Environ Model Assess (2016) 21:233–246
DOI 10.1007/s10666-015-9472-4

Eutrophication Prediction Using a Markov Chain Model: Application to Lakes in the Yangtze River Basin, China

Jiacong Huang [1] · **Junfeng Gao** [1] · **Yinjun Zhang** [2]

Received: 14 September 2014 / Accepted: 2 June 2015 / Published online: 11 June 2015

Abstract Lake eutrophication is harmful and difficult to predict due to its complex evolution. As an alternative to existing mechanistic models, a Markov chain model was developed to predict the development of lake eutrophication based on an 11-year dataset in 41 lakes of the Yangtze River Basin. This model was validated using a real-time update strategy and was demonstrated to be reliable. Based on the dataset, the lake eutrophication dynamics from 2000 to 2010 were analyzed. Lakes with different trophic states from 2011 to 2050 and their responses to different water management practices were simulated based on the developed model. The simulation results show that lake eutrophication would worsen from 2011 to 2040; however, eutrophication could be significantly alleviated by changing 100 km^2 of hypereutrophic lakes into eutrophic lakes per year from 2010 to 2020. The nutrient conditions in most of the lakes in the Yangtze River Basin show that phosphorus control would be more efficient than nitrogen control in eutrophication management practices. This case study demonstrates the utility of Markov chain models in using prior information to predict the long-term evolution of lake eutrophication at large spatial scales. The Markov chain technique can be easily adapted to predict evolutionary processes in other disciplines.

✉ Junfeng Gao
GaoJunf@niglas.ac.cn

[1] Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, 73 East Beijing Road, Nanjing 210008, China

[2] China National Environmental Monitoring Centre, 8 (B) Dayangfang Beiyuan Road, Chaoyang District, Beijing 100012, China

Keywords Eutrophication prediction · Markov chain · Eutrophication assessment · Yangtze River Basin

1 Introduction

Lake eutrophication is a global environmental problem, which has deleterious effects on aquatic ecosystems [23]. Predicting trends in lake eutrophication in the near future would help us understand the mechanisms of eutrophication in lake ecosystems, and thus provide strategies for lake managers to control lake eutrophication [16, 17]. For an individual lake, eutrophication indicators (e.g., chlorophyll *a*, total phosphorus, and nitrogen) can be predicted using a wide range of existing lake models [15, 21, 22, 32, 33]. However, lakes in watersheds are strongly affected by connecting rivers [14]. These lakes should be modeled as a dynamic system (termed "lake chains" in Hilt et al. [14]) rather than a group of individuals. Understanding the evolutionary patterns of lake eutrophication in a watershed is necessary to predict the dynamics of the entire system.

Markov chains were developed by a Russian mathematician Andrei Andreyevich Markov in 1907 and have a flexible structure for describing the evolution of systems based on prior information [3, 20, 38]. Markov chains, therefore, have been widely used in predicting the time scales over which the evolution of a dynamic system unfolds. For example, Markov chains have been used to study forests [6, 30, 38], landscapes [2, 12, 13], ecological networks [25], weather conditions [1], and the evolution of populations [9, 27]. Markov chain models provide a flexible transition probability matrix to describe the transformation of the system from one stage to the next [3]. This feature allows Markov chain models to predict the evolution of eutrophication of lakes in a watershed.

A majority of Chinese freshwater lakes are located in the Yangtze River Basin. During the last few decades, these aquatic ecosystems have experienced accelerated levels of eutrophication [19, 24]. Eutrophication control of these lakes has been among the top priorities for Chinese governments. Large-scale predictions of lake eutrophication over long-time scales could describe developing trends in the trophic state of lakes, and have the potential to support integrated watershed management. Such predictions are imperative for lake management agencies. However, the evolutionary pattern of lake eutrophication over such a large spatial scale is difficult to predict because there are a dearth of large datasets for the lakes in the Yangtze River Basin.

The main objective of this paper is to predict the long-term eutrophication dynamics of the lakes in the Yangtze River Basin. Based on an 11-year dataset for 41 lakes, a Markov chain model was developed by following the steps: (1) lake trophic states were assessed using a trophic state index; (2) a transition probability matrix was derived based on eutrophication assessment results from 2000 to 2010; and (3) the development of lake trophic states from 2011 to 2050 was predicted using the transition probability matrix. Performance of the Markov chain in this case study and the implications of the predictions for lake management were discussed.

2 Material and Methods

2.1 Study Area

The Yangtze River (length, 6300 km) is the longest river in China and the third longest river in the world, with a total drainage area of 1,808,500 km^2 (approximately 18.8 % of China's territory) (Fig. 1). The distribution of precipitation in the Yangtze River Basin is highly heterogeneous. The annual mean precipitation varies from 270–500 mm in the western region to 1600–1900 mm in the southeastern region [11].

There are 648 lakes with an area larger than 1 km^2 in the Yangtze River Basin. The total area of these lakes is 17, 178.5 km^2, accounting for approximately 21.2 % of the surface area of all of the lakes in China [31]. Most of these lakes are rather shallow. The Yangtze River and the lakes in the Yangtze River Basin have been listed in the Global Ecoregion 200 by the World Wildlife Fund (WWF) for conservation [43]. These lakes in the Yangtze River Basin provide important water resources for residential dwellings, industry, and irrigation. Unfortunately, many of these aquatic ecosystems have become eutrophic or are becoming eutrophic due to intensive ecological stress (e.g., high nutrient loading and extreme climate) in recent decades [19, 24]. To control lake eutrophication, many water quality-monitoring programs have been conducted by researchers and managers to obtain data on water quality in these aquatic ecosystems.

2.2 Data

A total of five indicators were used for eutrophication assessment of the lakes in the Yangtze River Basin, including chlorophyll *a* (Chl *a*, μg/l), total phosphorus (TP, mg/l), total nitrogen (TN, mg/l), chemical oxygen demand (COD, mg/l) and Secchi disk depth (SD, m). Water samples were collected and analyzed by the Ministry of Environmental Protection of the People's Republic of China. This sampling program was conducted at 173 sampling sites in 41 lakes (reservoirs) (Fig. 1 and Appendix A) of the Yangtze River Basin from 2000 to 2010. These lakes cover a total area of 12,741 km^2 and are a

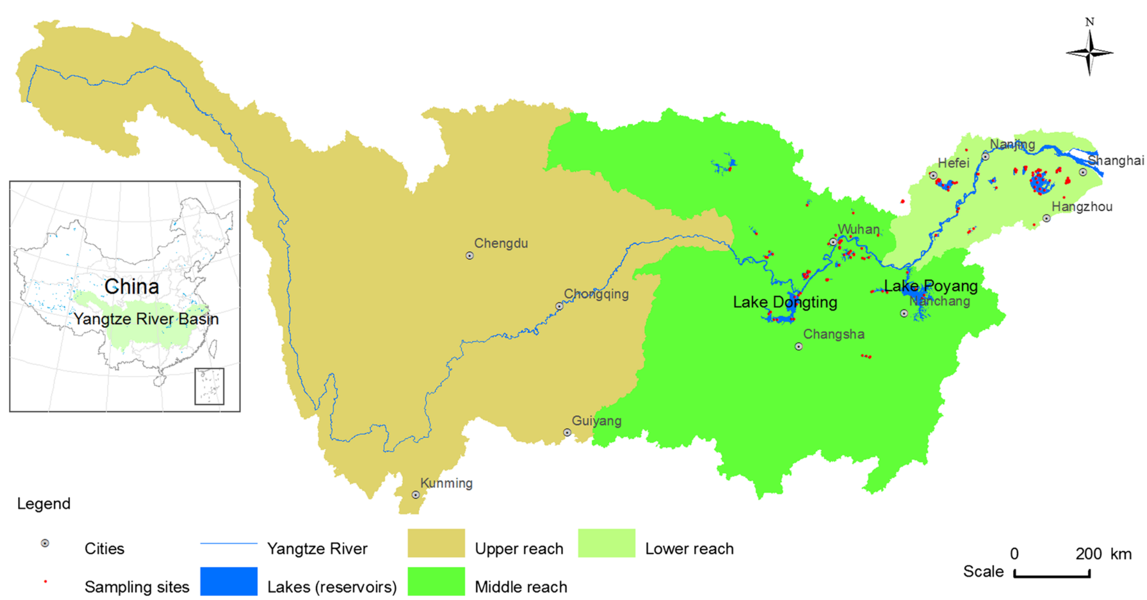

Fig. 1 Location of the Yangtze River Basin and sampling sites in 41 lakes (reservoirs)

concern of researchers and government officials due to their importance. The number of sampling sites for a lake was positively associated with its area and importance for the human population. In other words, more samples were collected from larger lakes that were more heavily used by the human population. During this 11-year period, water samples were collected seasonally between 2000 and 2006, and monthly between 2007 and 2010. Annual average values of these five indicators (i.e., Chl *a*, TP, TN, COD, and SD) were derived using seasonal and monthly values.

2.3 Lake Eutrophication Assessment

Many eutrophication assessment methods have been proposed to evaluate lake trophic state, e.g., the fixed boundary criteria developed by the OECD [35], Carlson's trophic state index [5], fuzzy analysis [28], and artificial neural networks [39]. Among these methods, OECD's criteria and Carlson's trophic state index have been most widely used. To obtain a more reliable assessment of lake trophic state among different regions, improved methods of indicator selection and parameter optimization have been developed [4, 18, 28, 37, 41, 44].

Carlson's trophic state index (TSI) has the advantage of providing continuous numerical classes to represent lake trophic state [5, 44], and was well suited for Chinese lakes [8, 18, 44–46]. The TSI ranges from 0 to 100 with a higher value implying more severe eutrophication, and can be calculated by:

$$\mathrm{TSI} = \sum_{j=1}^{m} w_j \mathrm{TSI}_j \tag{1}$$

where w_j is the weight of the j-th indicator for the TSI, TSI_j is the trophic state index of the j-th indicator, and m is the number of indicators. These assessment indicators for calculating TSI include basic and additional indicators. A relation-weighting index was used to determine the weights for all of these indicators [18, 45].

$$w_j = \frac{r_{ij}^2}{\sum_{j=1}^{m} r_{ij}^2} \tag{2}$$

where r_{ij} is the relation coefficient between the j-th indicator and the base indicator. TSI_j is calculated as follows [45]:

$$\mathrm{TSI}_j = 10\left(a_j + b_j \ln C_j\right) \tag{3}$$

where C_j is the measured value of the j-th indicator; a_j and b_j are constants, and can be calculated using the following equations:

$$a_j = -10 \frac{\ln C_{j\min}}{\ln C_{j\max} - \ln C_{j\min}} \tag{4}$$

$$b_j = 10 \frac{1}{\ln C_{j\max} - \ln C_{j\min}} \tag{5}$$

where $C_{j\min}$ is the measured lowest value of the j-th indicator, and $C_{j\max}$ is the measured highest value of the j-th indicator.

The results of assessment (TSI values) critically depend on indicator selection and the parameter values of w_j, a_j, and b_j. In this study, chlorophyll *a* concentration (Chl *a*) was selected as the base indicator due to its close relationship to lake trophic state [45]. Total phosphorus (TP), total nitrogen (TN), chemical oxygen demand (COD), and Secchi disk depth (SD) were selected as additional indicators given their well representation of lake trophic states in China [18, 45]. The constants (w_j, a_j, and b_j) for these assessment indicators were determined based on the large-scale survey data from a previous study [18]. These values of w_j, a_j, and b_j have been successfully used in eutrophication assessment of Chinese lakes [10]. Thus, TSI for lakes in the Yangtze River Basin can be calculated by [10]:

$$\begin{aligned} \mathrm{TSI} = {} & 0.1878\mathrm{TSI}_{\mathrm{TP}} + 0.1794\mathrm{TSI}_{\mathrm{TN}} + 0.1833\mathrm{TSI}_{\mathrm{COD}} \\ & + 0.1833\mathrm{TSI}_{\mathrm{SD}} + 0.2662\mathrm{TSI}_{\mathrm{Chl}} \end{aligned} \tag{6}$$

where TSI_j are obtained as follows [10]:

$$\mathrm{TSI}_{\mathrm{TP}} = 10(9.436 + 1.624\ln \mathrm{TP}) \tag{7}$$

$$\mathrm{TSI}_{\mathrm{TN}} = 10(5.453 + 1.694\ln \mathrm{TN}) \tag{8}$$

$$\mathrm{TSI}_{\mathrm{COD}} = 10(0.109 + 2.661\ln \mathrm{COD}) \tag{9}$$

$$\mathrm{TSI}_{\mathrm{SD}} = 10(5.118 - 2.117\ln \mathrm{SD}) \tag{10}$$

$$\mathrm{TSI}_{\mathrm{Chl}} = 10(2.5 + 1.086\ln \mathrm{Chl}) \tag{11}$$

The trophic states of 173 sampling sites at 41 lakes in the Yangtze River Basin (Fig. 1) from 2000 to 2010 were assessed using the TSI method. The TSI value (ranging from 0 to 100) was then evenly divided into 10 trophic states (1–10) with an interval of 10. The resulting TSI of a sampling site represents the trophic state of an area (S_m) determined by the area (S_{lake}) and the number of sampling sites for the lake (n_{lake}).

$$S_m = S_{\mathrm{lake}} / n_{\mathrm{lake}} \tag{12}$$

For example, there are 23 sampling sites in Lake Taihu, with an area of 2298 km^2. The resulting TSI of each sampling site represents the trophic state of an area (99.9 km^2) in Lake Taihu. Based on Eqs. 6–12, the lake area for each trophic state can be calculated.

2.4 Markov Chain for Eutrophication Prediction

The Markov chain is widely used to predict the evolution of a system, e.g., for forest [30, 38] and land use [2, 12]. In this study, this technique was adapted to predict the time scales of the evolution of lake eutrophication with a conceptual representation in Fig. 2. A Markov chain model was developed and applied by following the steps (Fig. 3);

1. Representation of lake trophic states

All of the lakes in the Yangtze River Basin were assumed to be a close system. Their trophic states at time t were represented by a matrix X_t.

$$X_t = [S_{1t} \quad S_{2t} \quad \dots \quad S_{nt}] \tag{13}$$

where S_{it} ($i \in \{1,2,\dots,n\}$) is the total area (km$_2$) of lakes with a trophic state of i at time t. n ($n=10$) is the number of lake trophic states. The predicted value of X_t using the Markov chain model is represented by $\hat{X}_t$.

$$\hat{X}_t = \left[\hat{S}_{1t} \quad \hat{S}_{2t} \quad \dots \quad \hat{S}_{nt}\right] \tag{14}$$

where $\hat{S}_{it}$ is the predicted value of S_{it} using the Markov chain model.

2. Description of lake trophic state dynamics

We assume that the predicted lake trophic state $\left(\hat{X}_{t+\Delta t}\right)$ only depends on the state of X and not on other previous states. Thus, the temporal dynamics of the lake trophic states $\left(\hat{X}_{t+\Delta t}\right)$ can be predicted using a transition probability matrix P.

$$\hat{X}_{t+\Delta t} = X_t P \tag{15}$$

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{16}$$

where $p_{ij}(p_{ij} > 0, \sum_{j=1}^{n} p_{ij} = 1,$ $i, j \in \{1, 2, \dots, n\})$ is the probability that lake trophic state changes from the i-th state to the j-th state in a time step (Δt=1 year in this study), i.e.,

$$p_{ij} = \frac{S_{it,\ j(t+\Delta t)}}{S_{it}} \tag{17}$$

where $S_{it,\ j(t+\Delta t)}$ is the total area (km^2) of lakes with a trophic state of i at time t, and with a trophic state of j at time $t+\Delta t$.

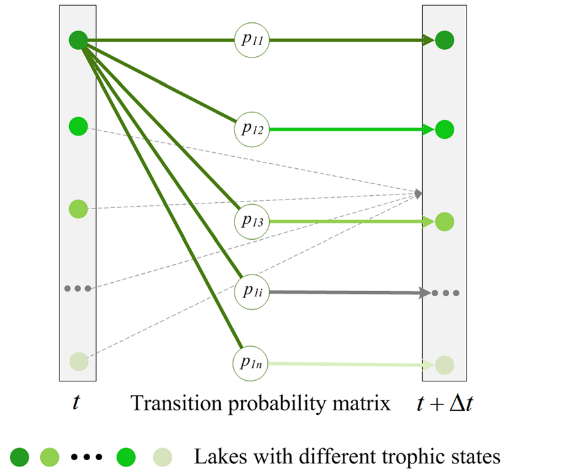

Fig. 2 Conceptual diagram of the Markov chain model for predicting lake trophic state. p_{ij} represents the probabilities that lake trophic state changes from the i-th state to the j-th state in a time step (Δt)

3. Estimation of the transition probability matrix

The element of the transition probability matrix (P) was calculated based on the lake eutrophication assessment results from 2000 to 2010 (mentioned in Section 2.3).

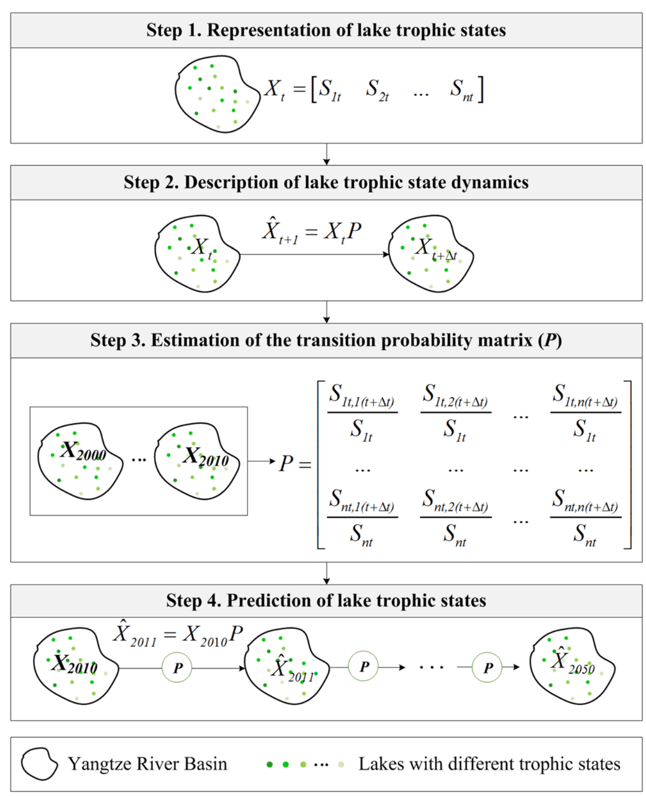

Fig. 3 Flow chart of the Markov chain model for predicting the trophic states of lakes in the Yangtze River Basin. P is the transition probability matrix for the dynamics of the lake trophic state. X_t represents the lake trophic state at year t. $\hat{X}_t$ represents the predicted lake trophic state at year t

$$p_{ij} = \frac{S_{it,j(t+1)}}{S_{it}} \quad (18)$$

where $S_{it,j(t+1)}(i,j \in \{1,2,\ldots,n\})$ is the area of lakes with an i-th trophic state at year t, and with a j-th trophic state at year $t+1$.

4. Prediction of lake trophic states

Based on the estimated P, a simulation (SimBase) was implemented to predict the trophic states of lakes in the Yangtze River Basin from 2011 to 2050 using Eq. 15. The initial condition of this simulation is from the assessment results of lake eutrophication for the year 2010 (X_{2010}). X dynamics are assumed to be homogeneous in time, i.e., P is constant during the simulation period.

5. Predicting the impacts of water management practices on lake trophic states

In water management practices, the hypereutrophic lakes are more of a concern than other lakes due to their deleterious impacts. Thus, lake managers have taken many measures (e.g., reducing nutrient inputs and sediment removal) to lower their trophic states. To evaluate the benefits of these measures on lake trophic states, another three simulations were implemented and compared with the simulation of SimBase. These three simulations (SimChange10, SimChange50, and SimChange100) assumed a specified area (10, 50, and 100 km_2, respectively) of lake where the trophic state of level 8 was improved to level 7 (Section 2.5) per year from 2010 to 2020. Such an assumption was described in the Markov chain model using a dynamic P by changing the values of $S_{8t,7(t+1)}$ and $S_{8t,8(t+1)}$ per year from 2010 to 2020.

$$S_{8t,7(t+1)} = S_{8t,7(t+1)} + s \quad (20)$$

$$S_{8t,8(t+1)} = S_{8t,8(t+1)} - s \quad (21)$$

where s ($s \in \{10, \ 50, \ 100 \text{ km}^2\}$) is the specified area of the lake where the trophic state of level 8 was improved to level 7 per year from 2010 to 2020.

2.5 Model Validation

The performance of a Markov chain model is highly determined by the transition probability matrix (P). Thus, a reliable estimation of P is needed to achieve a successful simulation. A real-time update strategy was used to validate the reliability of the Markov chain model (Section 2.4) from 2006 to 2010. In other words, the transition probability matrix was updated in real time during the validation period (Fig. 4). The area of lakes with different trophic states (levels 1–10) from measured data were compared with that from the estimated results of the Markov chain. Model fit was evaluated based on the following error statistics of ε_{it}:

$$\varepsilon_{it} = \frac{\left|\hat{S}_{it} - S_{it}\right|}{\sum_{i=1}^{n} S_{it}/n} \quad (22)$$

where S_{it} is the area of lakes with the trophic state level of i at year t. $\hat{S}_{it}$ is the estimated area of lakes with the trophic state level of i at year t. n is the number of trophic state levels (n= 10). ε_{it} is a dimensionless measure ranging from 0 to infinity.

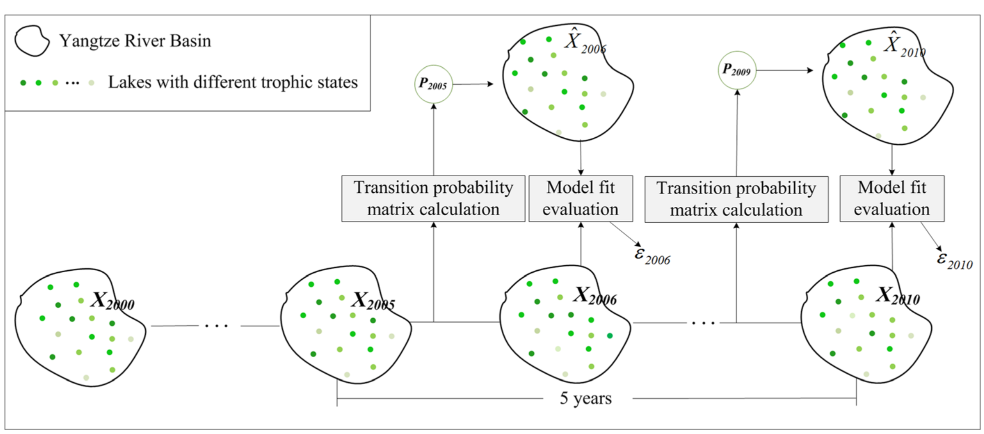

Fig. 4 Validation of the Markov chain model. P_t is the transition probability matrix for predicting lake eutrophication state in the year $t+1$. X_t represents the lake eutrophication state at year t. $\hat{X}_t$ represents the predicted lake eutrophication state at year t. ε_t is the deviation between X_t and $\hat{X}_t$ at year t

An ε_{it} value of 0 indicates that the estimated area ($\hat{S}_{it}$) matches the measured area (S_{it}).

3 Results

3.1 Lake Eutrophication From 2000 to 2010

The resulting TSI (ranging from 0 to 100) represents different trophic states: oligotrophic (0–30), mesotrophic (31–60), eutrophic (61–70), and hypereutrophic (71–100) [44]. The area percentages of these four lake trophic states from 2000 to 2010 are shown in Fig. 5. The area of hypereutrophic lakes decreased from 2006 to 2010, and the area percentage decreased from 14.83 to 4.58 %. However, the total area percentage of eutrophic and hypereutrophic lakes did not change significantly between 2002 and 2010. Because the two largest lakes (Lake Poyang and Dongting) in the Yangtze River Basin were mesotrophic, the area of mesotrophic lakes accounted for a large proportion of the area from 2000 to 2010. The area of the oligotrophic lakes decreased from 2007 to 2010, and the area percentage decreased from 7.65 to 1.27 %. There was a clear trend in oligotrophic lakes becoming more mesotrophic from 2007 to 2010.

The trophic state index (TSI) in this study was calculated by five indicators (Chl *a*, TP, TN, COD, and SD), suggesting that high trophic state may result from different environmental factors. For example, for Lake Poyang from 2003 to 2010, the average TSI values of TP and TN were 59.3 and 53.5, respectively, while the TSI value of Chl *a* was 16.9. The TSI value of TN was significantly higher than that of TP (Fig. 6).

The trophic states of lakes in the Yangtze River Basin are strongly affected by their locations and functions. Many of the lakes in developed and underdeveloped regions had an increasing trend toward eutrophication (e.g., Lakes Changdang, Chaohu, and Poyang in Fig. 7 a, b). However, lake eutrophication in the developed regions (e.g., Lake Changdang) was more severe than that in the underdeveloped regions (e.g., Lakes Chaohu and Poyang). For example, Lake Changdang, located in one of the most developed regions (Jiangsu Province) in China, had a TSI value of 70 in 2007, suggesting severe eutrophication. The eutrophication of many lakes in cities (e.g., Lake Xuanwu in Nanjing City and Lake Yushan in Maanshan City) was severe. However, eutrophication of these lakes was alleviated from 2000 to 2010 (Fig. 7 c). The trophic states of many lakes used for drinking water (e.g., Dongpu Reservoir in Anwei Province) were relatively low, and did not change significantly from 2000 to 2010.

3.2 Transition Probability Matrix for Predicting Lake Eutrophication in the Yangtze River Basin

The area ($S_{it,j(t+1)}$ in Eq. 18, km^2) of lakes from the *i*-th trophic state to the *j*-th trophic state was calculated based on the results of lake eutrophication assessment from 2000 to 2010. The assessment generated 10×10 matrix (Table 1), which should be read as follows: (1) for row 3, column 4, the total area of the lakes, whose trophic states changed from level 3 to 4 in a year, was 2493.4 km^2; (2) for row 3, column 5, the total area of the lakes, whose trophic states changed from level 3 to 5 in a year, was 98.9 km^2.

Based on Eq. 18, the transition probability matrix (*P*) was derived from the $S_{it,j(t+1)}$ matrix (Table 1). The value of *P* (Table 2) should be read as follows: (1) for row 7, column 4, there is no probability for lake trophic state to change from level 7 to 4 in a year; (2) for row 7, column 6, there is a probability of 0.122 for lake trophic state to change from level 7 to 6 in a year. The values on the diagonal line in Table 2 represent the probability of different lake trophic states to remain unchanged in a year. The values at the upper right of Table 2 represent the probability of different lake trophic states increasing in a year, and the values at the lower left of Table 2 represent the probability of different lake trophic states decreasing in a year. The

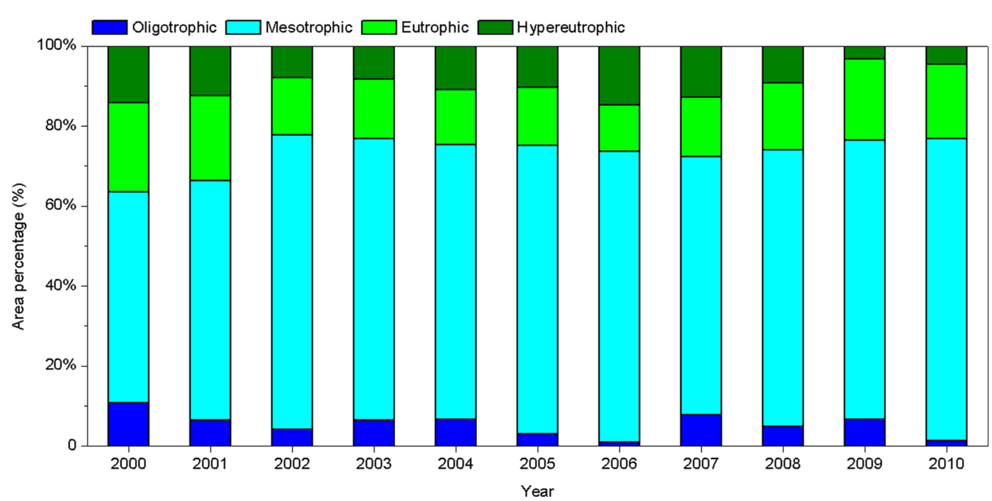

Fig. 5 Area percentages of four lake trophic states in the Yangtze River Basin from 2000 to 2010

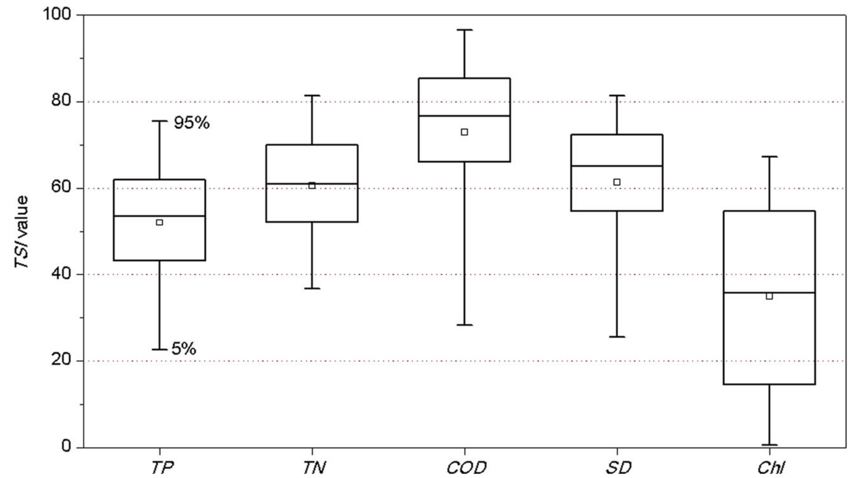

Fig. 6 Trophic state index (TSI) for five indicators (*Chl a* chlorophyll *a*, *TP* total phosphorus, *TN* total nitrogen, *COD* chemical oxygen demand, and *SD* Secchi disk depth) based on all of the samples at 173 sampling sites of 41 lakes (reservoirs) between 2000 and 2010

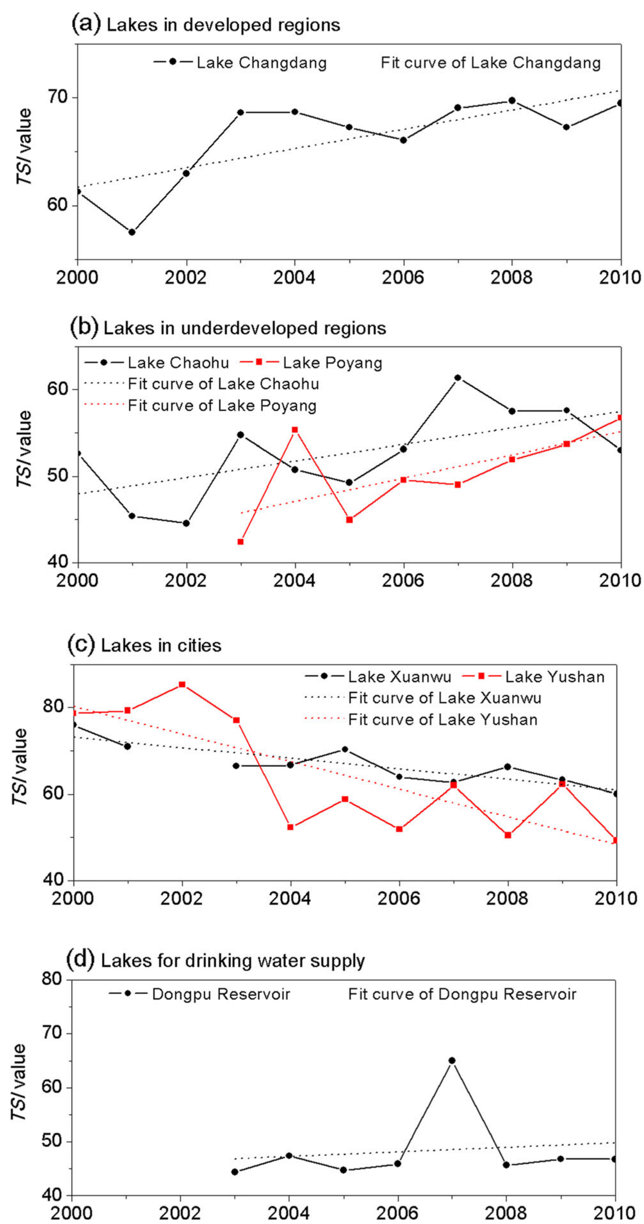

Fig. 7 The trophic state index (TSI) for different types of lakes from 2000 to 2010

Table 1 The area ($S_{it,j(t+1)}$ (km^2) in Eq. 18) of lakes from the *i*-th trophic state (row number) to the *j*-th trophic state (column number) in a year

Eutrophication level	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	0
2	0	0	117.4	11.6	0	0	0	0	0	0
3	0	0	2860.7	2493.4	98.9	0	0	0	0	0
4	0	0	1985.5	6843.2	4815.9	1444.2	0	0	0	0
5	0	0	149.5	2983.2	27,662.8	6380.4	41.6	0	0	0
6	0	0	112.4	495.9	5042.9	9584.6	1875.2	80.8	64.1	0
7	0	0	0	0	73.5	2106.0	12,854.1	2198.0	64.1	0
8	0	0	0	0	0	29.9	2993.4	7517.1	100.6	0
9	0	0	0	0	0	0	0	557.3	0	0
10	0	0	0	0	0	0	0	0	0	0

The area was calculated using the results from the lake eutrophication assessment from 2000 to 2010

values in Table 2 revealed that the low trophic states (levels 2–5) have a higher probability of increasing in any given year. However, the high trophic states (levels 6, 8, and 9) have a higher probability of decreasing in any given year.

3.3 Model Validation

The model fits of the Markov chain model during the validation period (2006–2010) were significantly different (Fig. 8). In 2006, the ε value (Eq. 22) varied from 0 to 214 % with an average value of 55.7 %, suggesting that the lake areas with different trophic states were not adequately predicted. However, it was clear that model fit was improved for the period from 2006 to 2010. In 2010, the estimated results agreed well with the measured data, with an average ε value of 10.9 %. This value (ε_{2010}) showed a significant decrease of 44.8 % compared with the ε value in 2006 (ε_{2006}).

3.4 Lake Eutrophication From 2011 to 2050

The prediction results of SimBase show that the trophic state of lakes in the Yangtze River Basin would worsen from 2011 to 2040 (Fig. 9) without implementing any mitigating measures. There is a strong tendency to change from a mesotrophic state to other trophic states from 2011 to 2025. The area of hypereutrophic-state lakes would increase, while the area of mesotrophic-state lakes would decrease. The area of eutrophic-state lakes would decrease slightly during the first 2 years with a minimum value of 2163.7 km^2 in 2013, and then increase afterward. The area of oligotrophic-state lakes would

Table 2 The estimated transition probability matrix using the results of lake eutrophication assessment from 2000 to 2010

Eutrophication Level	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0.910	0.090	0	0	0	0	0	0
3	0	0	0.525	0.457	0.018	0	0	0	0	0
4	0	0	0.132	0.453	0.319	0.096	0	0	0	0
5	0	0	0.004	0.080	0.744	0.171	0.001	0	0	0
6	0	0	0.007	0.029	0.292	0.554	0.109	0.005	0.004	0
7	0	0	0	0	0.004	0.122	0.743	0.127	0.004	0
8	0	0	0	0	0	0.003	0.281	0.707	0.009	0
9	0	0	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	0	0	0

The highest transition rate of each trophic level is colored dark gray. The values on the diagonal line represent the probability of different lake trophic states remaining unchanged throughout the year

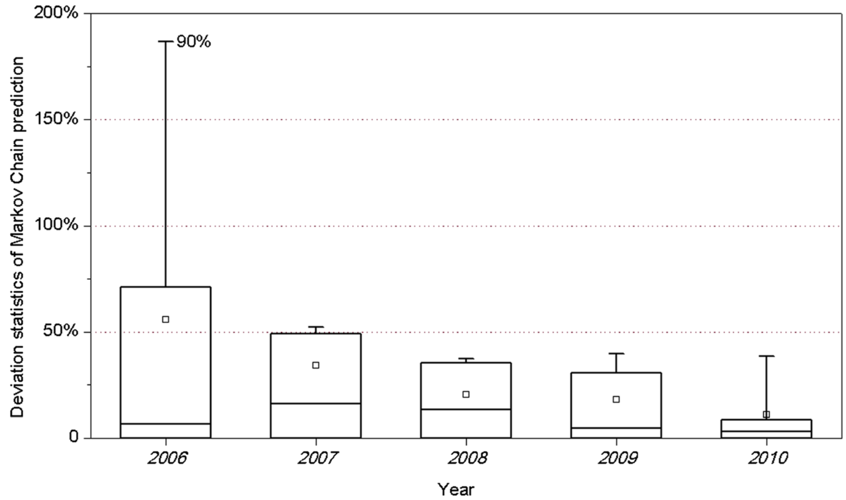

Fig. 8 Error statistics of the Markov chain model during the validation period (2006–2010)

increase rapidly in the first 7 years with a maximum value of 413.5 km^2 in 2017, and then decrease slowly afterward.

SimBase required 30 years (2011–2040) for lake eutrophication to achieve a nearly steady state (Fig. 9). The areas of lakes with different trophic states would remain stable in the 2040s with an average standard deviation of 4.2 km^2 (0.13 % of the total lake area). The stable lake areas of four different trophic states (oligotrophic, mesotrophic, eutrophic, and hypereutrophic) were 390.7; 8393.3; 2504.7; and 1261.1 km^2, respectively. Compared to the trophic states in 2010, the area of mesotrophic lakes significantly decreased by 1080.2 km^2, while the area of hypereutrophic lakes significantly increased by 686.2 km^2.

SimChange10, SimChange50, and SimChange100 showed significantly different results in predicting the areas of mesotrophic, eutrophic, and hypereutrophic lakes (Fig. 9). The area of hypereutrophic lakes decreased significantly with a stable area of 315.1 km^2 for the simulation of SimChange100. This area had a decrease of 946.0 km^2 in a simulation without any measures (SimBase). However, the areas of mesotrophic and eutrophic lakes increased. The area of oligotrophic lake showed no significant difference among these four simulations (Fig. 9 (a)).

4 Discussion

4.1 Performance of the Markov Chain

Lake eutrophication dynamics are complex, and strongly affected by a wide range of anthropogenic (e.g., land use and point-source pollution in its catchment area) and natural factors (geographic location, lake morphology, and climate conditions) [29, 34]. As an alternative to existing mechanistic models, a Markov chain model was developed to predict lake eutrophication in this study. The model validation results (Fig. 8) suggest that the Markov chain model we developed in this study is reliable and well suited to predict the eutrophication dynamics of a lake community at such a large spatial scale. The successful application of the Markov chain is primarily attributed to the following:

1. The strength of the Markov chain to embed prior information into the transition probability matrix — Evolutionary information from prior data of various disciplines can be easily described by the transition probability matrix. Moreover, this transition probability matrix could be updated in real time for a more reliable prediction.

2. The extensive dataset used in this study — The extensive data facilitated a reliable estimation of the transition probability matrix, and thus improved model performance. This conclusion was supported by the results of the model validation (Fig. 8). Model fit was relatively low in 2006, suggesting that prior information (2000–2005) was still not adequate to make reliable predictions for eutrophication dynamics. However, the fitting performance of this model improves from 2006 to 2010 due to the increasing amount of prior information used for the Markov chain model (Fig. 8). Thus, adequate prior information is needed to develop a reliable model.

4.2 Possible Applications of the Markov Chain

Although the Markov chain was used for a specific case in this study, it could be easily adapted for predicting dynamic

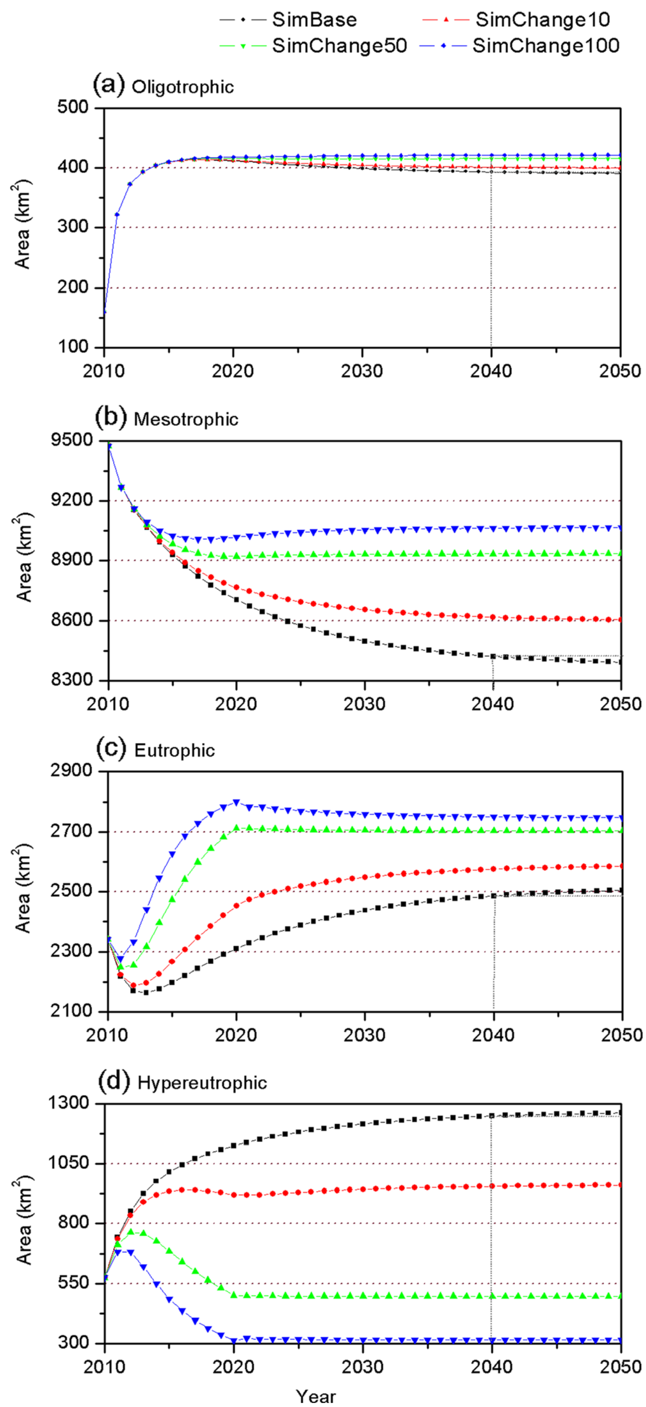

Fig. 9 The areas of lakes with different trophic states (oligotrophic, mesotrophic, eutrophic, and hypereutrophic) in the Yangtze River Basin from 2011 to 2050 predicted by the Markov chain model. SimChange10, SimChange50, and SimChange100 represent simulations that assumed a specified area (10, 50, and 100 km^2, respectively) of lake where a trophic state of level 8 was improved to level 7 per year from 2010 to 2020. SimBase is the simulation without any improvement

systems in aquatic sciences or even in other scientific disciplines. Markov chains are capable of being adapted widely, because their dynamics are rooted in statistics rather than in mathematical descriptions of the eutrophication process. The specific case in this study used the indicator of S_{it} (the lake area with a trophic state) to represent the state of a dynamic system (X). The transition probability matrix (P) represents the increase or decrease of lake area from one trophic level to another at each time step. Other indicators (e.g., phytoplankton biomass and species abundances) could be predicted with a redefinition of X and P following the steps in Fig. 2. Moreover, Markov chain models allow time steps to be manipulated; any time step (e.g., 1 year, 1 month, or 1 day) may be used, depending on the requirements for other specific case studies.

The manual measures were successfully coupled to the simulations of SimChange10, SimChange50, and SimChange100. This success demonstrated the potential of the Markov chain in modeling the impacts of water management strategies on lake eutrophication. However, to achieve a better Markov chain model, great efforts are required to quantify the relationship between these measures and lake areas with different trophic states. Moreover, other unexpected disturbances, such as natural (e.g., extreme climate) and human (e.g., large-scale ecological restorations) impacts, may alter the evolution of lake eutrophication considerably and were worth quantitatively coupling in this model. For example, there was a plan to build a dam at the outlet of Lake Poyang. This project would alter the hydrodynamic conditions in a large area of Lake Poyang significantly, and thus result in abrupt shifts in lake eutrophication dynamics. To overcome this weakness, system disturbances (e.g., nutrient reduction, natural disasters, ecological restorations, and extreme climate) could be included in a Markov chain based on the technical advances, such as the use of Markov decision processes [36]. However, the relationship between disturbance and its resulting change in nutrient concentration should first be quantified.

4.3 Implications for Lake Management

Lake eutrophication is highly concerned by lake managers and researchers due to its deleterious effects. The results of lake eutrophication assessment show that most of the lakes in cities (urban lakes) are experiencing serious levels of eutrophication due to the high loading of nutrients from waste water [19]. However, one encouraging fact is that the trophic states of many urban lakes were under control from 2000 to 2010 (Fig. 7 c). This finding is attributed to substantial inputs (e.g., ecological restorations and reductions in nutrient inputs) for improving water quality of urban lakes due to their important roles for cities. Compared with other types of lakes, lakes used for drinking water were protected from water contamination and had relatively low trophic-state level (Fig. 7 d).

Long-term predictions of lake eutrophication at a large spatial scale provide a good overview of lake eutrophication evolution in the Yangtze River Basin. The long periods required for lake eutrophication to achieve a steady state (Fig. 9) implied that the current trophic states of lakes in the Yangtze River Basin are in an unstable state. Although the trophic state predictions of the lakes in the Yangtze River Basin are pessimistic from 2011 to 2050 (Fig. 9), the simulation results of SimChange10, SimChange50, and SimChange100 revealed that some management strategies could be taken to ameliorate the eutrophication dynamics. Nitrogen and phosphorus control is one of the most important strategies for reducing eutrophication in aquatic ecosystems [7]. Assessment results (Fig. 6) show that the TSI value of TP is lower than that of TN. This implies that lake eutrophication in the Yangtze River Basin is more limited by phosphorus than by nitrogen. From this perspective, phosphorus control is more effective for mitigating lake eutrophication for most of the lakes in the Yangtze River Basin. However, for an individual lake, N:P ratio calculations are required for deducing that nitrogen or phosphorus is limiting eutrophication.

Water retention time is widely recognized as a key factor affecting lake eutrophication [40, 42]. Thus, water transfer has been used to flush pollutants out of lakes, resulting in the abrupt alleviation of lake eutrophication on short-time scales [26]. In contrast, dam-building projects increase water retention time of the lake and may therefore facilitate eutrophication.

5 Conclusions

A Markov chain model was developed to predict the time scale of the evolution of lake eutrophication in the Yangtze River Basin using an 11-year dataset. This study demonstrated the successful application of a Markov chain for long-term eutrophication prediction and provided an alternative method for predicting eutrophication evolution. The high performance of Markov chains depended on its strength of converting prior information into transition matrices. Markov chain techniques could thus be promising for uses in other disciplines with adequate prior information. The prediction results from our case study show a trend toward increased lake eutrophication in the Yangtze River Basin from 2011 to 2050. However, such trend could be gradually alleviated by taking appropriate controlling measures. These prediction results could help lake managers better understand the complex mechanisms of eutrophication and develop efficient eutrophication management strategies to control lake eutrophication.

Acknowledgments The project was financially supported by the National Basic Research Program of China (No. 2012CB417006) and "135" Key Program in Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences (No. NIGLAS2012135005). The authors would like to thank the Ministry of Environmental Protection of the People's Republic of China for providing the measured data for lake eutrophication assessment. The boundary data of the lakes in Yangtze River Basin were obtained from Scientific Data Sharing Platform for Lake and Watershed.

Appendix

Table 3 shows the statistical information (area, number of sampling sites, and trophic state index between 2000 and 2010) of lakes (reservoirs) surveyed in this study.

Table 3 Statistical information (area, number of sampling sites, and trophic state index between 2000 and 2010) of lakes (reservoirs) surveyed in this study

Lake	Area (km^2)	Number of sampling sites	Sampling period
Lake Poyang	3677.51	4	2003–2010
Lake Dongting	2630.78	10	2000–2010
Lake Taihu	2297.92	23	2000–2010
Lake Chao	769.32	12	2000–2010
Dangjiang reservior	609.34	3	2000–2010
Lake Liangzi	349.03	6	2001–2010
Lake Honghu	337.09	8	2001–2010
Lake Zhelin	248.74	4	2008–2010
Lake Shijiu	219.98	1	2006–2010
Lake Gehu	191.06	4	2000–2010
Lake Changhu	142.99	5	2001–2010
Lake Futou	141.19	2	2000–2010
Lake Changdang	119.04	4	2000–2010
Lake Yangcheng	116.57	7	2000–2010
Lake Shengjin	95.56	2	2003–2010
Zhanghe reservior	84.18	2	2001–2010
Lake Taiping	75.31	3	2000–2010
Lake Dazhi	73.53	5	2000–2010
Lake Dianshan	58.59	13	2000–2010
Xujiahe reservoir	55.97	2	2007–2010
Lake Baoan	45.26	2	2000–2010
Lake Tangxun	44.66	2	2000–2010
Lake Hougong	40.63	1	2001–2010
Lake Xiannv	39.45	4	2008–2010
Lushui reservior	39.19	4	2001–2010
Lake Gucheng	36.32	1	2006–2010
Lake Zhangdu	36.03	1	2000–2010
Lake Donghu	34.05	5	2000–2010
Bailianhe reservior	32.68	1	2000–2010
Fuqiaohe reservior	31.16	1	2000–2010
Lake Houhu	18.57	1	2000–2010
Dongpu reservior	15.53	2	2003–2010
Foziling reservoir	14.49	5	2001–2010
Chengxi reservoir	6.10	1	2000–2010
Qingshan reservoir	5.65	1	2000–2010
Lake Xuanwu	3.44	2	2000–2010
Lake Tianjing	1.41	3	2000–2010
Lake Guijia	1.09	3	2000–2010
Lake Yushan	0.68	10	2000–2010
Lake Nanhu	0.53	1	2001–2010
Huanglong reservior	0.32	2	2000–2010
In total	12,740.96	173	

References

1. Ailliot, P., & Monbet, V. (2012). Markov-switching autoregressive models for wind time series. *Environmental Modelling and Software, 30*, 92–101.
2. Baasch, A., Tischew, S., & Bruelheide, H. (2010). Twelve years of succession on sandy substrates in a post-mining landscape: a Markov chain analysis. *Ecological Applications, 20*, 1136–1147.
3. Balzter, H. (2000). Markov chain models for vegetation dynamics. *Ecological Modelling, 126*, 139–154.
4. Burns, N., McIntosh, J., & Scholes, P. (2009). Managing the lakes of the Rotorua district, New Zealand. *Lake and Reservoir Management, 25*, 284–296.
5. Carlson, R. E. (1977). A trophic state index for lakes. *Limnology and Oceanography, 22*, 361–369.
6. Chuang, C., Lin, C., Chien, C., & Chou, W. (2011). Application of Markov chain model for vegetation restoration assessment at landslide areas caused by a catastrophic earthquake in Central Taiwan. *Ecological Modelling, 222*, 835–845.
7. Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., Lancelot, C., & Likens, G. E. (2009). Controlling eutrophication: nitrogen and phosphorus. *Science, 323*, 1014–1015.
8. Duan, H., Zhang, Y., Zhang, B., Song, K., & Wang, Z. (2007). Assessment of chlorophyll *a* concentration and trophic state for Lake Chagan using Landsat TM and field spectral data. *Environmental Monitoring and Assessment, 129*, 295–308.
9. Forrest Hill, M., Witman, J. D., & Caswell, H. (2004). Markov chain analysis of succession in a rocky subtidal community. *The American Naturalist, 164*, 46–61.
10. Gao, J., Jiang, Z., Dou, H., et al. (2011). *Conservation and development of China's five largest freshwater lakes* (1st ed.). Beijing: Science Press **(In Chinese)**.
11. Gemmer, M., Jiang, T., Su, B., & Kundzewicz, Z. W. (2008). Seasonal precipitation changes in the wet season and their influence on flood/drought hazards in the Yangtze River Basin, China. *Quaternary International, 186*, 12–21.
12. Guan, D., Gao, W., Watari, K., & Fukahori, H. (2008). Land use change of Kitakyushu based on landscape ecology and Markov model. *Journal of Geographical Sciences, 18*, 455–468.
13. Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., & Hokao, K. (2011). Modeling urban land use change by the integration of cellular automaton and Markov model. *Ecological Modelling, 222*, 3761–3772.
14. Hilt, S., Köhler, J., Kozerski, H.-P., van Nes, E. H., & Scheffer, M. (2011). Abrupt regime shifts in space and time along rivers and connected lake systems. *Oikos, 120*, 766–775.
15. Huang, J., Gao, J., & Hörmann, G. (2012). Hydrodynamic-phytoplankton model for short-term forecasts of phytoplankton in Lake Taihu, China. *Limnologica, 42*(1), 7–18.
16. Huang, J., Gao, J., Hörmann, G., & Fohrer, N. (2014). Modeling the effects of environmental variables on short-term spatial changes in phytoplankton biomass in a large shallow lake, Lake Taihu. *Environmental Earth Sciences, 72*(9), 3609–3621.
17. Huang, J., Gao, J., Zhang, Y., & Xu, Y. (2015). Modeling impacts of water transfers on alleviation of phytoplankton aggregation in Lake Taihu. *Journal of Hydroinformatics, 17*(1), 149–162.
18. Jin, X., Liu, H., Tu, Q., Zhang, Z., & Zhu, X. (1990). *Lake eutrophication in China* (1st ed.). Beijing: China Environmental Science Press.
19. Jin, X., Xu, Q., & Huang, C. (2005). Current status and future tendency of lake eutrophication in China. *Science in China. Series C, Life Sciences, 48*, 948–954.
20. Johnson, J. C., Luczkovich, J. J., Borgatti, S. P., & Snijders, T. A. B. (2009). Using social network analysis tools in ecology: Markov process transition models applied to the seasonal trophic network dynamics of the Chesapeake Bay. *Ecological Modelling, 220*, 3133–3140.
21. Jørgensen, S. E. (1976). A eutrophication model for a lake. *Ecological Modelling, 2*, 147–165.
22. Jørgensen, S. E. (2010). A review of recent developments in lake modelling. *Ecological Modelling, 221*, 689–692.
23. Khan, F., & Ansari, A. (2005). Eutrophication: an ecological vision. *The Botanical Review, 71*, 449–482.
24. 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*, 662–668.
25. Leguerrier, D., Bacher, C., Benoît, E., & Niquil, N. (2006). A probabilistic approach of flow-balanced network based on Markov chains. *Ecological Modelling, 193*, 295–314.
26. Li, Y., Tang, C., Wang, C., Anim, D. O., Yu, Z., & Acharya, K. (2013). Improved Yangtze River diversions: are they helping to solve algal bloom problems in Lake Taihu, China? *Ecological Engineering, 51*, 104–116.
27. Link, W. A., Cam, E., Nichols, J. D., & Cooch, E. G. (2002). Of BUGS and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. *Jornal of Wildlife Management, 66*, 277–291.
28. Liou, Y., & Lo, S. (2005). A fuzzy index model for trophic status evaluation of reservoir waters. *Water Research, 39*, 1415–1423.
29. Liu, W., Zhang, Q., & Liu, G. (2010). Lake eutrophication associated with geographic location, lake morphology and climate in China. *Hydrobiologia, 644*, 289–299.
30. Logofet, D. O., & Lesnaya, E. V. (2000). The mathematics of Markov models: what Markov chains can really predict in forest successions. *Ecological Modelling, 126*, 285–298.
31. Ma, R., Yang, G., Duan, H., Jiang, J., Wang, S., Feng, X., Li, A., Kong, F., Xue, B., Wu, J., & Li, S. (2011). China's lakes at present: number, area and spatial distribution. *Science China Earth Sciences, 54*, 283–289.
32. Malmaeus, J. M., & Hakanson, L. (2004). Development of a lake eutrophication model. *Ecological Modelling, 171*, 35–63.
33. Mooij, W. M., Trolle, D., Jeppesen, E., Arhonditsis, G., Belolipetsky, P. V., Chitamwebwa, D. B. R., Degermendzhy, A. G., DeAngelis, D. L., De Senerpont Domis, L. N., Downing, A. S., Elliott, J. A., Fragoso, C. R., Jr., Gaedke, U., Genova, S. N., Gulati, R. D., Håkanson, L., Hamilton, D. P., Hipsey, M. R., 't Hoen, J., Hülsmann, S., Los, F. H., Makler-Pick, V., Petzoldt, T., Prokopkin, I. G., Rinke, K., Schep, S. A., Tominaga, K., Van Dam, A. A., Van Nes, E. H., Wells, S. A., & Janse, J. H. (2010). Challenges and opportunities for integrating lake ecosystem modelling approaches. *Aquatic Ecology, 44*, 633–667.
34. Müller, B., Lotter, A. F., Sturm, M., & Ammann, A. (1998). Influence of catchment quality and altitude on the water and sediment composition of 68 small lakes in Central Europe. *Aquatic Sciences, 60*, 316–337.
35. OECD. (1982). *Eutrophication of waters: monitoring, assessment and control*. Paris: OECD.
36. Schapaugh, A. W., & Tyre, A. J. (2013). Accounting for parametric uncertainty in Markov decision processes. *Ecological Modelling, 254*, 15–21.
37. Smith, V. (2003). Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research, 10*, 126–139.
38. Strigul, N., Florescu, I., Welden, A. R., & Michalczewski, F. (2012). Modelling of forest stand dynamics using Markov chains. *Environmental Modelling and Software, 31*, 64–75.
39. Strobl, R. O., Forte, F., & Pennetta, L. (2007). Application of artificial neural networks for classifying lake eutrophication status. *Lake and Reservoir Management, 12*, 15–25.

40. Verkhozina, V. A., Kozhova, O. M., & Kusner, Y. S. (2000). Hydrodynamics as a limiting factor in the Lake Baikal ecosystem. *Aquatic Ecosystem Health and Management, 3*, 203–210.
41. Walker, W. W. (1979). Use of hypolimnetic oxygen depletion rate as a trophic state index for lakes. *Water Resources Research, 15*, 1463–1470.
42. Wu, N., Schmalz, B., & Fohrer, N. (2011). Distribution of phytoplankton in a German lowland river in relation to environmental factors. *Journal of Plankton Research, 33*, 807–820.
43. WWF. (2013). http://wwf.panda.org/about_our_earth/ecoregions/ecoregion_list/ (accessed on September 30, 2013).
44. Xu, F., Tao, S., Dawson, R. W., & Li, B. (2001). A GIS-based method of lake eutrophication assessment. *Ecological Modelling, 144*, 231–244.
45. Xu, F., Wang, J., Chen, B., Qin, N., Wu, W., He, W., He, Q., & Wang, Y. (2011). The variations of exergies and structural exergies along eutrophication gradients in Chinese and Italian lakes. *Ecological Modelling, 222*, 337–350.
46. Yang, X., Wu, X., Hao, H., & He, Z. (2008). Mechanisms and assessment of water eutrophication. *Journal of Zhejiang University. Science. B, 9*, 197–209.