Saudi Journal of Biological Sciences (2017) **24**, 556–562

King Saud University

Saudi Journal of Biological Sciences

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Analysis of algae growth mechanism and water bloom prediction under the effect of multi-affecting factor

Li Wang *, Xiaoyi Wang, Xuebo Jin, Jiping Xu, Huiyan Zhang, Jiabin Yu, Qian Sun, Chong Gao, Lingbin Wang

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, 100048 Beijing, China

Received 13 October 2016; revised 5 January 2017; accepted 9 January 2017
Available online 24 January 2017

KEYWORDS

Chemical mechanism;
Algae growth;
Water blooms;
Prediction;
Multi-factor

Abstract The formation process of algae is described inaccurately and water blooms are predicted with a low precision by current methods. In this paper, chemical mechanism of algae growth is analyzed, and a correlation analysis of chlorophyll-a and algal density is conducted by chemical measurement. Taking into account the influence of multi-factors on algae growth and water blooms, the comprehensive prediction method combined with multivariate time series and intelligent model is put forward in this paper. Firstly, through the process of photosynthesis, the main factors that affect the reproduction of the algae are analyzed. A compensation prediction method of multivariate time series analysis based on neural network and Support Vector Machine has been put forward which is combined with Kernel Principal Component Analysis to deal with dimension reduction of the influence factors of blooms. Then, Genetic Algorithm is applied to improve the generalization ability of the BP network and Least Squares Support Vector Machine. Experimental results show that this method could better compensate the prediction model of multivariate time series analysis which is an effective way to improve the description accuracy of algae growth and prediction precision of water blooms.

1. Introduction

Algae bloom is a typical manifestation of eutrophication caused by extreme rise in nutrients such as nitrogen and phosphorus in lakes. Under the appropriate conditions of temperature, illumination, climate and hydrology, colored algae floaters are developed caused by explosive breeding in lakes (Yulia and Nadezhda, 2010). Thus, there is no clarity about the critical factors and mechanism of algae blooms, and some

* Corresponding author.
E-mail address: wangli@th.btbu.edu.cn (L. Wang).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sjbs.2017.01.026
1319-562X

effective technologies of prevention and treatment of algae blooms are still lacking in general. It is convenient for relevant departments to adopt countermeasures to reduce the hazards under the condition of predicting the occurrence of blooms accurately.

2. Related work

The method of time series analysis is a kind of vital mathematical tool to study stochastic process by means of using historical data to speculate the future trend of development of things (Luo and Wu, 2009; Wang et al., 2013, 2014, 2016; Jin, 2013; Jin et al., 2013a,b). However, a model that is suited for reflecting the relationship between outer affecting factors and data needs to be adopted to make compensation toward the prediction error of time series. To make time series prediction using neutral network is widely paid attention to which has access to success with its good nonlinear properties and distributed storage structure as well as high fault tolerance (Frank et al., 2001; Zhang and Qi, 2003). Li has adopted PNN to make prediction toward blooms which has provided an effective way for forecasting (Li et al., 2011). Zhu has adopted improved BP network to forecast the highest point of nonlinearity about chlorophyll (Zhu et al., 2011). Mulia has developed a hybrid model combined with ANN and GA (Genetic Algorithm) to make real-time forecast as well as mid-long term forecast (Mulia et al., 2013). However, traditional BP(Back Propagation) network and Least Squares Support Vector Machine have some shortcomings such as slow convergence speed and are plunged into local extreme that may lead to low precision unable to accurately describe the strength of the bloom (Xia et al., 2009). For this reason, KPCA (Kernel Principal Component Analysis) is applied to make dimension reduction toward affecting factors while GA is used to make optimization toward random initial weights. Finally, BP network and Least Squares Support Vector Machine are used to make compensation toward the prediction result gotten from time series which is an effective way to improve the accuracy and to realize complementarity between time series and intelligent method.

3. Material and methods

3.1. Biochemical mechanism analysis of algae bloom

Eutrophication is a phenomenon of water pollution caused by excessive nitrogen, phosphorus and other plant nutrients. Aquatic organisms, especially the large number of algae reproduction, change the amount of biomass population and destroy the ecological balance of the water body. The algae are synthesized by photosynthesis using the sunlight and inorganic compounds. The process of eutrophication is as follows.

$$106CO_2 + 16NO_3^- + HPO_4^{2-} + 112H_2O + 18H^+ + \text{energy} + \text{trace element} \rightarrow (CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 138O_2$$

Here $(CH_2O)_{106}(NH_3)_{16}(H_3PO_4)$ is chemical formula of algae.

The algae cell concentration shows the number of algae in the water and the intensity of the bloom. Before water blooms, algae growth rapidly, the concentration of algal cells increased continuously from tens of thousands to hundreds of millions per liter. Chlorophyll-a is one of the important components of algal cells. All algae contain chlorophyll-a. The content of chlorophyll-a is closely related to the type and quantity of algae in water, it is also related to the quality of water environment. Therefore, through the determination of chlorophyll-a in water, the number of algae in the water and water quality status is reflected.

Fig. 1 is the two-dimensional molecular structure of chlorophyll-a.

Water samples are collected from sampling point using a plastic bucket, and then the collected water samples are loaded into a transparent glass jar and placed in the incubator, cultured under constant temperature of 5000Lx, 26 °C. The algae density and chlorophyll-a of water sample are measured at the same time every day. Algae density is measured by blood cell counting plate microscopy. Chlorophyll-a is measured by the United States YSI company's 6600 type multi-function water quality monitor (Fig. 2).

The correlation between algae density and chlorophyll-a is studied. The growth curve was measured by the algae density and chlorophyll-a daily measurement data (Fig. 3).

Chlorophyll-a concentration and algal density data are shown as follows (Fig. 4).

Through regression analysis of chlorophyll-a and algal density data, the relationship between chlorophyll-a and algal density is linear, and there is a significant correlation between them. Fig. 5 shows the algae cell.

Through the process of photosynthesis, the main factors that affect the reproduction of the algae can be analyzed, such as the nutrient content of water body, light, water temperature and other physical and chemical factors. Nitrogen is a component of algae, phosphorus is directly involved in the process of photosynthesis, respiration, activation of enzyme system and energy conversion, both of which are essential for the growth of algae and the occurrence of water bloom. Algae bloom generally occurs in the condition of high temperature, small wind and slow lake flow. In addition, PH value in water is mainly affected by the content of carbon dioxide. The PH value changes, mainly due to the photosynthesis process by which algae absorb carbon dioxide and release oxygen, affecting the ion balance. Therefore, in this paper, the selected influence factors on the prediction of chlorophyll-a are as follows: PH value, oxygen consumption (OC), water temperature (T), turbidity, total nitrogen (TN), total phosphorus (TP), dissolved oxygen (DO), and so on.

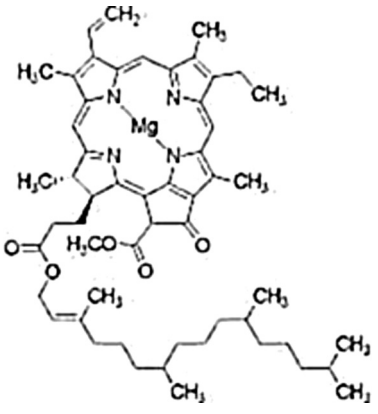

Figure 1 Molecular structure of chlorophyll-a.