WFL Publisher
Science and Technology
Meri-Rastilantie 3 B, FI-00980
Helsinki, Finland
e-mail: info@world-food.net

Journal of Food, Agriculture & Environment Vol.10 (2): 1245-1247. 2012

www.world-food.net

Analysis of lake vulnerability in the middle-lower Yangtze River catchment using C4.5 decision tree algorithms

Feng Wu [1], Xiangzheng Deng [1, 2]*, Qun'ou Jiang [1, 3], Jinyan Zhan [4] and Dongdong Liu [5]

*[1] Institute of Geographical Sciences and National Resources Research, China Academy of Science, Beijing 100101, China. [2] Center for Chinese Agricultural Policy, CAS, Beijing 100101, China. [3] Graduate University of Chinese Academy of Sciences, Beijing 100049, China. [4] State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, China. [5] School of Mathematics and Physics, China University of Geosciences, Wuhan, 430074, China. *e-mail: dengxz.ccap@igsnrr.ac.cn, zhanjy@bnu.edu.cn*

Received 4 February 2012, accepted 30 April 2012.

Abstract

Lake vulnerability is one of the significant issues in China, more and more lakes are polluted seriously. This paper comprehensively analyzes the vulnerable degree of lakes located in the middle-lower Yangtze River catchment. Firstly, the factors influencing lake vulnerability are explored. Then based on the references and materials collected, this study used the weight data and grading data to assess the vulnerable degree of lakes. Thirdly, the area percentage of impervious surface was employed to represent the intensity of human activities. Lastly, C4.5 decision tree algorithms were applied to get classification matrix which can be used to estimate the categories of all the lakes in the middle-lower Yangtze River catchment. The results show that lakes under rational exploitation accounted for 61.8%, and those over-exploited accounted only for 13.2%. As for the lake vulnerability, it can be deduced that low vulnerable lakes accounted for 32.4%, medium vulnerable lakes had the highest percentage with 45.6%, and accordingly, the amount of lakes with high vulnerability was lowest, only 22.1%. According to the vulnerability and exploited degree, the lakes were divided into three categories. Generally speaking, lakes within I category were rare, and amounts of lakes within II category were much more than other categories. Lakes within III category were affected by human activity mostly. The conclusions can provide significant information for policy formulation to mitigate lake vulnerability and restore lake function.

Key words: Lake vulnerability, middle-lower Yangtze River catchment, C4.5 decision tree algorithms.

Introduction

Lake pollution and vulnerability has become a global concerned problem. Since 1960's, plenty of countries began to be bothered by the problem of lake pollution, they all try to explore one effective method to prevent lake pollution and reduce the vulnerability [1, 2]. Historically speaking, most of the developed countries have experienced the tortuous process of "pollute first, clean up later" [3, 4]. Although the mitigation result for lake pollution is remarkable, it is on the expenses of lengthy and costly efforts. How to extend effective control and treatment for the numerous and various polluted lakes is urgent to be solved [3].

The lakes located in the middle-lower Yangtze River catchment, most of which are shallow and connected with river, are affected mostly by human activity, and their vulnerable degree is more concerned by experts and government. All of the five freshwater lakes are located in this region, so the environmental quality is more significant. According to field survey data, except Dongting Lake, the rests are all polluted to some extent [5, 6]. Amounts of previous researches have focused on lake eutrophication and vulnerability, and various pollution mitigation strategies have been proposed in China. However, most of those researches focus basically on water quality as far as the indexes of TP, TN, chlorophyll a, transparency are concerned. With less attention to human activities around the lake basin, it is quiet difficult to comprehensively assess the vulnerability of the lakes, and it is also not conducive to the classification of mitigation and management of lake pollution.

Case Study Area

Middle-lower Yangtze River catchment is one of the world's few regions congested with large shallow lakes. According to statistics, there are more than 70 lakes with the area bigger than 1 km^2, and the total area is 16,558km^2, accounting for 60% of the total area of freshwater lake in China. Due to the limits of data availability, this paper collects the data of 61 lakes. Among which, 11 lakes suffer from medium eutrophication (18%), 37 lakes are threatened by high eutrophication (60.7%), and 13 lakes are endangered by extreme high eutrophication (21.3%).

Theoretical Approaches

In recent years, amounts of classification methods are used in lake researches, such as fuzzy mathematical method, Gray correlation method, genetic algorithms and so on [7-10]. All of the methods have their own advantages and disadvantages, and the assessment results are diverse in the same region. C4.5 decision tree algorithm method is good at classification using information entropy of training data set as well as the sampling from unknown network flow [10]. It is beneficial to lake classification especially under incomplete data, which is helpful for our study.

Based on the analysis of lake vulnerability, this study selected the indicator index to represent the lake comprehensive vulnerability, and collected the relative data of more than 60 lakes in middle-lower Yangtze River catchment. According to the grading standard, this study estimated the comprehensive vulnerability of lakes in research area to indicate their resilience to pollution. On the other side, based on remote sensing technologies, we generated the percentage of impervious surface to describe the intensity of human activities. This paper utilized lake vulnerability and exploited degree to construct the matrix classification system, and then divided all of the lakes into different categories using C4.5 decision tree algorithms.

Data Processing

Lake vulnerability is impacted comprehensively by climate, regional geographical and hydrological characters, and human activities, all of which are influential to lake vulnerability. Combined with the principles of index selection, this study chose lake area, lake hydraulic retention time, stratification, characters of lakeside zone and lake eutrophication as the quantitative index of lake vulnerability.

Grading Standard and Weight

Two methods, which include expert assessment method and correlation analysis, were involved to estimate the weight of lake vulnerability of each index. Expert assessment method is a priori method, and correlation analysis is a posteriori method. Combining the results calculated by both methods, this paper ultimately calculated the grading standard and weight of each index (Table 1).

Indicator Index of Human Activity Intensity

At present, human activities have a significant impact on lake vulnerability. This study used impervious surface to represent the human activity. Impervious surfaces (IS) is an artificial surface in a built-up area. Due to its special features, it is considered to be an important factor in regional water quality deterioration. Meanwhile impervious surface changes can also influence non-point source pollution load. Carrying large amounts of sediments, nutrient pollutants, heavy metals, runoff of urban river could exacerbate non-point source pollution. Therefore, we defined the proportion of impervious surface area and lake drainage area as the quantitative indicator which stands for the impact of human activity.

Results and Discussion

Applying the above mentioned methodology, we calculated the vulnerable degree of the lakes and the exploited degree of lake area based on the sample data set with no default data. Calculated results are shown in Table 2. We divided middle-lower Yangtze River catchment into 9 regions.

We calculated the vulnerable degree as well as the exploited degree of lakes using statistical data and weight processing, and classified lake vulnerability using C4.5 decision tree algorithm.

The estimation results show that the number of lakes under rational exploitation takes up the highest percentage with 61.8%, and medium exploited lakes account for 25%, and the over exploited lakes take up 13.2%. It is clear that most of lakes in middle-lower Yangtze River catchment are affected by human activities. However, the degree is still not serious.

As for the vulnerability of lakes, it can be deduced that the lakes with low vulnerability take up 32.4%, medium vulnerable lakes have highest percentage with 45.6%, and on contrary, the number of lakes with high vulnerability is lowest, only 22.1%. It can be concluded that the distribution of lake vulnerability in middle-lower Yangtze River catchment is even, which is suitable for pollution mitigation according to classification.

C4.5 decision tree algorithms were applied to classify all the lakes into three categories. The lake under rational exploition and

Table 1. Lake vulnerability index and grading standard.

Index	Definition	Value	Unit	Grade	Note
Lake area	It is more sensitive to water quality for the small area lake	1-10	km^2	1	Very small lake/pool
		10-100		2	Small lake
		100-500		3	Medium lake
		500+		4	Large lake
Lake hydraulic retention time	The shorter the hydraulic retention time, the lower the cumulative pollutant concentration needed to be diluted	90+	Day	1	Residence period is very long
		30-90		2	Residence period is long
		15-30		3	Residence period is short
		1-15		4	Residence period is very short
Stratification	The lake with more obvious stratification is more sensitive to external P source	30.0+	metre	1	Obvious stratification
		13.5-30.0		2	Normal stratification
		4.5-13.5		3	Weak stratification
		0-4.5		4	No stratification
Characters of lakeside zone	The lake with larger slope lakeside is more sensitive to erosion	75-100%	Lakeside percentage within 300	1	Lakeside percentage with D slope（Slope is 12% or more）
		50-75%		2	
		25-50%		3	
		0-25%		4	
Trophic status index	Assessment of integrated eutrophication based on TN\TP\SD\chl-a	0-30	No	1	Low nutrition
		30-50		2	Medium nutrition
		50-70		3	High nutrition
		70+		4	Extreme high nutrition

Table 2. The results of estimation for sample lakes.

Lake name	Comprehensive vulnerable degree	Human activity intensity	Lake name	Comprehensive vulnerable degree	Human activity intensity
Poyang Lake	12.35	3.40%	Gaoyou lake	29.71	7.84%
Dongting Lake	12.68	2.77%	Red Lake	40.47	3.55%
Lake Tai	23.73	35.38%	Baoying Lake	52.96	8.11%
Chao Lake	43.88	10.44%	Kuncheng Lake	47.92	20.20%
Shengjin Lake	36.43	1.92%	Kuilei Lake	22.21	20.98%
Dianshan Lake	64.10	19.81%	Hongze Lake	59.73	15.26%
Wuchang Lake	36.24	7.97%	Yaogang Lake	54.38	2.85%

with low vulnerability was defined as category I, medium and high vulnerable area within medium exploited area was considered as category II and over exploited lake defined as category III. According to the results, fewer lakes belonged to category I including Hong Lake, Liangzi Lake, Junshan Lake and so on. The lakes within category II consisted of most of the lakes. The lakes within category III were Dianshan Lake, Kuncheng Lake, Baima Lake, and so on. Although the vulnerability degree is various within category III, all of the lakes are most affected by human activity.

Conclusions

This study estimated the lake vulnerability and exploition level, and classified lakes based on the lake vulnerability in middle-lower Yangtze River catchment using C4.5 decision tree algorithm. Results indicated that nearly two-thirds of the lakes' exploited level were comparatively low, one-third relatively higher. For the lakes which are over exploited, governments should formulate appropriate policies to increase environment awareness and reduce the effects of human activities on the Middle-lowerYangtze River catchment. As for the lake with medium and high vulnerability, government should focus on the lake restoration and mitigation. For the lake with low vulnerability, local government should pay much attention to combine environmental protection and ecological restoration and control pollutants emissions.

Acknowledgements

This research was supported by the State Water Pollution Control and Management of Major Science and Technology (2009ZX07106-001); the National Key Programme for Developing Basic Science (2010CB950904) and the National Scientific Foundation of China (70873118; 41171434). Data support from the projected funded by the Ministry of Science and Technology of China (2010GXS5B163; 2008BAC43B01; 2008BAK50B05; 2008BAK50B06) are also greatly appreciated.

References

[1]Wu, M., Liu, Z. Y. and Li, Z. Y. 2008. Evaluation of lake eutrophication based on Bayes method. Journal of Chendu University of Information Technology **23**(5):533-536.

[2]Li, Z. Q., Wang, S. L. and Yang, M. 2010. Lake protection and utilization based on ecological sensitivity analysis - A case study of Futou Lake in Hubei Province **19**(6):714-718.

[3]Zhang, Y., Wang, Y. and Ye, W. H. 2002. Surface water quality changes in Shanghai City in past 20 years. Acta Scientiae Circumstantiae **22**(2):247-251.

[4]Wu, H. J. and Guo, S. L. 2001. The effect of hydrological regime on phytoplankton community. Advances in Water Science **12**(1):52-55.

[5]Zhang, Y. Z., Zheng, B. H. and Fu, G. 2006. On the assessment methodology and standards for nutrition status in channel type reservoirs based on zoning of eutrophication sensitivity. Acta Scientiae Circumstantiae **26**(6):1016-1021.

[6]Pu, D. W. 2010. Simulation of flow rate and energy change in different observation site in Namucuo Lake. Tibet Science and Technology **10**:27-31.

[7]Lathrop, R. C. and Lillie, R. A. 1980. Thermal stratification of Wisconsin lakes. Transactions of the Wisconsin Academy of Sciences, Arts and Letters **68**:90-96.

[8]Jin, D. Y., Liu, B. and Du, J. D. 2004. Present status of lakeshore and ecological restoration of east part of Lake Erhai, Yunnan Province. Research of Environmental Sciences **17**:80-85.

[9]Liu, J. M. and Deng, X. Z. 2011. Influence of different land use on urban microenvironment in Beijing City, China. Journal of Food, Agriculture & Environment **9**(3&4):1005-1011.

[10]Wang, M. C., Liu, X. Q. and Zhang, J. H. 2002. Evaluate method and classification standard on lake eutrophication. Environmental Monitoring in China **18**(5):47-49.