

## Correlation analysis of urban lakes water quality indexes\*

Bo-Yang Sun<sup>†</sup> and Xiao-Hua Yang

*School of Environment, Beijing Normal University,  
Beijing, 100875, China*

*<sup>†</sup>E-mail: sunboyang@aliyun.com*

Xiao-Juan Chen

*Changjiang Water Resources Protection Institute, Wuhan, 430051, China*

*E-mail: chenxiaojuan8442@163.com*

Urban lakes water quality and quantity is affected by many factors. And there is a correlation between the various indicators of water quality and quantity. This paper selects the water temperature, pH, dissolved oxygen, chemical oxygen demand, permanganate, BOD, ammonia nitrogen as the main indexes, by correlation analysis of these indicators, find variation of Urban Lakes water quality and quantity and making use of monitoring data more reasonable. Research shows that water quality and quantity indexes showed significant cyclical and correlation, the highest correlation coefficient is 0.94.

*Keywords:* Urban Lakes; Water Quality and Quantity; Correlation Analysis.

### 1. Introduction

Urban water environment problem has become increasingly serious, especially in the urban landscape lake. Due to the low velocity of water flow and urban storm, water quality problems are often more severe than the river. In sunny days, the water quality in urban lakes will be slowly deteriorated[1]. And when after the storm, road debris and pollutants flowing into the lake, resulting in a sharp deterioration of the water quality. Previous studies focus on the type and concentration of pollutants, but few to study the intrinsic link between pollutants and causes[2]. Therefore, this article as a study area in Houhai (Beijing), analyzed correlation of water quality and quantity data in Houhai. Find the law and contact between the various indicators of water quality and quantity.

---

\* This work is supported by State Key Laboratory of Earth Surface Processes and Resource Ecology.

<sup>†</sup>This work was supported by the Project of National Natural Foundation of China (No.51079004,51379013), and the State Key Program of National Natural Science of China (No. 41530635).

## 2. Method

There are many methods of correlation analysis, the method more used is Pearson correlation coefficient and canonical correlation. Taking into account the characteristics of water quality data, water quality data in Houhai is equidistant measurement data, therefore, as used herein, the Pearson correlation coefficient method to analyze water quality and quantity.

The correlation coefficient was first proposed by Karl Pearson, the correlation coefficient is a statistical indicator to reflect the close correlation degree between the variables. Pearson correlation coefficient is calculated as follows [3]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{Y})^2}} \quad (1)$$

Where  $r$  is the correlation coefficient;  $x_i$  and  $y_i$  is the  $i$ -th measured value;  $\bar{X}$  and  $\bar{Y}$  represent the mean of the two sets of data (X and Y).

## 3. Correlation Analysis of Water Quality Data in Houhai

Houhai refers to a lake in Xicheng District of central Beijing, China. It is larger than other lakes nearby in central Beijing. It is a popular tourist attraction of China. In this article, taken Houhai as study area, water quality data taken from the lake outlet. Study area diagram is as Figure 1.



Fig. 1. Houhai Lake.

The main water quality indexes is water temperature, pH, dissolved oxygen, chemical oxygen demand, permanganate, BOD and ammonia nitrogen. The monthly data from 2009.01-2011.12. Sampling point is located in the lake outlet. Water quality sampling data as shown below (Figure 2).

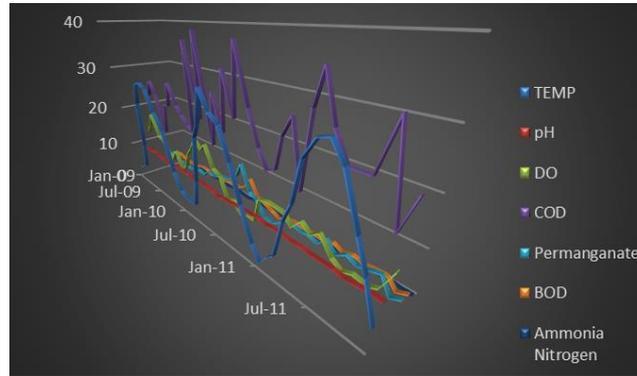


Fig. 2. Water quality monitoring data in Houhailake.

Above each set of data, calculate the correlation coefficient of every two sets of data, correlation coefficient matrix obtained as follows (Table 1).

Table 1 Correlation Coefficient Matrix.

	TEMP	pH	DO	COD	Permanganate	BOD	Ammonia Nitrogen
TEMP	1.00	-0.52	-0.73	0.11	0.01	0.08	0.27
pH	-0.52	1.00	0.38	-0.33	-0.43	-0.47	-0.38
DO	-0.73	0.38	1.00	-0.09	0.07	0.04	-0.12
COD	0.11	-0.33	-0.09	1.00	0.37	0.32	0.24
Permanganate	0.01	-0.43	0.07	0.37	1.00	0.94	0.21
BOD	0.08	-0.47	0.04	0.32	0.94	1.00	0.23
Ammonia Nitrogen	0.27	-0.38	-0.12	0.24	0.21	0.23	1.00

As shown in Figure 2, water quality data showed a periodic variation. This is mainly due to the Beijing rainfall concentrated in summer and seasonal temperature changes significantly.

The correlation coefficient ranges from -1 to +1 inclusive, where 1 is total positive correlation, 0 is no correlation, and -1 is total negative correlation [4]. As shown in Table 1, both have positive correlation and negative correlation between the quality indexes. Water temperature and pH showed a significant negative correlation, the lower the temperature the higher the pH value. Water temperature and DO showed a more significant negative correlation too. Apart from DO, pH have negative correlation with all other indexes. The correlation coefficient between Permanganate and BOD is 0.94, it showed a significant positive correlation. Explanation permanganate play an important role in biochemical reactions.

#### **4. Conclusion**

The above study shows that water quality and quantity indexes showed significant cyclical and correlation. There are both positive and negative correlation. Position and the corresponding pollution sources vary from lake to lakes. But through correlation analysis, we can always find certain regularity. In this article, the monitoring points is less and monitoring kinds of pollutant to be increased. This needs further study to help us find variation of Urban Lakes water quality and quantity and making use of monitoring data more reasonable.

#### **5. References**

1. Y. Mei, X.H. Yang, C. Chang, J. He and R. Jiang, Comprehensive Assessment of Pollutant Removal in Bioretention, *Advanced Science Letters* 10(2012) 698-699.
2. Y. Mei, X.H. Yang, C. Chang, J. He and R. Jiang, Set Pair Analysis of Optimal Choice of Bioretention Media, *Advanced Science Letters* 10(2012) 687-689.
3. K. Pearson, Notes on regression and inheritance in the case of two parents, *Proceedings of the Royal Society of London* 58 (1895) 240-242.
4. National Council on Measurement in Education, [http://www.ncme.org/ncme/NCME/Resource\\_Center/](http://www.ncme.org/ncme/NCME/Resource_Center/)
5. S. Y. Huang, M. H. Lee, C. K. Hsiao, Nonlinear measures of association with kernel canonical correlation analysis and applications, *Journal of Statistical Planning and Inference* 139 (7) (2009) 2162.