Study on Dissolved Oxygen, Ammonia-nitrogen and Permanganate Index in Water of Lake Dianchi

Lucen Li
College of Computer Science, Chong Qing University, Chongqing, 400044, China
angela@cas-harbour.org

Abstract. The water environment of Lake Dianchi in Yunnan province is threatened by eutrophication, which leads to the decrease of dissolved oxygen in water. Meanwhile, as the index reflecting the pollution of organic and inorganic oxidizable substances in water, permanganate index is also affected. China has been developing the automatic monitoring system for a series of water quality factors, such as dissolved oxygen and permanganate index. Additionally, the monitoring and prediction of ammonia-nitrogen, one of the water eutrophication sources, are also important. This study analyzed the water quality factor data in Lake Dianchi from 2015 to 2018 and implemented the prediction based on the Long Short-Term Memory neural network algorithm. It is found that the water quality of Lake Dianchi will be in a relatively ideal state in the future, but it cannot be improved further by itself. The Kunming government should implement more targeted improvement measures according to the characteristics of water quality factors, such as the seasonality of Lake Dianchi. Those measures will improve the ecological service quality of Lake Dianchi effectively, ensure the daily water safety of people in Kunming, and bring more economic values to the city.

1. Introduction
China is a country with many lakes. With stable conditions, lakes have the abilities to provide water for producing and living, rendering aquatic resources, regulating atmospheric and so on[1]. Water pollution is a crucial threat to maintaining the values of lakes. Lake Dianchi is the primary water source in Yunnan Province. With the rapid development of economy and society, the water quality in Lake Dianchi declines and pollution is increasingly severe. Since the 1990s, the overall water quality in Lake Dianchi becomes worse than type V[2].
Table 1 Basic Items of Environmental Quality Standards for Surface Water Unit: mg/L

| Number | Standard value | Type | I | II | III | IV | V |
|--------|----------------|------|---|----|-----|----|----|
| 1      | Temperature (°C) |      | Water temperature changed by human should be limited to: maximum rise of weekly mean temperature ≤ 1 maximum drop of weekly mean temperature ≤ 2 |
| 2      | PH Value | 6 ~ 9 |
| 3      | Dissolved oxygen | ≥ 7.5 | 6 | 5 | 3 | 2 |
| 4      | Permanganate index | ≤ 2 | 4 | 6 | 10 | 15 |
| 5      | Chemical oxygen demand (COD) | ≤ 15 | 15 | 20 | 30 | 40 |
| 6      | Five days' biochemical oxygen demand (BOD5) | ≤ 3 | 3 | 4 | 6 | 10 |
| 7      | Ammonia-nitrogen (NH3-N) | ≤ 0.15 | 0.5 | 1.0 | 1.5 | 2.0 |
| 8      | total phosphorus (TP) | ≤ 0.01 | 0.025 | 0.05 | 0.1 | 0.2 |
| 9      | total phosphorus (TN) | ≤ 0.2 | 0.5 | 1.0 | 1.5 | 2.0 |

Water pollution has changed the value structure of Lake Dianchi. From 2006 to 2015, the average annual ecosystem service value of Lake Dianchi was 386.55 billion yuan and the water pollution in Lake Dianchi caused a loss of 1.523 billion yuan to the economic and social development in Yunnan Province. The principal economic losses mainly came from the cost of water recharge and cyanobacteria removal[1].

Previous researchers have found that Lake Dianchi was eutrophicated and polluted. In 2018, the water quality of Caohai of Lake Dianchi raised to class IV[3], which means that the treatment of Lake Dianchi has made some achievements in local prevention and treatment. However, due to the complexity and uncertainty of the ecological restoration process of lake water, there is no fundamental solution to the pollution control in Lake Dianchi. Therefore, this research aimed to analyze the feasibility of maintaining Permanganate index, NH3-N and Dissolved oxygen of Lake Diachi that meet the requirement of type IV, and upgrading water quality to type III in the future.

Molecular oxygen in air dissolves in water becomes the dissolved oxygen, which is usually called DO. The amount of DO in water is an index to measure the self-purification ability of water. If DO in water is consumed and it takes a short time to restore to the initial state, it is indicated that the water has a strong self-purification ability or the water pollution is not serious. Permanganate index refers to the amount of consumed oxidant when treating water samples with potassium permanganate as oxidant under certain conditions. It represents the total amount of reducing substances (mainly organic pollutants) in the water sample that can be oxidized by potassium permanganate. The higher the permanganate index is, the more serious the water is polluted by organic matter[4]. NH3-N is a nutrient in water, which can lead to water eutrophication. It is a major oxygen consuming pollution in water, which is toxic to fish and other aquatic organisms[5].

2. Literature review
Eutrophication is one of the significant challenges to the lake environment in China. Eutrophication is a kind of water pollution caused by excessive nitrogen, phosphorus and other plant nutrients. It causes algae and other plankton to proliferate rapidly, resulting in reduced dissolved oxygen in the water. Monitoring and controlling the content of NH3-N, DO and Permanganate index is an important measure to protect the water quality of Lake Dianchi. At present, many scholars have made lots of researches on the water quality index that affects eutrophication and its modeling methods.

Yang Yinkang and Zou Xuexian[6] selected environmental quality assessment parameters,
determined the assessment standards of various pollutants, calculated the comprehensive assessment map of environmental quality and obtained the pollution situation of each part of Lake Dianchi according to the map. Therefore, this study selected three parameters: Permanganate index, NH$_3$-N and DO, to determine the pollution condition of Lake Dianchi. Moreover, Chen Yuling and Zhang Shisen[7] selected four main limiting factors raised by Snicedder: the limit values of total phosphorus concentration, chlorophyll concentration, transparency and saturation of dissolved oxygen in the deep water layer as the discriminant of the eutrophication process of Lake Dianchi. Through the experiment, Chen Yuling established a phosphorus model to describe the total phosphorus concentration over time. The model was calibrated through the measured data of 1988, and verified with the measured data of 1989. It used the mean square root verification error to represent the simulation accuracy. Chen Canyun, Zhang Baoxu and Li Yuliang[8] established a model to determine apparent settling velocity S of total phosphorus on the basis of the measured data of the Miyun Reservoir from 1985 to 1995 and used the apparent settling velocity S to predict the change of TP concentration. Therefore, this study focused on the change trend of observed data of Permanganate index, NH$_3$-N and DO to predict these three parameters of Lake Dianchi. Liu Liping[9] adopted chlorophyll concentration, total phosphorus concentration and water transparency as the discriminant, used the modified Carlson index to evaluate the eutrophication level of Lake Dianchi and analyzed the possible causes from 1988 to 1999.

As the rise of more water quality control schemes in Lake Dianchi, more experts studied the water environment of Lake Dianchi. Liang Zhongyao and Liu Yong[10] took Waihai of Lake Dianchi as the analysis object, selected eight detection points to detect the water quality indicators and used STL and RSI method to decompose the water quality indicators to study and analyze the change trend. Therefore, this study identified the variation trend of water quality index in Lake Dianchi based on the long time series of water quality data. In addition, many scholars have predicted the trend of water quality according to different models. For example, Zou Zhihong and Wang Lejuan[11] applied the Gray Markov model to predict the eutrophication trend of Lake Taihu, Lake Dianchi and Lake Chaohu. Yang Zhiying[12] established a BP network model based on the observed data of six environmental quality assessment parameters and used the model to predict the pollution trend of Lake Honghu. Moreover, Zhang Hucai and Chang Fengqin[13] used the detected data of four points of Dianchi Lake from April to September in 2015 to analyze the transformation of each parameter from the dry season to the rainy season, as well as the characteristics and spatial changes of the rainy season. This study also studied water quality changes under different seasons and analyzed the fluctuation of water quality in Lake Dianchi. Chen Yingyi, Cheng Qianqian[14] used PCA to select the water quality parameters and established a prediction model for dissolved oxygen in water for aquaculture based on long short-term memory neural. Through the experiment, the PCA-LSTM network model was proved to meet the actual demand of predicting dissolved oxygen. Therefore, this study adopted the LSTM neural network to predict Permanganate index, NH$_3$-N and DO in Lake Dianchi.

This project focused on Permanganate index, NH$_3$-N and Dissolved oxygen (DO) in Lake Dianchi, used scientific computing tools to analyze the pollution condition of Lake Dianchi, and employed Deep learning prediction model (LSTM) to evaluate the feasibility of improving water quality of the whole Lake Dianchi to break from type IV in the future. If the water environment in Lake Dianchi is improved in a certain situation in the computer simulation, it means that it might be a feasible solution to the pollution treatment of Lake Dianchi. Meanwhile, it can provide the Kunming Government with more precise guidance in pollution treatment and strategic planning. Moreover, it is helpful for Kunming Government to reduce economic losses in water treatment and restore the ecosystem services of Lake Dianchi.

3. Research methods

3.1. Long Short-Term Memory neural network

LSTM was proposed in 1997 by Sepp Hochreiter and Jürgen Schmidhuber[15]. It is an special kind of
Recurrent Neural Networks (RNN). Standard RNN cannot deal with long minimal time lags between relevant signals. Unlike RNN, a common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. LSTM adds or removes information to the cell state through the gate structure and the cell can remember values over arbitrary time intervals. Therefore, LSTM has ability to deal with the issues that are highly related to time series.

3.2. LSTM water quality factor prediction model

3.2.1. The data sample
This project took the surface water quality monitoring data of Lake Dinachi, Guanyin Mountain site, in Kunming, Yunnan province as the research object, and used LSTM to construct the prediction model of water quality factors.

China environmental monitoring station uses automatic water quality monitoring system to detect water quality. At present, most of the monitoring instruments in China are mainly used to monitor the comprehensive pollution index, and there is a lack of instruments to monitor individual pollutants, such as the instruments to monitor the organic pollutants[16].

This project took Permanganate index, NH$_3$-N and DO value from 2015 to 2018 as the input parameters to the model and predicted their change trend. Partial data of selected data samples are shown in Table 2.

| time          | name          | PH value | DO(mg/L) | NH$_3$-N(mg/L) | Permanganate index(mg/L) |
|--------------|--------------|----------|----------|---------------|------------------------|
| 2017-08-01 00:00:00 | Guanyin Mountain | 8.36     | 5.56     | 0.52          | 12.10                   |
| 2017-08-01 04:00:00 | Guanyin Mountain | 8.20     | 5.19     | 0.61          | 12.60                   |
| 2017-08-01 08:00:00 | Guanyin Mountain | 7.98     | 4.77     | 0.65          | 11.08                   |
| 2017-08-01 12:00:00 | Guanyin Mountain | 7.65     | 3.76     | 0.70          | 10.43                   |

In order to ensure the stability of time series data and the high efficiency of the model, the data is first normalized and converted to [0, 1]. This project adopted the Min-Max scaling method. The formula (1) is shown as follows:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

3.2.2. Test results
Using the Python 3.6.3 as programming language, this project was implemented under the development environment of Anaconda 4.3.30 and based on TensorFlow 1.13.1. Parameters related to the experiment were set as follows: LSTM time step 3, hidden size of hidden neurons 32, batch size 10, num_layers 1, training_steps 2000. This project predicted the water quality data of the next 12 months based on water quality data samples from 2015 to 2018.

In order to display the prediction results of the LSTM algorithm for various water quality factors visually, the real data and prediction data graphs of three factors are drawn respectively, as shown in Figure 1-3. The figures show the change trend of each factor in the next 12 months. Since the time step set is 3, the data of the first 3 months are omitted from the figure.
Fig. 1. DO prediction results

Fig. 2. NH$_3$-N prediction results
3.2.3. The model evaluation indices

This project adopted the unⅣersal mean square error (MSE) to evaluate the regression prediction model. The calculation formula is shown in formula (2).

\[
MSE(y, y') = \frac{1}{n} \sum_{i=1}^{n} (y_i - y'_i)^2
\]  

(2)

The calculation results of MSE of each water quality factor are shown in Table 3. It can be seen from the table that the relatⅣe error of the water quality prediction model based on LSTM is low.

| predictors         | DO   | NH₃-N | Permanganate index |
|--------------------|------|-------|-------------------|
| MSE                | 5.386| 1.408 | 0.0423            |

4. Experimental summary

Water pollution has affected the ecosystem service of Lake Dianchi severely and caused serious economic losses. This project analyzed three main water quality factors in Lake Dianchi in the past three years and predicted the change trend of each factor in the next twelve months based on LSTM algorithm. Four results are summarized as followed:

1. Both actual and predicted values of NH₃-N meet the requirements of type Ⅲ. Although it might fluctuate in the next 12 months, most of the time is still lower than 0.5 and has reached the class Ⅲ water quality requirement. Therefore, the ammonia nitrogen is not the main pollutant in Lake Dianchi.

2. Only one of the predicted values of DO does not meet the requirements of type Ⅲ. It suggests that Lake Dianchi still has the self-purification ability and DO is less likely to be reduced to type Ⅳ standard in the future.

3. Permanganate index is the main pollution factor. The measured data of the water quality of Lake Dianchi over thirty-six months are uniformly over 6, which does not meet the requirements of class Ⅲ water quality. In addition, more than half of the time does not meet the requirements of class Ⅳ water quality. Moreover, it shows a certain seasonality, which is at a high level in the middle of each year. And this seasonality is also reflected in the predicted results. Therefore, the Kunming government should pay more attention to the seasonality of permanganate index, and adjust the preventive measures. The feasibility of breaking away from type Ⅳ and upgrading to type Ⅲ is high, if the Kunming government attaches importance to the control of permanganate index.
4. The prediction model of water quality based on LSTM shows prediction results with small errors, which provides a guiding method for the development of automatic water quality monitoring system in China. By increasing the prediction function of water quality monitoring system, it can find and improve the deficiencies of equipment timely and provide better monitoring scheme.

5. Conclusion
This project used the data measured by the automatic surface water monitoring system from the Ministry of Ecology and Environmental of the People’s Republic of China. The function of automatic water monitoring system is not perfect and the data defaults and exceptions are common, which increases the difficulty of data analysis. If water quality data are measured in a more regulated approach, the data will be more accurate, which will improve the accuracy of prediction and provide more specific guidance to water improvement and the development of automatic water monitoring.

References
[1] Wei, G. Du, Z.P., Yan, C.A., Chen,Y. (2019) Evaluating net ecosystem services value of a polluted lake: A case study of Lake Dianchi. Acta Ecologica Sinica, 39 (5).
[2] Feng, N., Mao, F., Li, X.Y., Zhang, A.D. (2010) Research on Ecological Security Assessment of Dianchi Lake. Environmental Science, 31 (2).
[3] Monthly report on water quality monitoring of lakes and plateaus. (2018) Yunnan Provincial Department of Ecological Environment.
[4] Lai, J.F., Wu, J.N., Liu, L.B. (2018) Study on the correlation of surface water dissolved oxygen with permanganate index and chemical oxygen demand. City and Town Water Supply.
[5] Jiang, R., Zeng, Y.H., Wang, Q. (2013) Research progress of ammonia nitrogen wastewater treatment technology Environmental Science and Management, 38 (06): 131-134.
[6] Yang, Y.K., Zou, X.X. (1982) Study on environmental quality evaluation of Lake Dianchi. Acta Academia Medicine Kunming.
[7] Chen, Y.L., Zhang, S.S. (1991) A Simulation Model for Eutrophication Prediction in Lake Dianchi. Shuili Xuebao.
[8] Chen, Y.C., Zhang, B.X., Li, Y.L.(1998) Nutrient analysis and prediction of Miyun reservoir. Shuili Xuebao.
[9] Liu, L.P. (2001)Analysis on Eutrophication of Dianchi Lake and its Control Measures. Yunan Environmental Science.
[10] Liang, Z.Y., Liu, Y., Sheng, H., et al. (2014) Identification of water quality trend in Lake Dianchi using an integrated STL and RSI approach Acta Science Circumstantiae, 34(3).
[11] Zou, Z.H., Wang, L.j. (2009) Gray Markov forecast of lake eutrophication. Acta Scientiae Circumstantiae, 29(2).
[12] Zhang, Q. (2016) Prediction of Water Quality Index of Honghu Lake based on Back Propagation Neural Network Model, 14(2).
[13] Zhang, H.C., Chang, F.Q., Duan, L.Z., et al. (2017) Water quality characteristics and variations of Lake Dianchi. Advanced in Earth Science, 32 (6).
[14] Chen, Y.Y., Cheng, Q.Q., et al. (2018) Principal component analysis and long short-term memory neural network for predicting dissolved oxygen in water for aquaculture. Transactions of the Chinese Society of Agricultural Engineering, 34(17).
[15] Hochreiter, S., Schmidhuber, J. (1997) Long Short-Term Memory. MIT Press.
[16] Yang, H.D. (2019) Interpretation of Application and Development Trend of Surface Water Environment Automatic Monitoring Technology. Energy Conservation & Environmental Protection.