Check for
updates

Trend analysis and modeling of nutrient concentrations in a preliminary eutrophic lake in China

Xinnan Tong · Xinze Wang · Zekun Li ·
Pingping Yang · Ming Zhao · Kaiqin Xu

Abstract Accurately measuring and estimating trends and variations in nutrient levels is a significant part of managing emerging eutrophic lakes in developing countries. This study developed an integrated approach containing Seasonal Trend Decomposition using Loess (STL) and a dynamic nonlinear autoregressive model with exogenous input (NARX) network to decompose and estimate the nutrient concentrations in Lake Erhai, a preliminary eutrophic lake in China. The STL decomposition results indicated that total nitrogen (TN) concentration of Lake Erhai progressively descended from 2006 to 2014, except for some agriculture area. The total phosphorus (TP) concentration showed an increasing

trend from 2006 to 2013 and then decreased in 2014, but in the area near the tourist attractions, TP increased continuously from 2011 to 2014. Seasonal variations in TN and TP indicated that the lowest water quality of Lake Erhai occurred from July to October. Based on results obtained with STL, TP was selected as the sensitive parameter, as it showed a significant deterioration trend, and the area near the tourist attractions was selected as the sensitive area. Three variables (DO, pH, and water temperature) were selected as input parameters to estimate TP using the dynamic NARX model. The NARX modeling results demonstrated that it can accurately estimate TP concentrations with low root-mean-square error (0.0071 mg/L). The study establishes a new approach to better understand trends and variations in nutrient levels and to better refine estimates by identifying more easily accessible physical parameters in a preliminary eutrophic lake.

X. Tong · X. Wang (✉)
School of Environmental Science and Engineering, Shanghai Jiao Tong University, No. 800 Dongchuan Rd, Shanghai 200240, China
e-mail: xinzewang@sjtu.edu.cn

X. Tong
e-mail: tongxinnan@sjtu.edu.cn

Z. Li · P. Yang · M. Zhao
Dali Environmental Monitoring Station, Dali 671000 Yunnan, China

K. Xu
Research Center for Material Cycles and Waste Management, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan

Introduction

Eutrophication is a worldwide environmental problem, originating from the excessive accumulation of nutrients in water bodies. The problem has drawn significant attention since the 1970s (Bouwman et al. 2002; Smith 2003); however, the issue emerged much later in China, along with the country's economic growth in the 1990s (Le et al. 2010). Things have become worse in the new

century, with sudden algal bloom break-outs in several important lakes, threatening ecological safety, and human health (Guo 2007; Goreau 2008). To face this challenge, the Chinese government proposed a national project in 2006, called the Major Science and Technology Program for Water Pollution Control and Treatment. The program was funded at a level of more than 12 billion RMB. Several rivers, lakes, and reservoirs were selected, with the goal of obtaining specific restoration experience in basins with different natural characteristics and social economic conditions.

Of the studied water bodies, Lake Erhai was selected as a representative preliminary eutrophic lake. Agricultural non-point pollution was identified as the main pollution source (Shang and Kong 2014). With the rapid increase in population and economic development, the water quality in Lake Erhai rapidly deteriorated after the 1990s. In 1996 and 2003, two severe algal blooms broke out in part of the lake. After two consecutive "Five-Year Programs" involving integrated remediation, the water quality of Lake Erhai has improved significantly in the last decade. In 2015, its water quality was graded as class II to III, according to the Water Environment Quality Standard of China (GB3838-2002). This means it meets the standard for a drinking water source (Wang et al. 2015).

However, protecting preliminary eutrophic lakes in developing countries often leads to conflicts with local economic development. The eutrophication process may accelerate during high-speed economic development. In the Lake Erhai basin, the local gross domestic product (GDP) growth rate in 2015 was 8.26%, which was faster than the average level in China (6.9%) and in the USA (2.6%). Thus, there remains a risk of future water quality deterioration. Moreover, both the positive and negative experiences in the management of the lake over the past 10 years should be carefully considered. Understanding the variations in water quality in Lake Erhai can help inform reasonable eutrophication control policies and environmentally friendly economic development strategies.

Until now, trends and variations in water quality for most lakes, especially in developing countries, have been demonstrated directly using raw monitoring data. The periodicity and stochasticity of variations in water quality are generally not accounted for, inhibiting the accurate assessment of its eutrophication state. Seasonal Trend Decomposition using Loess (STL) (Cleveland et al. 1990) has been demonstrated to be a reliable tool for analyzing long-term trends and variations (Sellinger et al. 2007; Shamsudduha et al. 2009; Wang et al. 2014a, b). STL uses a locally weighted running-line smoother to fit data, which can address nonlinear and local trends. Qian et al. (2000) and Stow et al. (2015) analyzed long-term and seasonal trends of nutrient concentrations and loads of two rivers in USA, respectively. The results indicate that STL method can exclude stochastic fluctuations in raw data and can accurately decompose water quality data into long-term and seasonal trends. With the development of different modeling tools, the modeling of variations in water quality has drawn increased attention. Of the well-developed techniques, (artificial neural network (ANN) has been shown to be a useful water quality modeling tool (Chebud et al. 2012; Khalil and Adamowski 2014). With ANN, the complicated, laborious, expensive, and time-consuming process of generating water quality data could be reduced by using more easily collected hydraulic data (Chang et al. 2015; Tomić et al. 2016). However, the previous researches always use raw monitoring data to forecast the water quality with ANN (Park et al. 2015; Xu et al. 2005). To date, no work has modeled water quality using an ANN method, which using STL composing data as input data.

This study analyzed Lake Erhai's nutrient concentration data from 2006 to 2014, the longest time span seen in the literature on this topic (Yin et al. 2011; Jiao et al. 2013). STL was applied to analyze long-term trends and seasonal variations of TN and TP concentrations of 11 stations in Lake Erhai. Sensitive water quality parameters and sensitive sampling sites were selected using STL, revealing different long-term variations and patterns compared to other indices. Finally, the most sensitive parameter in the sensitive sampling sites was modeled using an ANN method. The study established a new approach to better understand and estimate trends and variations in nutrient levels, using more easily accessible physical parameters in a preliminary eutrophic lake.

Methods

Study area and data collection

Lake Erhai is in southwest China (25° 25′–26°16′N, 99°32′–100°27′E) and is the seventh largest freshwater body in China. Its surface area covers approximately 250 km² , with an average depth of 10.2 m, and a storage

volume of 28.2×10^8 m^3 (Zhang et al. 2015). The northern part of the basin is in an agricultural region. Three primary rivers (Miju River, Yongan River, Luoshi River) flow into the lake in that area, carrying many agricultural pollutants (Wang et al. 2013). The western region of the basin consists of a narrow plain between Cang Mountain and Lake Erhai, where both agriculture and tourism are highly developed. Eighteen rivers, called the Cangshan Eighteen Creek, flow into the west part of the lake. The east side of Lake Erhai consists of sharp embankments. The southern part of the basin is downtown of Dali City; there is significant domestic and tourism pollution in this region (Fig. 1).

Figure 1 shows the distribution of sampling sites. Each water sample was a mixture of water collected at 0.5 m below the water surface and water collected from the bottom of the lake at 0.5 m above the sediment. In the study period, the water samples were collected two times a month (1st and 15th) from 2006 to 2014. Table 1 displays the preliminary statistics of the water quality data and the data processing.

The analysis scheme

This study developed a systematic analysis scheme to assess spatial-temporal water quality using artificial intelligence and statistics techniques (Fig. 2). A STL analysis was conducted to identify the sensitive water quality parameter and the sensitive area. In our research, the sensitive water quality parameter was defined as the water pollutant whose concentration value changed the most. The sensitive area was defined as the area where the sensitive water quality parameter changed the most. The nonlinear autoregressive model with exogenous inputs (NARX) dynamic ANN was used to model the concentration of the sensitive water quality parameter in the sensitive area and the input parameters were selected using correlation analysis. The main method is described as follows.

STL

A time series of monthly environmental monitoring data at a selected location is defined as the sum of two components: one high-frequency seasonal component and one low-frequency long-term component (or trend component). Each individual observation is decomposed as

$$Y_{year,month} = T_{year,month} + S_{year,month} + R_{year,month} \qquad (1)$$

In this expression, $Y_{year,month}$ is the observed value for a given year and month; $T_{year,month}$ is the frequency of variation in the data, together with non-stationary, long-term changes in the levels. $S_{year,month}$ is defined the variation in the data at or near the seasonal frequency. The term $R_{year,month}$ is the remaining variation in the data beyond the seasonal and trend components. The STL model applies one continuous loses line for the smoothed, long-term component, and 12–month specific loses lines for the seasonal component. The method of choosing smoothing parameters can be found in the reference (Cleveland et al. 1990). We chose window widths of 12 months and 48 months, respectively, in order to represent the window widths of the seasonal and long-term components. Periodic time windows were selected to reveal trends. A detailed description of the operational processes and parameter setting are provided in Cleveland et al. (1990).

Correlation analysis

Selecting input parameters for the water quality estimation model is an important part of the process. A larger number of input parameters can improve prediction accuracy; however, excess input data can lead to more redundancy in computing, hindering the simulation (Chang et al. 2015). The correlation analysis can reveal the internal correlation relationship between water quality parameters. Because of this function, the correlation analysis can be used to select the input parameters for ANN model. This study used a correlation analysis (Pearson correlation test) to select the input factors. The trend data (decomposed by STL) was used for the correlation analysis instead of the observation data, which is different from other research.

NARX dynamic ANN

The NARX dynamic ANN is one of the most widely used dynamic ANN. It retains data from a former operation and feeds it into the next data operation, giving it a dynamic feature, while also keeping complete information.

The NARX network contains input layers, hidden layers, and output layers. Recurring connections from the output may delay several unit times to form new inputs. The mathematical structure of the
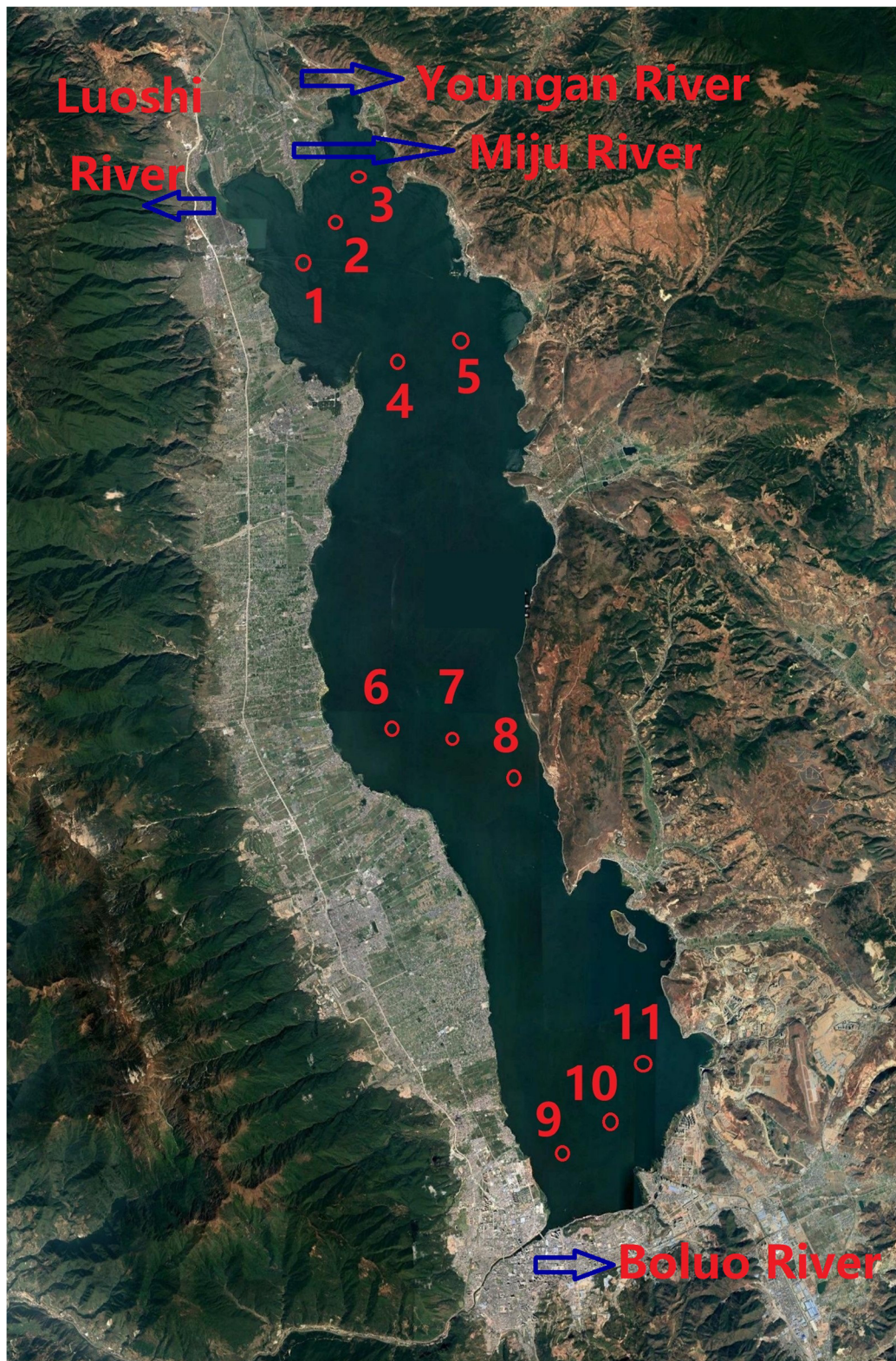
Fig. 1 Map of Lake Erhai and the location of inflow rivers (Shang et al. 2012)

Table 1 Basic statistics of water quality data obtained from 11 monitoring stations along Lake Erhai (January 2006–December 2014)

Parameters	Unit	Min	Max	Mean	SD[a]
Temperature	°C	8.0	26.0	17.3	4.5
pH	–	8.7	8.6	8.7	0.1
DO	mg/L	7.43	7.28	6.60	0.86
COD	mg/L	3.40	3.24	2.80	0.43
TN	mg/L	0.455	0.741	0.533	0.135
TP	mg/L	0.018	0.032	0.023	0.008
NH$_3$-N	mg/L	0.149	0.363	0.140	0.101

NARX network is shown in the following Eq. 1 (Chang et al. 2015):

$$z(t) = f[z(t-1), ..., z(t-dz); U(t)]$$

In this expression, U(t) and z(t) denote the input vector and output value of the model at a discrete time step t, respectively. The term f(·) indicates the nonlinear activation function that needs to be approximated by a learning algorithm.

The NARX network can be trained by two modes. The first mode is the series-parallel (SP) mode, where the output's regressor in the input layer is formed only by actual values of the system's output, d(t):

$$z(t) = f[d(t-1), ..., z(t-dz); U(t)]$$

The second mode is the parallel (P) mode, where estimated outputs are fed back into the

Fig. 2 Study flow of the proposed systematic analysis scheme

output's regressor in the input layer and can be mathematically represented as Eq. (1). The NARX network can be trained in the SP mode to construct the relationship between actual and estimated values of the target variable. Then the constructed NARX network in the P mode is applied to the unrecorded period for improving estimation performance with the recurrent information (the estimated values derived from the model).

The performance of the NARX dynamic ANN model is evaluated using common measures of goodness of fit: RMSE (root-mean-square error) and R (the correlation coefficient between the estimated value and observed value). The minimization of RMSE and maximization of R between experimental and modeled values was the target of appropriate estimation model, shown as follows (Young et al. 2015):

$$R = \frac{(1/N)\sum_{i=1}^{N}\left[(Y_m)_i - \overline{Y_m}\right]\left[(Y_0)_i - \overline{Y_0}\right]}{\sqrt{(1/N)\sum_{i=1}^{N}\left[(Y_m)_i - \overline{Y_m}\right]^2} \times \sqrt{(1/N)\sum_{i=1}^{N}\left[(Y_0)_i - \overline{Y_0}\right]^2}}$$

$$RMSE = \sqrt{(1/N)\sum_{i=1}^{N}\left[(Y_m)_i - (Y_0)_i\right]^2}$$

$$\overline{Y_m} = \left(\sum_{i=1}^{N}(Y_m)_i\right)/N$$

$$\overline{Y_0} = \left(\sum_{i=1}^{N}(Y_0)_i\right)/N$$

In this expression, Y_m is the estimated value, Y_0 is the observed value, and N is the number of data points.

Statistical analysis

For nitrogen long-term data, significant differences among stations and years were evaluated through a parametric one-way analysis of variance (ANOVA). Significant ($p < 0.05$) differences were detected by a multiple Tukey comparison test. Statistical analyses were performed using the commercial software SPSS version 19.0 (SPSS, Inc., Chicago, IL, USA) (Fig. 3).
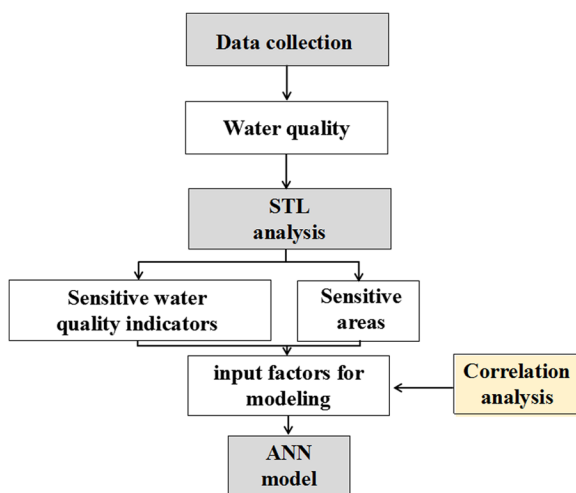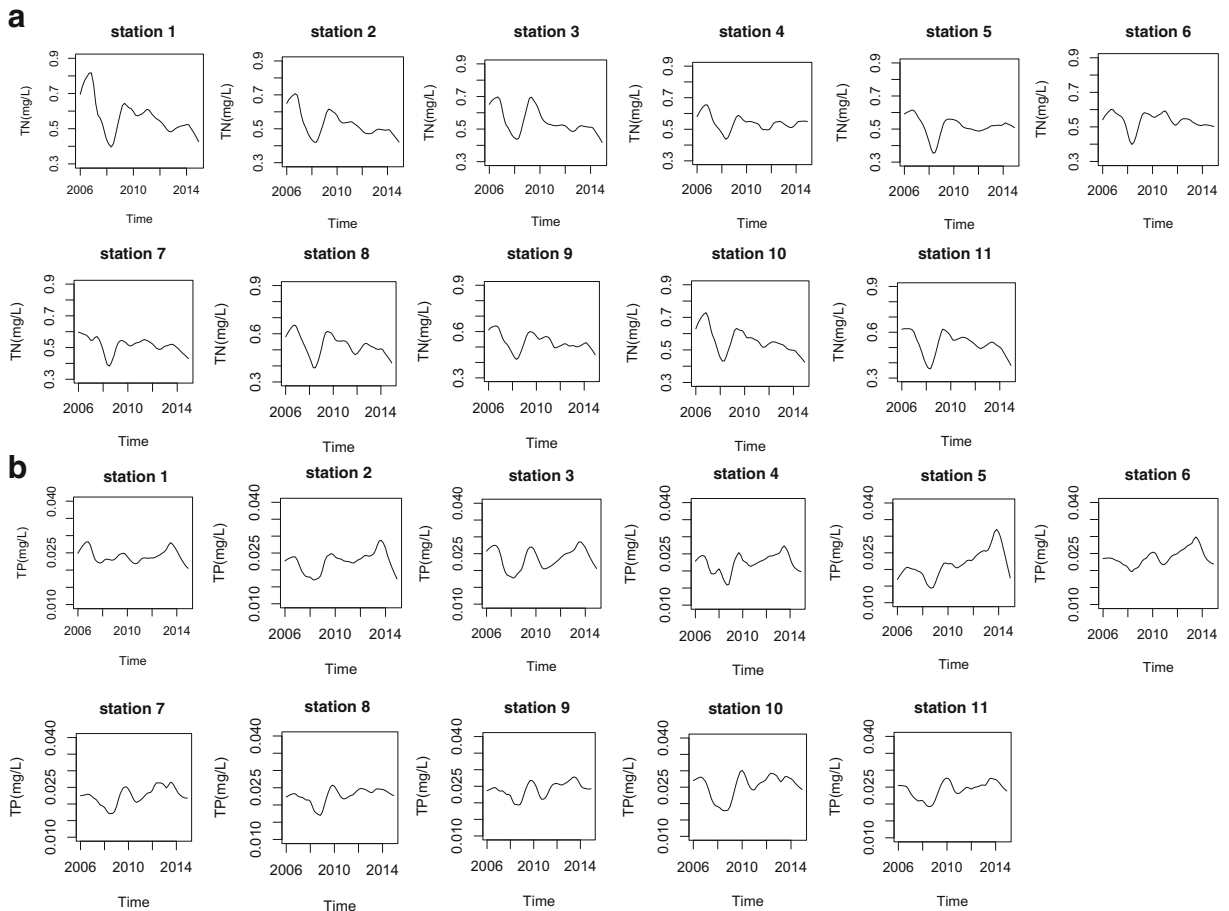
Fig. 3 Spatial patterns of the long-term trends in TN and TP concentrations (in mg/L). Each column represents one sampling station

Results and discussion

Long-term trends in nutrient concentrations at the sampling sites

Long-term trends in total nitrogen (TN) concentrations appear almost the same across all sites. Between 2006 and 2014, TN concentrations trended progressively downward, with the exception of a sharp fluctuation in 2007–2008. The decreasing tendency of the TN concentration reflects the impacts of governance during these years. Starting in 2003, the local government initiated a water environment remediation project for three main inflow rivers in the northern basin (Miju River, Yongan River, Luoshi River). At the same time, estuarine wetlands for these three inflow rivers and a 3.3-km^2 lakeside zone were constructed. These measures effectively reduced the TN load from the inflow rivers and agriculture production near the lake. Several

decentralized waste treatment systems were also built to control rural domestic sewage from nearby villages. According to a survey, by 2010, there were 45 soil purification tanks in the basin (Zhang et al. 2013).

In 2007–2008, the TN concentrations at all sampling sites reached their minimum values. The values then increased to a new higher value that remained lower than levels before 2007. This outcome was likely due to the larger rainfall amount in 2007 (931 mm), which was higher than the multiyear average value (673 mm). The sudden change in river flow may have led to significant improvements in river water quality and impacts on ecosystems (Chang et al. 2015). This may have led to a decline in the TN concentration.

For the spatial comparison, TN concentrations at site 1 were significantly higher compared to other sites (Table S1; $P < 0.05$). This site was located near the northern lakeshore of Lake Erhai, a traditional agricultural and livestock rising area. This area includes 58% of

the total farmland and 70% of the total livestock in the basin (Zhang et al. 2014; Lu et al. 2015). As a result, the northern region contributed the most agriculture non-point pollution produced among the whole basin. It was estimated that the three main inflow rivers (Miju River, Yongan River, Luoshi River) in the northern basin created approximately 47% of the TN load into the lake (Yan et al. 2005).

Because of the many restoration measures listed above, the TN concentration at the north stations (stations 1, 2, and 3) declined significantly between 2009 and 2014 ($p < 0.05$; Table S2). While TN concentrations remained almost the same across most of the sites, the trends in the TN concentrations differed at sites 4 and 5. Unlike other sites, the TN concentrations at sites 4 and 5 remained stable, rather than declining after 2010. These sites were near the northwest lakeshore of Lake Erhai, near a large farming village (Xizhou Town). In this area, rainfall, farmland irrigation water, and rural domestic wastewater flow into the lake through rivers or overland flow. This creates a large nutrient load. Our results indicated that while much has been done, the measures taken to prevent agriculture non-point pollution in the north-western region have not been sufficient. Stricter agricultural non-point pollution control policies should be implemented in this area in the future.

The TP concentrations at sites 1–5, which are in an agriculture and livestock region, progressively increased from 2006 to 2013. Concentrations then decreased in 2014. Studies showed that in the Lake Erhai basin, increased livestock and rural wastewater were the main sources of TP, creating 60% and 19% of the total TP load, respectively (Shang and Kong 2014; Wang et al. 2014a, b). From 2011 to 2014, livestock breeding in Lake Erhai basin developed more rapidly than previously. In 2014, meat and milk production increased 9.2% and 22%, respectively, compared to 2011 levels. However, livestock wastewater and animal dung were not well treated. Livestock wastewater was discharged into the lake without treatment, and the animal dung was habitually stacked on the roadside. Nutrient elements were then brought into the lake through rivers or overland runoff.

The inadequate governance of livestock pollution appears, therefore, to have been the dominant factor driving the increased TP before 2014. However, since 2013, there has been an increased exploration of livestock waste collection and recycling. In 2014, three livestock waste collection points were put into use,

leading to the collection of 120,000 t of livestock waste. As a result, TP concentrations have declined since then; however, the TP concentrations at sites 9, 10, and 11 were higher than other sites (Table S1). This may be related to the geographical position of these sites. Sites 9, 10, and 11 are located in the southern part of Lake Erhai, which adjoins downtown Dali City. The tail water from the Xiaguan municipal sewage plant is the main source of the TP load in this region. The modeling indicates that in downtown Dali City, the treatment ratio of urban sewage was 93.65% and the TP load of the tail water was 48 t/a (Bai et al. 2015). This may help explain the higher TP concentration in this area.

Unlike other sites, the TP concentration at site 6 rose consistently between 2006 and 2014. Site 6 was located near the west lakeshore of Lake Erhai, next to Dali Ancient Town, a famous scenic spot. In recent years, as tourism has developed, the number of tourists in Dali has increased rapidly (from 6 million in 2010 to 9 million in 2014). This increase of tourists may have led to the increased nitrogen and phosphorus in the water (Thebault and Qotbi 1999; Karydis and Kitsiou 2012). Thus, the government should promote ecological tourism and strengthen wastewater treatment in this area.

Seasonal patterns of nutrient concentrations at the monitoring stations

The former study showed that for spatial distribution, the west part of Lake Erhai contains most TN and TP contents, in percentages of 51% and 60%, respectively (Bai et al. 2015). This is because of the higher population density, more developed agriculture, and increases in the livestock industry. In this study, the TN long-term concentrations of stations in west part of Lake Erhai (stations 1, 4, 6) were higher than those stations in the center and east part of Lake Erhai (Table S1). Therefore, we chose the stations 1, 4, 6, and 9 to do seasonal pattern analysis. Thus, the addition of long-term trend data and seasonal data from these sites were selected to analyze the seasonal patterns.

In our research, TN concentration trends followed a regular seasonal pattern, with the highest value occurring from July to October (Fig. 4; Table S3). This seasonal pattern may relate to the rainfall pattern and crop rotation system in Lake Erhai basin. The climate of the Lake Erhai basin is consistent with a typical low-latitude plateau and is a subtropical and southwest monsoon type (Tang et al. 2012). Seasons alternate
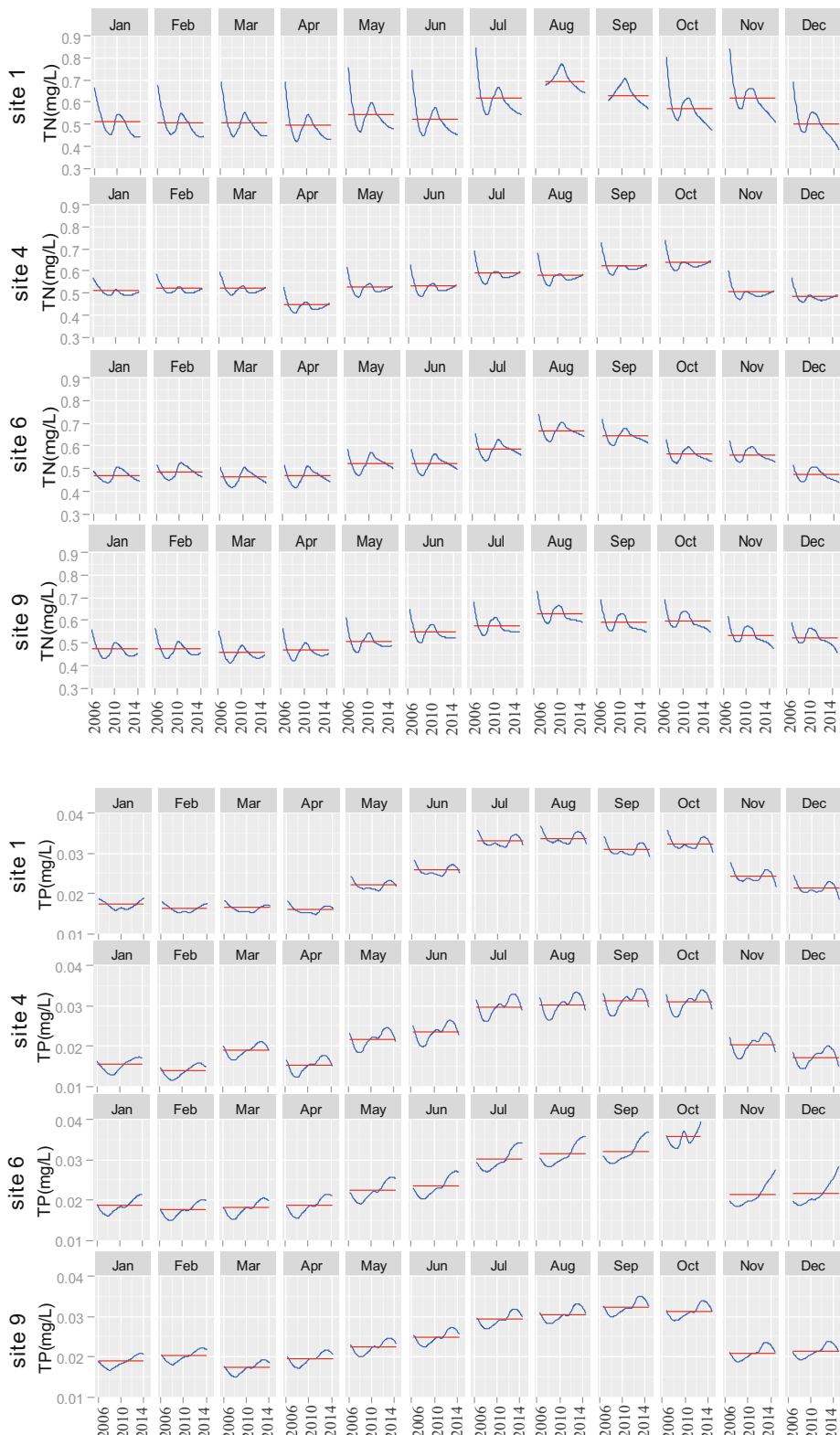
Fig. 4 Seasonal patterns in the long-term nutrient concentration (in mg/L) trends at sites 1, 4, 6, and 9 for TN and TP. Each column represents the long-term trend in a specific month (calculated as the sum of first two terms in Eq. 1 for the month). The red line represents the long-term mean for each month

significantly between the dry and rainy seasons. Historical data indicate that 85–96% of the rainfall is concentrated between May and October, with a monthly peak of 217 mm in August (Dearing et al. 2008).

A crop rotation system has been established in the study region for more than 100 years. Spring crops, including rice, maize, and tobacco, are planted from June to September. Winter crops, including broad beans, garlic, barley, and rape, are planted from September to April of the following year (Shang et al. 2012). Therefore, recently applied fertilizer is scoured by abundant rainfall runoff (Chen et al. 2016), which finally converges into Lake Erhai. This may be the dominant factor driving the increasing trend in TN concentration from July to September. For winter crops, fertilizer is applied in October. Garlic needs more fertilizer than other crops. More specifically, the amount of nitrogen fertilizer applied to garlic is four to five times the amount it takes up (Tang et al. 2012). Therefore, although October is the tail of the rainy season, the large amount of fertilizer may induce a substantial TN load into Lake Erhai.

The TP concentration experiences a similar seasonal pattern as TN. The TP concentration was significantly higher from July to October, and lower in other months (Fig. 4; Table S3). The seasonal variations in TP concentrations were also due to rainfall patterns in the basin. However, unlike TN, the TP concentrations remained high throughout the rainy season. Although the rainfall leads to both scouring and dilution, an additional reason explains this outcome. The P releasing from sediments have become an important source of P loading in lakes (Boström et al. 1988; Caraco et al. 1989). For example, Liu et al. reported that the release flux of TP at the sediment-water interface of Lake Erhai is estimated at 114.2 t/a (2015). Jin et al. indicated that the low DO value of overlying water may cause anaerobic P release at the sediment-water interface of the lake (2006). In this

study, Fig. 5 shows that oxygen concentrations in Erhai Lake were higher than 5.5 mg/L, indicating aerobic conditions, but from July to October, the DO concentrations were lower than other months. Considering that the water sample was a mixture of surface water and overlying water, there may be an anaerobic condition in sediment-water interface, which may be the reason for higher P in these months. The P release may also be affected by stream and wind (Klein 2008). Therefore, during the rainy season, disturbances from rainfall, river discharge, and DO may have led to phosphorus releases and the high TP concentration from July to October (Pant and Reddy 2001; Kim et al. 2003).

The TP concentration at site 6 increased persistently between July and October, reaching its highest value in October. This was different from other sites. This may be due to the sudden increase in tourist population between October 1 and 7, the National Day Holiday of China. For example, during the National Day Holiday in 2014, Dali received approximately 450,000 tourists, an average of 60,000 tourists each day. This was three times the mean value for that year. These high numbers of tourists may produce significant levels of domestic wastewater and solid waste, resulting in a higher TP concentration in October. This result indicated that, despite agriculture non-point pollution, pollution caused by tourism has become an emerging pollution source in the Lake Erhai basin.

The relationship between TN and TP concentration and rainfall can highlight nitrogen and phosphorus sources (Duan and Zhang 1999). Non-point source pollution is significantly affected by rainfall; however, point source pollution is relatively stable. The seasonal TN and TP concentration patterns in site 9 held at steadier levels than at other stations. This may because pollutants in the southern part of the Lake Erhai basin mainly come from Boluo River and tailwater from
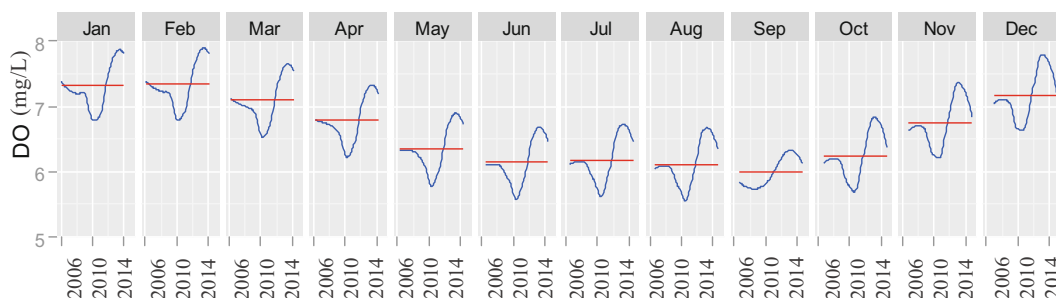

Fig. 5 Seasonal patterns for DO. The red line represents the long-term mean for each month
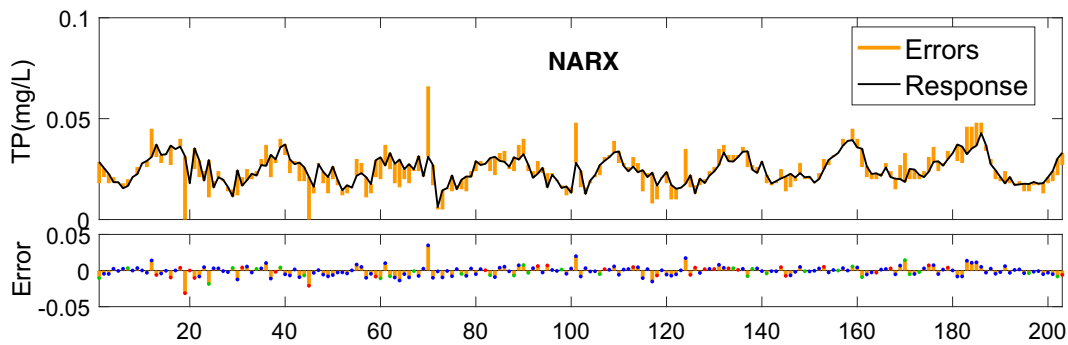
Fig. 6 Performance comparison of regional TP concentrations

WWTPs. The nitrogen and phosphorus concentrations are simultaneously impacted by non-point and point source pollution, making the effect of rainfall less clear.

Modeling of the long-term TP trends at site 6 using an NARX-ANN approach

Long-term TN trends at all sites and long-term TP trends at most sites decreased during the study period; however, TP at site 6 showed a significant increasing trend across these years. Therefore, site 6 and its TP levels were selected as the sensitive area and sensitive water quality parameter, respectively. Afterward, a correlation analysis was conducted before the ANN analysis to select the input parameters.

We selected the input indicators by examining the correlation between TP and other water quality parameters at site 6. Trend data (decomposed by STL) and observed data were used for the correlation analysis. In the observed data analysis, TP was not significantly correlated with other parameters. However, when conducting the trend data analysis, TP was significantly correlated with COD, DO, pH, and water temperature. This may because trend-based data can reflect variation tendencies more accurately than observed data. As such, a correlation analysis based on trend data can more clearly reveal the correlation between parameters (Chang et al. 2015). Although TP was correlated with COD, DO, pH, and water temperature, based on data availability, we selected the more easily accessible physical parameters as input variables: DO, pH, and water temperature (Table 2).

Using the previous three selected factors as input variables and the TP concentrations as the output variable, this study adopted NARX-ANN to estimate the TP concentration at site 6. The neural network was trained using 70% of the data (2006–2014), and the validation and testing was conducted using 15% of the data. Table 3 presents the model performance in the testing stages. The RMSE was sufficiently low, and the R value was sufficiently high, to demonstrate that the model performed wells. Figure 6 shows the modeling time series of TP concentrations associated with NARX-ANN. The results demonstrated that (a) this model fit the time series of TP well and (b) the model captures the peak value during the rainiest season (June–September). This

Table 2 Correlation of water quality parameter pairs at site 6 (January 2006–December 2014)

	Parameter pair	Correlation coefficient
Trend data	TP, TN	0.073
	TP, NH$_3$-N	− 1.171
	TP, COD	0.399**
	TP, DO	0.570**
	TP, pH	− 0.716**
	TP, water temperature	− 0.206*
Observed data	TP, TN	− 0.079
	TP, NH$_3$-N	− 0.133
	TP, COD	− 0.032
	TP, DO	− 0.085
	TP, pH	− 0.177
	TP, water temperature	0.090

*$\theta < 0.05$, significance level Sig(θ) (significant correlation); **$\theta < 0.01$, significance level Sig(θ) (highly significant correlation)

Table 3 Performance of TP concentration modeling at site 6 in the testing stages

Model	Node number	RMSE	R
NARX	3	0.0071	0.716

is critical for water pollution control. In summary, the NARX network coupled with the correlation analysis can delineate TP concentrations using COD, DO, and pH.

Conclusion

This study developed an integrated approach containing STL and a dynamic nonlinear autoregressive model with exogenous input (NARX) network to decompose and estimate the nutrient concentrations in Lake Erhai, a preliminary eutrophic lake in China. The outcome shows the positive effects of integrated remediation over the past 10 years, as well as deficiencies. From 2006 to 2014, TN was successfully controlled, as its concentration progressively descended in most part of Lake Erhai. In the area near the tourist attractions, TP increased continuously from 2011 to 2014, which means that the government should not only strengthen the agriculture wastewater treatment but also promote ecological tourism. TP was successfully modeled to assess variations in sensitive parameters at sensitive sites, facilitating the modeling of nutrient concentrations using more easily obtained physical parameters.

References

Bai, X. Y., Hu, C. Z., & Pang, Y. (2015). Pollution load, distribution and characteristics of low-polluted water in Lake Erhai watershed. *Journal of Lake Science, 27*(2), 200–207.

Bouwman, A. F., Van Vuuren, D. P., Derwent, R. G., & Posch, M. (2002). A global analysis of acidification and eutrophication of terrestrial ecosystems. *Water, Air, and Soil Pollution, 141*(1–4), 349–382.

Boström, B., Andersen, J. M., Fleischer, S., & Jansson, M. (1988). Exchange of phosphorus across the sediment-water interface. In *Phosphorus in freshwater ecosystems* (pp. 229–244). Dordrecht: Springer.

Caraco, N. F., Cole, J. J., & Likens, G. E. (1989). Evidence for sulphate-controlled phosphorus release from sediments of aquatic systems. *Nature, 341*(6240), 316–318.

Chang, F. J., Tsai, Y. H., Chen, P. A., Coynel, A., & Vachaud, G. (2015). Modeling water quality in an urban river using hydrological factors data driven approaches. *Journal of Environmental Management, 151*, 87–96.

Chebud, Y., Naja, G. M., Rivero, R. G., & Melesse, A. M. (2012). Water quality monitoring using remote sensing and an artificial neural network. *Water, Air, & Soil Pollution, 223*(8), 4875–4887.

Chen, C., Gao, M., Xie, D., & Ni, J. (2016). Spatial and temporal variations in non-point source losses of nitrogen and phosphorus in a small agricultural catchment in the Three Gorges region. *Environmental Monitoring and Assessment, 188*(4), 257.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics, 6*(1), 3–73.

Dearing, J. A., Jones, R. T., Shen, J., Yang, X., Boyle, J. F., Foster, G. C., Crook, D. S., & Elvin, M. J. D. (2008). Using multiple archives to understand past and present climate human environment interactions: the Lake Erhai catchment, Yunnan Province, China. *Journal of Paleolimnology, 40*(1), 3–31.

Duan, S., & Zhang, S. (1999). The variations of nitrogen and phosphorus concentrations in the monitoring stations of the three major rivers in China. *Scientia Geographica Sinica, 19*(5), 411–416.

Goreau, T. J. (2008). Fighting algae in Kaneohe bay. *Science, 319*(5860), 157–157.

Guo, L. (2007). Doing battle with the green monster of Taihu Lake. *Science, 317*(5842), 1166–1166.

Jiao, L. X., Zhao, H. C., Wang, S. R., & Gao, D. T. (2013). Characteristics of temporal and spatial distribution of phosphorus loading in Erhai Lake in 2010. *Research of Environmental Sciences, 26*(5), 534–539.

Jin, X., Jiang, X., Yao, Y., Li, L., & Wu, F. C. (2006). Effects of light and oxygen on the uptake and distribution of phosphorus at the sediment–water interface. *Science of the Total Environment, 357*(1), 231–236.

Karydis, M., & Kitsiou, D. (2012). Eutrophication and environmental policy in the Mediterranean Sea: a review. *Environmental Monitoring and Assessment, 184*(8), 4931–4984.

Khalil, B., & Adamowski, J. (2014). Comparison of OLS, ANN, KTRL, KTRL2, RLOC, and MOVE as record-extension techniques for water quality variables. *Water, Air, & Soil Pollution, 225*(6), 1–19.

Kim, L. H., Choi, E., & Stenstrom, M. K. (2003). Sediment characteristics, phosphorus types and phosphorus release rates between river and lake sediments. *Chemosphere, 50*(1), 53–61.

Klein, J. D. (2008). From ditch to delta, nutrient retention in running waters. Dissertation, Wageningen University.

Le, C., Zha, Y., Li, Y., Sun, D., Lu, H., & Yin, B. (2010). Eutrophication of lake waters in China: cost, causes, and control. *Environmental Management, 45*(4), 662–668.

Liu, W., Wang, S., Zhang, L., Ni, Z., Zhao, H., & Jiao, L. (2015). Phosphorus release characteristics of sediments in Erhai Lake and their impact on water quality. *Environmental Earth Sciences, 74*(5), 3753–3766.

Lu, S. Y., Zhang, W. T., Xing, Y., Qu, J. T., Li, K., Zhang, Q., & Xue, W. (2015). Spatial distribution of water quality parameters of rivers around Erhai Lake during the dry and rainy seasons. *Environmental Earth Sciences, 74*(11), 7423–7430.

Pant, H. K., & Reddy, K. R. (2001). Phosphorus sorption characteristics of estuarine sediments under different redox conditions. *Journal of Environmental Quality, 30*(4), 1474–1480.

Park, Y., Chu, K. H., Park, J., Cha, S. M., King, J. H., et al. (2015). Development of early-warning protocol for predicting chlorophyll-*a* concentration using machine learning models in freshwater and estuarine reservoirs, Korea. *Science of the Total Environment, 502*(2015), 31–41.

Qian, S. S., Borsuk, M. E., & Stow, C. A. (2000). Seasonal and long-term nutrient trend decomposition along a spatial gradient in the neuse river watershed. *Environmental Science and Technology, 34*(21), 4474–4482.

Sellinger, C. E., Stow, C. A., Lamon, E. C., & Qian, S. S. (2007). Recent water level declines in the Lake Michigan−Huron system. *Environmental Science & Technology, 42*(2), 367–373.

Shamsudduha, M., Chandler, R., Taylor, R., & Ahmed, K. (2009). Recent trends in groundwater levels in a highly seasonal hydrological system: the Ganges-Brahmaputra-Meghna Delta. *Hydrology and Earth System Sciences, 13*(12), 2373–2385.

Shang, X., & Kong, H. N. (2014). Countrtmeasures for the non-point pollution control in the watershed of Lake Erhai, China. Our lakes: from the present towards a future perspective (pp. 91-101). Kyoto: Shoukadoh.

Shang, X., Wang, X., Zhang, D., Chen, W., Chen, X., & Kong, H. (2012). An improved SWAT-based computational framework for identifying critical source areas for agricultural pollution at the lake basin scale. *Ecological Modelling, 226*, 1–10.

Smith, V. H. (2003). Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research, 10*(2), 126–139.

Stow, C. A., Cha, Y., Johnson, L. T., Confesor, R., & Richards, R. P. (2015). Long-term and seasonal trend decomposition of Maumee river nutrient inputs to western Lake Erie. *Environmental Science and Technology, 49*(6), 3392–3400.

Tang, Q. X., Ren, T. Z., Wilko, S., Liu, H. B., Lei, B. K., Lin, T., et al. (2012). Study on environmental risk and economic benefits of rotation systems in farmland of Erhai Lake Basin. *Journal of Integrative Agriculture, 11*(6), 1038–1047.

Thebault, J. M., & Qotbi, A. (1999). A model of phytoplankton development in the Lot River (France). Simulations of scenarios. *Water Research, 33*(4), 1065–1079.

Tomić, A. N. Š., Antanasijević, D. Z., Ristić, M. Đ., Perić-Grujić, A. A., & Pocajt, V. V. (2016). Modeling the BOD of Danube River in Serbia using spatial, temporal, and input variables optimized artificial neural network models. *Environmental Monitoring and Assessment, 188*(5), 1–12.

Wang, Z., Liu, C. X., Liao, J., Liu, L., Liu, Y. H., Huang, X., et al. (2014a). Nitrogen removal and N_2O emission in subsurface vertical flow constructed wetland treating swine wastewater:

Effect of shunt ratio. *Ecological Engineering Journal, 72*(2014), 446–453.

Wang, Z., Xie, J., Fang, D., Wu, D. Y., Kong, H. N., Liu, A. W., et al. (2013). Nitrogen pollution load in farmlands under two typical farming systems in the north of Lake Erhai region. *Journal of Ecology and Rural Environment, 29*(5), 625–629.

Wang, F., Wang, X., Zhao, Y., & Yang, Z. (2014b). Long-term periodic structure and seasonal-trend decomposition of water level in Lake Baiyangdian, northern China. *International journal of Environmental Science and Technology, 11*(2), 327–338.

Wang, S., Zhang, L., Ni, L., Zhao, H., Jiao, L., Yang, S., Guo, L., & Shen, J. (2015). Ecological degeneration of the Erhai Lake and prevention measures. *Environmental Earth Sciences, 74*(5), 3839–3847.

Xu, Y. F., Ma, C. Z., Liu, Q., Xi, B. D., Qian, G. G., Zhang, D. Y., Hou, S. L., et al. (2005). Method to predict key factors affecting lake eutrophication: anew approach based on support vector regression model. *International Biodeterioration & Biodegradation, 102*(2015), 308–315.

Yan, C. Z., Jin, X. C., Zhao, J. Z., Shen, B., Li, N. B., Huang, C. Z., et al. (2005). Ecological protection and sustainable utilization of Erhai Lake, Yunnan. *Environmental Sciences, 26*(5), 38–42.

Young, C. C., Liu, W. C., & Hsieh, W. L. (2015). Predicting the water level fluctuation in an alpine lake using physically based, artificial neural network, and time series forecasting models. *Mathematical Problems in Engineering, 2015*, 1-11.

Yin, Y., Chu, Z., Zhao, M., Li, Z., Ye, B., & Jin, X. (2011). Spatial and temporal changes in water quality in aquatic-terrestrial ecotone of Lake Erhai. *China Environmental Science, 31*(7), 1192–1196.

Zhang, L., Wang, S., Li, Y., Zhao, H., & Qian, W. (2015). Spatial and temporal distributions of microorganisms and their role in the evolution of Erhai Lake eutrophication. *Environmental Earth Sciences, 74*(5), 3887–3896.

Zhang, L., Wang, S., & Wu, Z. (2014). Coupling effect of pH and dissolved oxygen in water column on nitrogen release at water-sediment interface of Erhai Lake, China. *Estuarine, Coastal and Shelf Science, 149*, 178–186.

Zhang, Z., Wang, X. Z., LI, Y. H., et al. (2013). Removal of endocrine disrupting chemicals in rural sewage by soil trench systems. *Research of Environmental Sciences, 26*(2), 173–180.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.