Li YP, Gardner WS (2011a) Controlling harmful cyanobacterial blooms in a hyper-eutrophic lake (Lake Taihu, China): the need for a dual nutrient (N & P) management strategy. Water Res 45(5):1973–1983

Paerl HW, Hall NS, Calandrino ES (2011b) Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. Sci Total Environ 409(10):1739–1745

Qin BQ, Xu PZ, Wu QL, Luo LC, Zhang YL (2007) Environmental issues of Lake Taihu, China. Hydrobiologia 581(1):3–14

Rapala J, Sivonen K, Lyra C, Niemela SI (1997) Variation of microcystins, cyanobacterial hepatotoxins, in Anabaena spp. as a function of growth stimuli. Appl Environ Microbiol 63(6):2206–2212

Razmkhah H, Abrishamchi A, Torkian A (2010) Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: a case study on Jajrood River (Tehran, Iran). J Environ Manag 91(4):852–860

Richardson CJ, King RS, Qian SS, Vaithiyanathan P, Qualls RG, Stow CA (2007) Estimating ecological thresholds for phosphorus in the Everglades. Environ Sci Technol 41(23):8084–8091

Routh J, Choudhary P, Philip AM, Kumar B (2009) A sediment record of recent of nutrient loading and trophic state change in Lake Norrviken, Sweden. J Paleolimnol 42(3):325–341

Ruggieri N, Castellano M, Capello M, Maggi S, Povero P (2011) Seasonal and spatial variability of water quality parameters in the Port of Genoa, Italy, from 2000 to 2007. Mar Pollut Bull 62(2):340–349

Singh KP, 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(12):3980–3992

Stefanidis K, Papastergiadou ES (2007) Aquatic vegetation and related abiotic environment in a shallow urban lake of Greece. Belg J Bot 140(1):25–38

Unmesh CP, Sanjay KS, Prasant R, Binod BN, Dinabandhu B (2006) Application of factor and cluster analysis for characterization of river and estuarine water systems: a case study: Mahanadi River (India). J Hydrol 331(3–4):434–445

Vézie C, Rapala J, Vaitomaa J, Seitsonen J, Sivonen K (2002) Effect of nitrogen and phosphorus on growth of toxic and nontoxic Microcystis strains and on intracellular microcystin concentrations. Microb Ecol 43(4):443–454

Wang M, Shi W (2008) Satellite-observed algae blooms in China's Lake Taihu. Eos Trans AGU 89(22):201–202

Wang D, Feng X, Zhou L, Hao J, Xu X (2008) Relationship between blue algal bloom and water temperature in Lake Taihu based on MODIS. J Lake Sci 20(2):173–178 Chinese

Xu H, Yang LZ, Zhao GM, Jiao JG, Yin SX, Liu ZP (2009) Anthropogenic impact on surface water quality in Taihu Lake Region, China. Pedosphere 19(6):765–778

Xu H, Paerl HW, Qin BQ, Zhu GW, Gao G (2010) Nitrogen and phosphorus inputs control phytoplankton growth in eutrophic Lake Taihu, China. Limnol Oceanogr 55(1):420–432

Zhou YL, Yu M (2004) The occurrence damage and control of algal. Bull Biol 39(6):11–14 Chinese

Zhu GW (2008) Eutrophic status and causing factors for a large, shallow and subtropical Lake Taihu, China. J Lake Sci 20(1):21–26 Chinese

Zhu GW (2009) Spatio-temporal distribution pattern of water quality in Lake Taihu and its relation with cyanobacterial blooms. Resour Environ Yangtze Basin 18(5):439–445