

Study on water pollution change rule and source in Taihu Lake Basin

Song wei

Tianjin Research Institute for Water Transport Engineering, M.O.T., Tianjin 300456, China

Email:s56wei@gmail.com

Abstract: The main pollutants of water environment in Taihu (COD, NH₃-N and DO) are studied in this paper. The statistical analysis characteristics of the weekly mean concentration of pollutants in seven automatic monitoring stations in Taihu Lake Basin were studied by statistical methods, and the correlation between the meteorological data of seven stations and pollutants was analyzed. The results show that, the normal distribution is the water quality index of each site, and the fit type selection is most consistent with the significant relationship. DO weekly mean concentrations were better correlated at several sites except for individual sites, CODMn weeks mean concentration, the correlation between the sites is very small, the correlation between NH₃-N weeks mean concentration at all sites showed no correlation or low linear correlation, These results shows that the DO pollution sources of these stations have some relevance, but the pollution sources of CODMn and NH₃-N are not related closely, and only a few sites are related.

1. Introduction

In recent years, great achievements have been made in China's economic construction, but the deterioration of the ecological environment is accompanied by economic growth, among which water related problems are the most prominent. It is urgent to strengthen the prevention and control of water pollution, to reduce the harm to human health and the environment caused by water pollution, and to formulate effective control strategies for pollutants[1]. The cause and treatment of water pollution in Taihu has always been a hot issue. Yin Yan,Zhang Yunlin and so on [2][3] have successively studied the temporal and spatial distribution of COD and the correlation with other elements in the northern Taihu Lake region and the whole lake area, Yang Dingtian and Chen Weimin [4] analyzed the temporal and spatial distribution of COD and its related factors in Meiliang Bay,Taihu, Xie Rongrong and so on[5] have studied the emission changes of COD and NH₃-N in Taihu, and have determined the total reduction rate of these two pollutants. In this paper, the main pollutants of water environment in Taihu are taken as the object of study. By analyzing the changing trend and correlation of various pollutants evaluation indexes, the trend of the function of some pollution indexes is obtained, and based on the analysis of the sources of the major pollutants, the relationship between the various influencing factors is found, which provides the research and policy basis for controlling and reducing pollution at the source.

2. Research Area and Data

2.1 Research Area

Taihu called Zhenze, also known as wuhu, Located at longitude 119, 52 ', 32', 120, 36 ', 10', latitude 30,



55 ' , 40' - 31 ' , 32' 58', Is China's third largest freshwater lake, Lake area of 2338km², drainage area of 36985km². there are 48 islands and 72 peaks. The Taihu regional map is shown in figure 1.

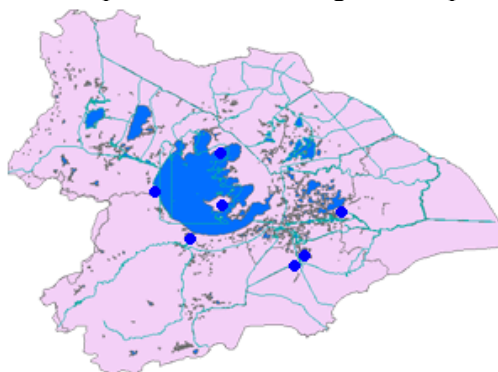


Figure 1 Taihu lake regional map

2.2 Research Data

The scope of this research is Jiangsu Wuxi Shazhu the water quality automatic monitoring station, Jiangsu Yixing Lanshanzui water quality automatic monitoring station, Jiangsu Suzhou Xishan water quality automatic monitoring station, Zhejiang Huzhou Xintang harbor water quality automatic monitoring station, Shanghai Qingpu Jishui harbor water quality automatic monitoring station, Zhejiang Jiaxing Wangjiangjing River water quality automatic monitoring station and Zhejiang Jiaxing Xielu harbor water quality automatic monitoring station. there are seven key water quality monitoring stations of the Taihu Lake basin. Table 1 gives the latitude, longitude and cross section information of the seven monitoring stations.

Table 1 location of seven water quality automatic monitoring stations in Taihu Lake Basin

Point name	Point latitude and longitude	Rivers
Jiangsu Wuxi Shazhu monitoring station	East longitude 120° 13' 46" north latitude 31° 23' 58"	Taihu
Jiangsu Yixing Lanshanzui monitoring station	East longitude 119° 54' 43" north latitude 31° 12' 56"	Taihu
Jiangsu Suzhou Xishan monitoring station	East longitude 120° 12' 0" north latitude 31° 13' 0"	Taihu
Zhejiang Huzhou Xintang harbor monitoring station	East longitude 120° 50' 20" north latitude 30° 50' 50"	Xintang harbor river
Shanghai Qingpu Jishui harbor monitoring station	East longitude 120° 54' 06" north latitude 31° 06' 56"	Jishui harbor river
Zhejiang Jiaxing Wangjiangjing River monitoring station	East longitude 120° 42' 31" north latitude 3° 53' 06"	Beijing Hangzhou the Grande Canale
Zhejiang Jiaxing Xielu harbor monitoring station.	East longitude 120° 42' 31" north latitude 30° 53' 06"	Xielu harbor river

3. Statistical distribution characteristics of water quality indicators in Taihu Lake Basin

3.1 Determine the weekly concentration optimal statistical distribution model of water quality evaluation index

The basic normal distribution, uniform distribution, Poisson distribution and exponential distribution are used in the statistical distribution model, the weekly concentration data of each site in 2011 were analyzed by representative data, the optimum distribution model of three water quality indexes (mainly DO, CODMn and NH₃-N) was determined, and two kinds of goodness of fit test methods, KS test and Binomial test, were used to test the accuracy.

The fit test results about the index concentration of Jiangsu Wuxi Shazhu the water quality automatic monitoring stations in Taihu basin as shown in table 2~4.

Table 2 in Wuxi, Jiangsu Sha Zhu automatic water quality monitoring stations DO concentration goodness of fit test results in 2011

SN	Original hypothesis	Test method	Sig.	conclusion
1	The categories defined by DO $\leq$ 11.94 and >11.94 occur with probabilities of 0.5 and 0.5	Single sample Binomial test	0.000	Reject null hypothesis
2	The distribution of DO is normal, with an average value of 9.56 and a standard deviation of 2.74	Single sample KS test	0.230	Keep null hypothesis
3	The distribution of DO is uniform, with a minimum value of 5.87 and a maximum value of 18	Single sample KS test	0.000	Reject null hypothesis
4	The distribution of DO is Poisson, with an average value of 0.00	Single sample KS test	—	beyond computation
5	The distribution of DO is exponential, with an average value of 9.56	Single sample KS test	0.000	Reject null hypothesis

Table 3 in Wuxi, Jiangsu Sha Zhu automatic water quality monitoring stations CODMn concentration goodness of fit test results in 2011

SN	Original hypothesis	Test method	Sig.	conclusion
1	The categories defined by CODMn $\leq$ 3.75 and >3.75 occur with probabilities of 0.5 and 0.5	Single sample Binomial test	0.127	Keep null hypothesis
2	The distribution of CODMn is normal, with an average value of 3.60 and a standard deviation of 0.58	Single sample KS test	0.530	Keep null hypothesis
3	The distribution of CODMn is uniform, with a minimum value of 2.60 and a maximum value of 4.90	Single sample KS test	0.097	Keep null hypothesis
4	The distribution of CODMn is Poisson, with an average value of 0.00	Single sample KS test	—	beyond computation
5	The distribution of CODMn is exponential, with an average value of 3.60	Single sample KS test	0.000	Reject null hypothesis

Table 4 in Wuxi, Jiangsu Sha Zhu automatic water quality monitoring stations NH₃-N concentration goodness of fit test results in 2011

SN	Original hypothesis	Test method	Sig.	conclusion
1	The categories defined by NH ₃ -N $\leq$ 0.28 and >0.28 occur with probabilities of 0.5 and 0.5	Single sample Binomial test	0.000	Reject null hypothesis
2	The distribution of NH ₃ -N is normal, with an average value of 0.24 and a standard deviation of 0.09	Single sample KS test	0.265	Keep null hypothesis

3	The distribution of NH ₃ -N is uniform, with a minimum value of 0.05 and a maximum value of 0.51	Single sample KS test	0.000	Reject null hypothesis
4	The distribution of NH ₃ -N is Poisson, with an average value of 0.00	Single sample KS test	—	beyond computation
5	The distribution of NH ₃ -N is exponential, with an average value of 0.24	Single sample KS test	0.000	Reject null hypothesis

Further analysis of the other six sites, the analysis of significant differences, it is worth the water quality evaluation index, the best fitting distribution of concentration values are all normal distribution.

3.2 Parameter estimation of optimum concentration model of water quality index

Maximum likelihood (MLE) is used to estimate the statistical parameters of pollutant concentration under optimum distribution. In the 3.1 section, the concentration optimal distribution model of water quality evaluation parameters in Taihu basin has been established as normal distribution. The parameters of these distributions are estimated at a confidence level of 95%, i.e., the level of significance is 5%. This shows that the parameter value exceeds the given value with a probability of only 5%. Table 5 to 7 respectively estimates the estimated parameters of the best concentration model for the water quality evaluation parameters of the Taihu Lake Basin using maximum likelihood.

Table 5 The site of the Taihu Basin DO weekly concentration under the the normal distribution parameter estimation in 2011

Site number	Mean value in distribution	Standard deviation	Sig	conclusion
1	9.56	2.74	0.230	Keep null hypothesis
2	7.80	1.95	0.850	Keep null hypothesis
3	9.56	2.74	0.230	Keep null hypothesis
4	9.17	3.11	0.125	Keep null hypothesis
5	6.88	2.23	0.045	Reject null hypothesis
6	4.46	0.97	0.913	Keep null hypothesis
7	4.49	1.71	0.889	Keep null hypothesis

As shown in table 5, the DO concentrations in seven sites of the Taihu Basin is vary greatly. Since the concentration of DO in the water quality automatic monitoring station of Qingpu Shanghai Jishui harbor is higher than that of 5~6mg/L, the normal fitting can not agree with the significant level of 5%. The asymptotic significance value is 0.045, close to 0.05, which can basically satisfy the normal distribution. Except for the water quality automatic monitoring station of Qingpu Shanghai Jishui harbor, the normal fitting of the other six stations is good, that is, the difference assumption is completely established.

Table 6 The site of the Taihu Basin CODMn weekly concentration under the the normal distribution

Site number	Mean value in distribution	Standard deviation	Sig	conclusion
1	—	—	—	Keep null hypothesis
2	—	—	—	Reject null hypothesis
3	—	—	—	Keep null hypothesis
4	—	—	—	Keep null hypothesis
5	—	—	—	Keep null hypothesis
6	—	—	—	Keep null hypothesis

—	— —	— —	—	Keep null hypothesis
—	— —	—	— —	Keep null hypothesis

As shown in table 6, under the condition that the concentration of CODMn in the water quality automatic monitoring station of Yixing Lanshanzui in Jiangsu was 0.05 in the significant level, reject null hypothesis, that is to say, the concentration difference is not obvious at this significance level, but the asymptotic significance value is close to 0.042 and 0.05, and the other distribution fitting is far less than the normal distribution. The best fitted normal distribution is fitted.

Table 7 the site of the Taihu Basin NH₃-N weekly concentration under the the normal distribution parameter estimation in 2011

Site number	Mean value in distribution	Standard deviation	Sig	conclusion
1	—	—	—	Keep null hypothesis
2	—	—	—	Keep null hypothesis
3	—	—	—	Keep null hypothesis
4	—	—	—	Keep null hypothesis
5	— —	—	—	Keep null hypothesis
6	— —	—	—	Keep null hypothesis
7	— —	—	—	Keep null hypothesis

As shown in table 7, compared with the concentration of the first two groups, the average concentration of NH₃-N is small and the standard deviation is small, the asymptotic significance values are not less than 0.05, that is, under normal distribution, under the condition that the significance level is 0.05, the NH₃-N concentration data can be tested by significance test.

4. Variation trend of water quality index concentration in Taihu Basin

4.1 Characteristics of weekly variation of water quality in Taihu Basin

This paper mainly identifies three water quality indicators (mainly DO, CODMn and NH₃-N), reference <Environmental quality standards for surface water>(GB 3838-2002)[7] form State Bureau of environmental protection.

In addition to the emission sources, the concentration index of water pollution is also affected by meteorological conditions such as temperature, rainfall and so on. Therefore, in a certain area, the concentration of water quality index shows a tendency to change with time. Tables 8 to 10 describe the basic statistics of concentration values from December 27, 2010 to December 25, 2011 (the year of 2011).

Table 8 the site of taihu lake watershed DO weeks are concentration of basic statistics in 2011

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
average	9.564	7.795	7.096	9.170	6.875	4.464	4.490
standard error	0.381	0.271	0.213	0.432	0.309	0.135	0.237
median	8.690	7.740	6.645	7.980	6.185	4.460	4.435
mode	10.500	9.890	8.110	13.500	6.560	4.690	6.020
standard deviation	2.744	1.952	1.537	3.114	2.225	0.974	1.709
variance	7.531	3.812	2.362	9.695	4.951	0.949	2.921

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
kurtosis	0.930	0.274	0.022	-0.621	-0.257	0.858	-0.259
skewness	1.142	-0.279	0.761	0.544	0.790	-0.134	0.186
region	12.130	9.730	5.980	12.690	9.070	4.840	7.210
minimum value	5.870	2.470	4.820	3.910	3.330	1.860	1.540
Maximum value	18.000	12.200	10.800	16.600	12.400	6.700	8.750
sum	497.330	405.350	368.980	476.830	357.500	232.150	233.470
count	52	52	52	52	52	52	52

Notes: the names of each site are described in Section 3.2.

Table 9 the site of taihu lake watershed CODMn weeks are concentration of basic statistics in 2011

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
average	3.596	4.054	4.035	4.588	3.998	6.165	7.956
standard error	0.081	0.206	0.125	0.163	0.087	0.183	0.071
median	3.500	3.700	4.000	4.800	4.000	6.200	7.900
mode	3.300	3.500	4.300	4.800	4.000	6.600	7.600
standard deviation	0.585	1.487	0.899	1.153	0.625	1.319	0.515
variance	0.342	2.210	0.808	1.329	0.391	1.739	0.265
kurtosis	-0.803	3.118	-0.232	.523	0.596	0.314	0.549
skewness	0.250	1.688	-0.001	-0.688	0.297	-0.493	0.213
region	2.300	7.000	4.200	5.300	3.200	6.200	2.600
minimum value	2.600	1.800	1.800	1.700	2.500	2.400	6.900
Maximum value	4.900	8.800	6.000	7.000	5.700	8.600	9.500
sum	187.000	210.800	209.800	229.400	207.900	320.600	413.700
count	52	52	52	50	52	52	52

Table 10 the site of taihu lake watershed NH3-N weeks are concentration of basic statistics in 2011

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
average	0.012	0.836	0.161	0.431	1.769	1.617	1.363
standard error	0.230	0.079	0.011	0.026	0.079	0.088	0.092
median	0.260	0.765	0.140	0.410	1.755	1.705	1.415
mode	0.088	0.900	0.120	0.230	2.240	2.120	1.340
standard deviation	0.008	0.573	0.079	0.188	0.571	0.632	0.662
variance	2.231	0.328	0.006	0.035	0.326	0.400	0.438
kurtosis	1.086	2.544	4.910	-0.611	0.609	-1.097	-0.925
skewness	0.460	1.356	2.010	0.375	0.730	-0.279	0.120
region	0.050	2.830	0.410	0.820	2.720	2.370	2.360
minimum value	0.510	0.060	0.060	0.060	0.860	0.370	0.220
Maximum value	12.590	2.890	0.470	0.880	3.580	2.740	2.580
sum	497.330	43.490	8.390	22.390	91.970	84.100	70.900
count	52	52	52	52	52	52	52

4.2 Correlation analysis with the concentration index of water quality in Taihu Basin

The correlation coefficient is an index describing the degree of dispersion between the two sets of

measurement variables. Used to determine whether changes in two sets of measurement variables are related, that is, whether a larger value of a variable is associated with a larger value of another variable (positive correlation) ; Or whether a smaller value of a variable is associated with a larger value of another variable(negative correlation) ; Or the values in the two variables are not related to each other (the correlation coefficient is approximately zero). Let (X, Y) be a binary random variable, then the correlation coefficient is expressed as :

$$r_{XY} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (1)$$

Formula : r_{XY} is a correlation coefficients of random variables X and Y. r is a numerical characteristic that measures the degree of linear correlation between X and Y in a random variable, Its range of value is [-1,1]. The greater the |r| value, the higher the degree of linear correlation between variables; The closer the |r| value is to 0, the lower the degree of linear correlation between variables. Generally can be divided into three grades : $|r| < 0.3$ is low linear correlation; $0.3 \leq |r| < 0.8$ is significant correlation; $0.8 \leq |r| < 1$ is highly linear correlation.

Tables 11 to 13 shows the correlation analysis of weekly mean concentration of Taihu Lake Basin in 2011.

Table 11 DO between sites related to the weekly average concentration of the Taihu Lake Analysis

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
Site 1	1.000						
Site 2	0.716	1.000					
Site 3	0.481	0.581	1.000				
Site 4	0.604	0.743	0.614	1.000			
Site 5	0.588	0.494	0.353	0.653	1.000		
Site 6	0.175	0.185	0.228	0.400	0.466	1.000	
Site 7	0.551	0.543	0.412	0.531	0.303	0.273	1.000

Table 12 CODMn between sites related to the weekly average concentration of the Taihu Lake Analysis

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
Site 1	1.000						
Site 2	0.238	1.000					
Site 3	-0.085	0.139	1.000				
Site 4	-0.109	0.102	0.141	1.000			
Site 5	-0.338	-0.167	-0.046	0.176	1.000		
Site 6	-0.246	0.017	0.001	-0.163	0.400	1.000	
Site 7	-0.379	-0.041	0.142	0.147	0.352	0.327	1.000

Table 13 NH3-N between sites related to the weekly average concentration of the Taihu Lake Analysis

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
Site 1	1.000						
Site 2	0.018	1.000					
Site 3	0.145	0.267	1.000				
Site 4	-0.356	-0.121	0.071	1.000			
Site 5	0.335	-0.081	0.081	0.012	1.000		
Site 6	0.287	0.227	0.172	-0.194	0.330	1.000	
Site 7	0.264	0.161	0.312	-0.305	0.357	0.771	1.000

You can see the data from tables 10 to 13: In addition to site 6 and site 1, 2, 3, 7 correlation is smaller, the other parts of the Taihu Lake Basin DO week mean concentration is better correlation, it

shows that these stations are related to the source of the change of dissolved oxygen. About CODMn, the correlation between each sites in the Taihu Lake basin was very small, only sites 5 and sites 1, 6, and 7 were included, with low linear correlations between site 6 and site 7. About NH₃-N, the correlations between each sites were uncorrelated or low linearly correlated, among them, the correlation between site 6 and site 7 was close to significance. Based on the previous analysis, it is concluded that the sources of ammonia nitrogen between site 6 and site 7 may be consistent, and bring about similar diffusion effects.

5. Conclusion

In this paper, statistical characteristics of the weekly mean concentrations of CODMn, NH₃-N and DO in the seven automatic monitoring stations of Taihu Lake Basin from December 27, 2010 to December 25, 2011 (2011 cycle) were studied by statistical methods, the correlation analysis of pollutants in the seven stations was carried out, and the main conclusions were as follows:

(1) On the whole, the normal distribution is fitted well for each site water quality index concentration, that most consistent with significant relationships. Although some data can satisfy other distribution types, but from the significance difference value, they all chose the normal distribution.

(2) The correlation analysis between the water quality indexes at each site, in addition to the site 6 and site 1, 2, 3, 7, they are less correlation, several other sites in Taihu basin between the weekly average concentrations of DO have good correlation, suggesting that these stations as a result of changes in dissolved oxygen sources have certain relevance. The correlation between CODMn and the stations in Taihu basin is very small. Only 5 and 1, 6 and 7 of site 6 and site 7 are low linear correlation. About CODMn, the correlation between the stations in the Taihu Lake basin is very small. Only between 5 and 1, 6, 7, and between site 6 and site 7 is a low linear correlation. About NH₃-N, The correlation between sites showed no correlation or low linear correlation, specifically, the site 7 and site 3, 4, 5, 6, site 1 and site 4, 5, site 3 and site 7, site 5 and site 6 is low linear correlation, among them, the correlation between site 6 and site 7 was close to significance. It is considered that the source of ammonia nitrogen between site 6 and site 7 may be consistent and bring about similar diffusion effects.

References

- [1] Strategy research group of sustainable development of Chinese Academy of Sciences. China Sustainable Development Study I in 2007[M]. Beijing: Science Press, 2012.
- [2] Zhang Yunlin, Yang Longyuan, Qin Boqiang, et al. Spatial Distribution of COD and the Correlations with Other Parameters in the Northern Region of Lake Taihu [J]. Environmental Science, 2008, 29(6): 1457-1462.
- [3] Yin Yan, Zhang Yunlin, Shi Zhiqiang, et al. Temporal-Spatial Variations of Chemical Oxygen Demand and Biochemical Oxygen Demand in Lake Taihu [J]. Acta Scientiae Circumstantiae, 2010, 30(12): 2544-2552.
- [4] Yang Dingtian, Chen Weimin. Analysis of temporal and spatial distribution and related factors of CODMn in Meiliang Bay of Taihu lake [J]. 2007 Chinese Sustainable Development Forum Chinese Sustainable Development Symposium, 2007(4): 358-362.
- [5] Xie Rongrong, Pang Yong, Xu Xintong, et al. Research on pollutant discharge tendency in Taihu Lake Basin of Jiangsu Province and determine of cutting total volume quota [J]. Water Resources Protection, 2015(6): 165-169.
- [6] Gao Junfeng, Jiang Zhigang. Conservation and Development of China's Five Largest Freshwater Lakes[M]. Beijing: Science Press, 2012.
- [7] GB. 3838-2002, Environmental quality standard for surface water[S]. Beijing: China Environmental Science Press, 2002.