Neural Comput & Applic (2016) 27:19–26
DOI 10.1007/s00521-013-1538-0

Freshwater algal bloom prediction by extreme learning machine in Macau Storage Reservoirs

Inchio Lou · Zhengchao Xie · Wai Kin Ung · Kai Meng Mok

Received: 25 September 2013 / Accepted: 16 December 2013 / Published online: 3 January 2014

Abstract Understanding and predicting dynamic change of algae population in freshwater reservoirs is particularly important, as algae-releasing cyanotoxins are carcinogens that would affect the health of public. However, the high complex nonlinearity of water variables and their interactions makes it difficult in modeling its growth. Recently, extreme learning machine (ELM) was reported to have advantages of only requirement of a small amount of samples, high degree of prediction accuracy and long prediction period to solve the nonlinear problems. In this study, the ELM-based prediction and forecast models for phytoplankton abundance in Macau Storage Reservoir are proposed, in which the water parameters of pH, SiO_2, and some other water variables selected from the correlation analysis were included, with 8-year (2001–2008) data for training and the most recent 3 years (2009–2011) for testing. The modeling results showed that the prediction and forecast (based on data on the previous 1st, 2nd, 3rd and 12th months) powers were estimated as approximately 0.83 and 0.90, respectively, showing that the ELM is an effective new way that can be used for monitoring algal bloom in drinking water storage reservoir.

Keywords Algal bloom · Phytoplankton abundance · Extreme leaning machine · Prediction and forecast models

I. Lou · Z. Xie (✉) · K. M. Mok
Faculty of Science and Technology, University of Macau, Taipa, Macau SAR
e-mail: zxie@umac.mo

W. K. Ung
Laboratory and Research Center, Macao Water Co. Ltd, Taipa, Macau SAR

1 Introduction

Freshwater algal bloom is one of the water pollution problems that occur in eutrophic lakes or reservoirs due to the presence of excessive nutrients. It has been found that most species of algae (also called phytoplankton) can produce various cyanotoxins including *microcytins, cylindrospermopsis* and *nodularin*, which have directly impact on the water treatment processes and consequently affect the health of public [1]. Thus, it is of great importance to understand the population dynamics of algae in the raw water storage units. However, modeling the algae population in such a complicated system is a challenge, as the physical, chemical and biological processes as well as the interaction among them are involved, resulting in the highly nonlinear relationship between phytoplankton abundance and various water parameters.

Computational artificial intelligence techniques have been developed as the efficient tools in recent years for predicting (without considering time series effect) or forecasting (considering time series effect) algal bloom. Previous studies [2] have used the principle component regression (PCR), i.e., principal component analysis (PCA) followed by multiple linear regressions (MLR), to predict chlorophyll-*a* levels, the fundamental index of phytoplankton. However, the intrinsic problem of PCR is that the variables data set used as the input of the model has high complex nonlinearity, expecting that PCR alone is inadequate for prediction and the prediction results were unsatisfactory. With the development of artificial intelligence models, artificial neural network (ANN) such as back propagation (BP) was applied to predict the algal bloom by assessing the eutrophication and simulating the chlorophyll-*a* concentration. ANN is a well-suited method with self-adaptability, self-organization and error tolerance,

which is better than PCR for nonlinear simulation. ANN has been used for predicting the chlorophyll concentration [4–6]. However, this method has such limitations as requirement of a great amount of training data, difficulty in tuning the structure parameter that is mainly based on experience, and its "black box" nature that is difficult to understand and interpret the data [2, 3].

Considering the drawbacks of the both methods, recently extreme learning machine (ELM) is thought as the best solution. ELM is a simple and efficient learning algorithm that was developed recently. In the name of ELM, extreme means that its learning speed is extremely fast while it has higher generalization than the gradient-descent-based learning [7], and it has been shown that ELM could be 10 times faster compared with some traditional algorithms such as BP [8]. Furthermore, ELM can be used to solve issues like local minima, improper learning rate and over-fitting which are very possible in traditional ANN [5, 7]. Examples [4–7] also showed that ELM possesses a superior performance than other conventional algorithms on different benchmark problems from both regression and classification areas. There are some existing works [9–11] with using ELM, and through these works, it can be seen that ELM could have a very good performance for some engineering applications. By far, as the best knowledge of authors, there is no existing application of using ELM on prediction or forecast the phytoplankton abundance in algal blooms.

In this study, it is attempted to develop an ELM-based predictive model to simulate the dynamic change of phytoplankton abundance in Macau Reservoir given a variety of water variables. The measured data from 2001 to 2011 were used to train and test the model. The present study will lead to better understanding of the algal problems in Macau, which will help to develop later guidelines for forecasting the onset of algae blooms in raw water resources.

2 Materials and methods

2.1 MSR and water parameters measurement

Macau is situated 60 km southwest of Hong Kong, and experiences a subtropical seasonal climate that is greatly influenced by the monsoons. The difference of temperature and rainfall between summer and winter is significant though not great. Macau Main Storage Reservoir (MSR) (Fig. 1), located in the east part of Macau peninsula, is the biggest reservoir in Macau with the capacity of about 1.9 million m^3 and the water surface area of 0.35 km^2. It is a pumped storage reservoir that receives raw water from the West River of the Pearl River network and can provide

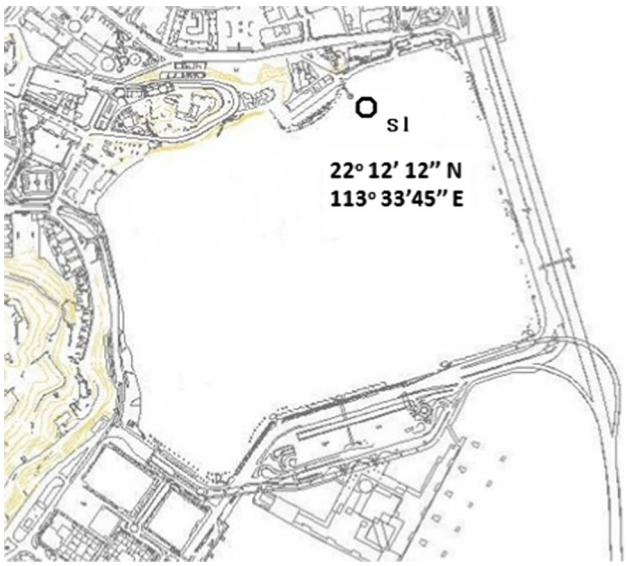

Fig. 1 Location of the MSR

water supply to the whole areas of Macau for about 1 week. MSR is particularly important as the temporary water source during the salty tide period when high salinity concentration is caused by intrusion of sea water to the water intake location. In recent years, there were reports (Macao Water Co. Ltd., unpublished data) that the reservoir experienced algal blooms and the situation appeared to be worsening.

Macau Water Supply Co. Ltd. is responsible for water quality monitoring and management. Location in the inlet of the reservoir was selected for sampling. Samples were collected in duplicate monthly from May 2001 to February 2011 at 0.5 m from the water surface. A total of 23 water quality parameters, including hydrological, physical, chemical and biological parameters, were monitored monthly. Precipitation was obtained from Macau Meteorological Center (http://www.smg.gov.mo/www/te_smgmail.php). Imported volume, exported volume and water level were recorded by the inlet and outlet flow meters, based on which the hydraulic retention time (HRT) can be calculated. Turbidity, temperature, pH, conductivity, chloride (Cl^-), sulfate (SO_4^{2-}), silicon (SiO_2), alkalinity, bicarbonate (HCO_3^-), dissolved oxygen (DO), ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-), total nitrogen (TN), phosphorus (PO_4^{3-}), total phosphorus (TP), suspended solid, total organic carbon (TOC) and UV_{254} and iron (Fe) were measured according to the standard methods [12, 13]. The phytoplankton samples were fixed using 5 % formaldehyde and transported to laboratory for microscopic counting.

In this work, correlation analysis was conducted to identify the water parameters which were significantly correlated with phytoplankton abundance. Only the

parameters with the correlation coefficients greater than 0.3 are selected as inputs in the ELM models. It was also noted that the parameters selected in forecast models are different from those in the prediction models, as the water parameters in previous data were also used in the correlation analysis. In our study, two types of forecast models were used, depending on the monthly data used as the inputs. Forecast model 1 was based on the last 3-month data, while forecast model 2 was based on the last 3-month data as well as the previous 12th month data, i.e., the previous 1st, 2nd and 3rd and 12th month data. The purpose of adding the previous 12th month data in the forecast model 2 is to take the historical effect the last year that have similar environmental conditions, as the environmental conditions, such as temperature, influence the growth of phytoplankton.

2.2 Extreme learning machine

ELM originally was proposed as a learning scheme for single-hidden-layer feed-forward neural networks (SLFNs). Then, it was extended to the generalized SLFNs where the hidden layer needs not be neuron alike [14, 15]. In the past, gradient-descent-based approaches were used for feed-forward neutral networks, and all parameters need to be tuned, which usually take a long time. While for ELM which the basic idea is that, the model has only one hidden layer, and the parameters of this hidden layer, including the input weights and biases of the hidden nodes, need not to be tuned. On the contrary, these hidden nodes parameters are assigned randomly, which means that they may be independent of the training data [15]. It should be pointed out that bigger the number of hidden nodes, the higher the accuracy of the built model. However, the modeling and computation time will also be elongated. The random input weights affect the performances significantly. However, it is not easy to find a proper set of initial weights. It is recommended to run more epochs and enlarge the training data set if possible. After these input weights and hidden layer biases are assigned randomly, SLFNs can be treated as linear system, and output weights which link hidden layer to the output layer can be calculated using generalized inverse operation [16–18]. References [4, 5, 7] proposed and proved the theory of ELM. In order to make it clear on ELM and its application in water treatment prediction, here the fundamental theory of ELM in [7] will be briefly reintroduced first as follows:

Consider a training data set $\boldsymbol{D}$ of N arbitrary distinct samples $(\boldsymbol{x}_i, \boldsymbol{t}_i)$, where $\{\boldsymbol{x}_i\} \in R^m$ is the $m \times 1$ input vector and $\{\boldsymbol{t}_i\} \in R^n$ is the $n \times 1$ target vector. Standard SLFNs with $\tilde{N}$ nodes and activation function $g(x)$ are mathematically modeled a

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(\boldsymbol{x}_j) = \sum_{i=1}^{\tilde{N}} \beta_i g(\boldsymbol{w}_i, b_i, \boldsymbol{x}_j) = \boldsymbol{o}_j, \quad 1 \le j \le N \tag{1}$$

where $\boldsymbol{w}_i$ is weight vector connecting the ith hidden node and the input nodes, β_i is the weight vector connecting the ith hidden node and the output nodes, and b_i is the threshold of the ith hidden node.

Since the goal is to find the relation between $\boldsymbol{x}_i$ and $\boldsymbol{t}_i$, if the SLFNs can approximate the training data with zero error (i.e., $\sum_{j=1}^{\tilde{N}} \boldsymbol{o}_j - \boldsymbol{t}_j = 0$), then there exists β_i, $\boldsymbol{w}_i$ and b_i such that Eq. (2) is satisfied.

$$\sum_{i=1}^{\tilde{N}} \beta_i g(\boldsymbol{w}_i, b_i, \boldsymbol{x}_j) = \boldsymbol{t}_j, \quad 1 \le j \le N \tag{2}$$

The above N equations can be written compactly as

$$\boldsymbol{H}\beta = \boldsymbol{T} \tag{3}$$

where

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}(\boldsymbol{x}_1) \\ \vdots \\ \boldsymbol{h}(\boldsymbol{x}_N) \end{bmatrix} = \begin{bmatrix} g(\boldsymbol{w}_1, b_1, \boldsymbol{x}_1) & \cdots & g(\boldsymbol{w}_{\tilde{N}}, b_{\tilde{N}}, \boldsymbol{x}_1) \\ \vdots & \ddots & \vdots \\ g(\boldsymbol{w}_1, b_1, \boldsymbol{x}_N) & \cdots & g(\boldsymbol{w}_{\tilde{N}}, b_{\tilde{N}}, \boldsymbol{x}_N) \end{bmatrix}_{N \times \tilde{N}}, \tag{4}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times n} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times n} \tag{5}$$

$\boldsymbol{H}$ is called the hidden layer output matrix of SLFN. $\boldsymbol{h}(\boldsymbol{x}) = g(\boldsymbol{w}_1, b_1, \boldsymbol{x}), \ldots, g(\boldsymbol{w}_{\tilde{N}}, b_{\tilde{N}}, \boldsymbol{x})$ is called the hidden layer feature mapping. The ith column of H is the ith hidden node output with respect to inputs $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$. The ith row of $\boldsymbol{H}$ is the hidden layer feature mapping with respect to the ith input $\boldsymbol{x}_i$.

According to the proofs in [4, 5], if the activation function is infinitely differentiable, the input weight vectors $\boldsymbol{w}_i$ and hidden layer biases b_i can be randomly assigned. Moreover, these parameters are not necessarily tuned, and the hidden layer output matrix $\boldsymbol{H}$ can actually remain unchanged once random values have been assigned in the beginning of learning.

Different from traditional learning algorithms, ELM tends to reach not only the smallest training error but also the smallest norm of output weights [19]:

$$\text{Minimize} : \|\boldsymbol{H}\beta - \boldsymbol{T}\|^2 \text{ and } \|\beta\| \tag{6}$$

Then, if the number $\tilde{N}$of hidden neurons is equal to the number N of distinct training samples (i.e., $\tilde{N} = N$), the matrix $\boldsymbol{H}$ is square and invertible, which means that the

output weights β can be analytically calculated by simply inverting $\boldsymbol{H}$, and thus the SLFNs can approximate these training samples with zero error. However, most of the times, the number of hidden nodes is much less than the number of distinct training samples (i.e., $\tilde{N} \ll N$), and thus $\boldsymbol{H}$ is a non-square matrix and there may not exist β_i, $\boldsymbol{w}_i$ and b_i, and Eq. (3) cannot be satisfied. Fortunately, since $\boldsymbol{w}_i$ and b_i are fixed, Eq. (3) becomes a linear system, and the smallest norm least square method can be used instead of the standard optimization method to estimate the output weights.

$$\beta = \boldsymbol{H}^{\dagger}\boldsymbol{T} \tag{7}$$

where $\boldsymbol{H}^{\dagger}$ is the Moore–Penrose pseudoinverse of matrix $\boldsymbol{H}$ [16], which can be calculated using the orthogonal projection method [18]:

$$\boldsymbol{H}^{\dagger} = \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}} \text{ when } \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} \text{ is nonsingular} \tag{8}$$

$$\text{or } \boldsymbol{H}^{\dagger} = \boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\right)^{-1} \text{ when } \boldsymbol{H}\boldsymbol{H}^{\mathrm{T}} \text{ is nonsingular} \tag{9}$$

where the superscript T means matrix transposition.

Based on this learning algorithm, the training time can be extremely fast because only three calculation steps are required: (1) randomly assign hidden nodes parameters; (2) calculate the hidden layer output matrix $\boldsymbol{H}$; and (3) calculate the output weight β. Moreover, since the output weights are calculated analytically using inverse matrix, it ensures that the results are global and hence better prediction accuracy and generalization performance can be achieved. After training, the output function of ELM for an unseen vector $\boldsymbol{X}$ (take one output node case as an example) can be expressed as:

$$f(\boldsymbol{X}) = \boldsymbol{h}(\boldsymbol{X})\beta \tag{10}$$

2.3 Performance indicators

The performance of models was evaluated using the following indicators: square of correlation coefficient (R^2) that provides the variability measure for the data reproduced in the model; mean absolute error (MAE) and root mean square error (RMSE) that measure residual errors, providing a global idea of the difference between the observation and modeling. The indicators were defined as below by Eqs. 11–15.

$$R^2 = 1 - \frac{F}{F_o} \tag{11}$$

$$F = \sum\left(Y_i - \hat{Y}_i\right)^2 \tag{12}$$

$$F_o = \sum\left(Y_i - \overline{Y_i}\right)^2 \tag{13}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2 \tag{14}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2} \tag{15}$$

where n is the number of data; Y_i and $\overline{Y_i}$ are observation data and the mean of observation data, respectively, and $\hat{Y}_i$ is the modeling results.

3 Results and discussion

The correlation of $\log_{10}$ phytoplankton and water parameters for forecast model and prediction model are shown in Table 1. Parameters with correlation coefficients greater than 0.3 (highlighted in bold) will be retained in the models. It was also noted that the parameters selected in forecast models are different from those in the prediction models, as the water parameters in previous data (past

Table 1 Correlation analysis of prediction and forecast model

Parameters	Prediction model	Forecast model Time lagged (month)			
		$t-1$	$t-2$	$t-3$	$t-12$
Turbidity	−0.03	0.00	−0.01	−0.06	−0.25
Temperature	0.19	0.21	0.19	0.14	0.22
pH	**0.49**	**0.42**	**0.38**	**0.33**	**0.33**
Conductivity	−0.08	0.01	0.14	0.21	−0.24
Cl^-	0.01	0.10	0.22	0.28	−0.16
SO_4^{2-}	−0.03	0.03	0.14	0.22	−0.28
SiO_2	**0.33**	**0.31**	0.16	0.04	−0.08
Alkalinity	**−0.34**	**−0.30**	−0.21	−0.12	**−0.36**
HCO_3^-	**−0.46**	**−0.40**	**−0.32**	−0.24	**−0.38**
DO	**0.39**	**0.35**	**0.34**	**0.31**	0.18
NO_3^-	−0.29	−0.22	−0.22	−0.15	**−0.35**
NO_2^-	−0.10	−0.08	−0.02	0.03	−0.23
NH_4^+	0.11	0.10	0.08	0.25	0.05
TN	**0.68**	**0.60**	**0.53**	**0.46**	0.23
UV_{254}	**0.56**	**0.55**	**0.48**	**0.47**	−0.07
Fe	−0.14	−0.06	−0.04	−0.08	−0.27
PO_4^{3-}	0.02	0.06	0.06	0.03	0.11
TP	0.08	0.05	0.02	0.00	−0.21
Suspended solid	**0.31**	**0.35**	**0.31**	0.23	−0.10
TOC	**0.38**	**0.33**	0.29	**0.35**	0.07
HRT	−0.12	−0.11	−0.13	−0.16	0.10
Water level	0.13	0.05	0.01	−0.02	0.10
Precipitation	−0.09	0.05	0.11	0.06	−0.05
Phytoplankton abundance	–	**0.82**	**0.71**	**0.62**	0.24

Table 2 Performance indexes of the prediction and forecast models

Performance index	Prediction model		Forecast model			
	Accuracy performance (training set)	Generalization performance (testing set)	Accuracy performance (training set)		Generalization performance (testing set)	
			$t-1, 2, 3$	$t-1, 2, 3, 12$	$t-1, 2, 3$	$t-1, 2, 3, 12$
R^2	0.82492	0.83217	0.8637	0.9009	0.8702	0.9033
RMSE	0.25798	0.30129	0.2246	0.2236	0.3643	0.3166
MAE	0.19976	0.23663	0.1794	0.1597	0.2565	0.2258

record) were also used in the correlation analysis. In the forecast models of ELM, phytoplankton abundance (t) is a function of water parameter ($t-1$), water parameter ($t-2$) and water parameters ($t-3$), where $t-1$, $t-2$, $t-3$ and $t-12$ represent the 1, 2, 3 and 12 months prior to time t. Thus, there were only nine parameters used in the prediction models, and 23 time-lagged parameters selected for the forecast models.

After the correlation analysis, it comes to the testing of the models invoked two parts, the accuracy performance and the generalization performance. Accuracy performance is to test the capability of the model to predict the output for the given input set that originally used to train the model, while generalization performance is to test the capability of the model to predict the output for the given input sets that were not in the training set. In order to prevent the model that is memorizing the inputs instead of generalized learning, both performance checks need to be considered. In the present research, the performance indexes for ELM-based models were averaged with 50 runs.

In the application of ELM in this work, for the predication models, after the correlation analysis, nine parameters, such as pH and SiO_2, are selected as the independent variables, and phytoplankton abundance is selected as the induced variable (target value). Then, the data from May 2005 to December 2008 are used to train the model, and data from January 2009 to February 2011 are used to test the model. In the training process, the cross-validation approach as mentioned previously is adopted to obtain the optimal combination of parameters for the testing. Specifically, the training data are divided into 10 about the same size groups that 9 groups for training and the rest 1 group is used to test the model trained by the previous 9 groups' data. Then, this (9 groups for training and 1 for group testing) is repeated for 9 times (10 times in total). And then, parameters of the one process, which has the best testing performance in these 10 repeats, will be used as the optimal parameters combination in the 'real' testing process, which has the data from January 2009 to February 2011. The forecast model basically follows the same steps of the prediction model, while the only difference between these two models is that effect of time series is included in the forecast model. So, in the forecast model, only the previous 3 months' data are included in the training process.

The performance of prediction and forecast models was shown in Table 2. The results indicated that the ELM was successful in the prediction and forecast phytoplankton abundance in MSR, with the R^2 greater than 0.82 for both training and testing data sets. Compared to the prediction model, ELM had better performance with the R^2 of 0.8637 (0.8702), RMSE of 0.2246 (0.3643) and MAE of 0.1794 (0.2565) for training (testing), suggesting that the historical water parameters including the phytoplankton abundance have effect on the prediction, which can improve the prediction power. Furthermore, when including the previous 12th month data as input in the forecast model, the prediction power of the forecast model can increase up to 0.9 with the RMSE of 0.2236 (0.3166) and MAE of 0.1597 (0.2258) for training (testing). These results further confirmed the historical effects on the model accuracy and generalization performance, and also implied that take the previous 12th month data as memorizing learning can improve the prediction power in the forecast model. Besides, further compared with our previous study [22] for forecast of phytoplankton abundance using support vector

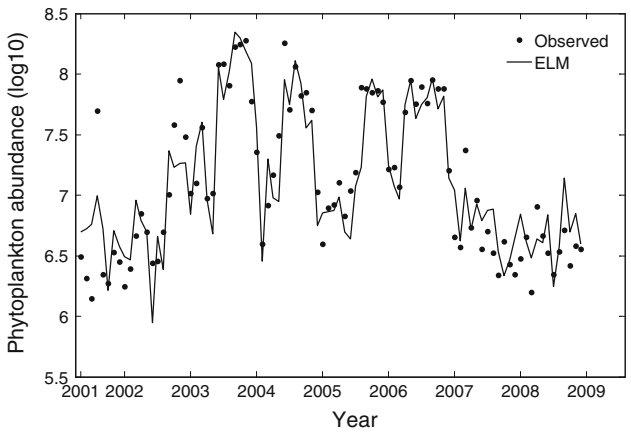

Fig. 2 Observed and predicted phytoplankton level for the training and validation data set of the prediction models

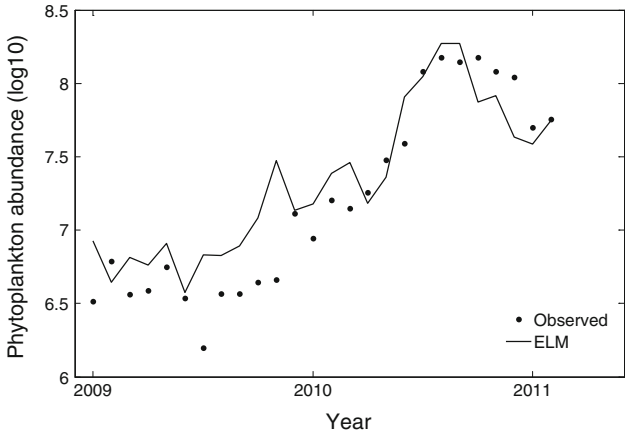

Fig. 3 Observed and predicted phytoplankton level for the testing data set of the prediction models

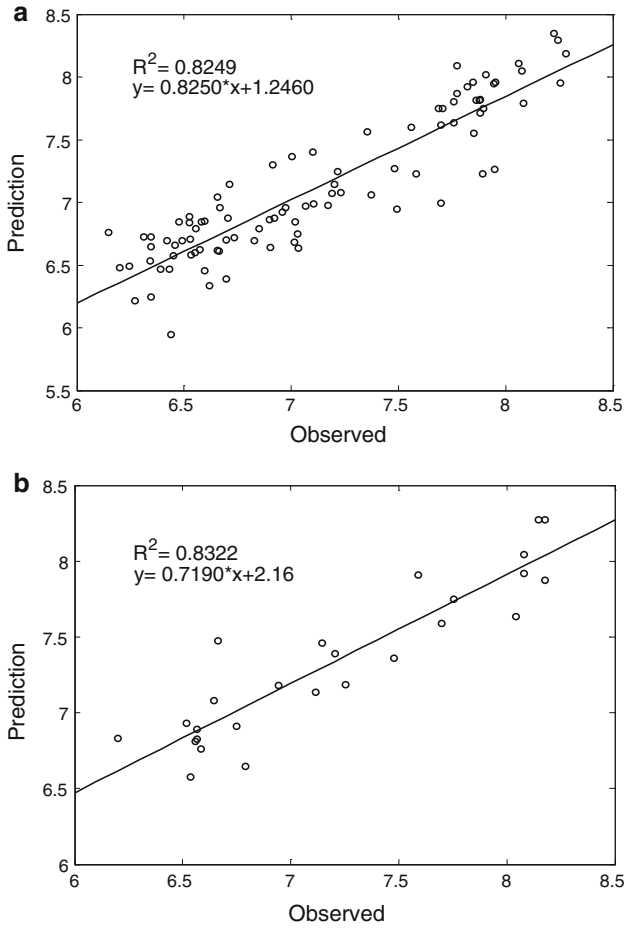

Fig. 4 ELM results for the training and validation (**a**) and testing (**b**) data set of the prediction model

machine (SVM) with R^2 of 0.86, the present study using ELM has better prediction power with R^2 of 0.9.

The observed data versus the modeling data are shown in Figs. 4, 7 and 10, and the observed and modeling

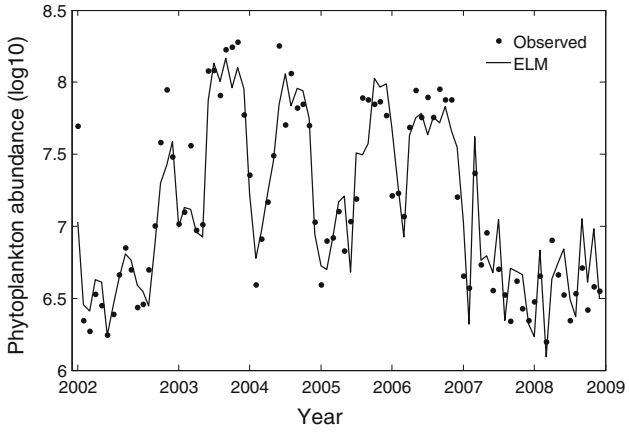

Fig. 5 Observed and predicted phytoplankton level for the training and validation data set of the forecast model 1 that based on the previous 1st, 2nd and 3rd months data

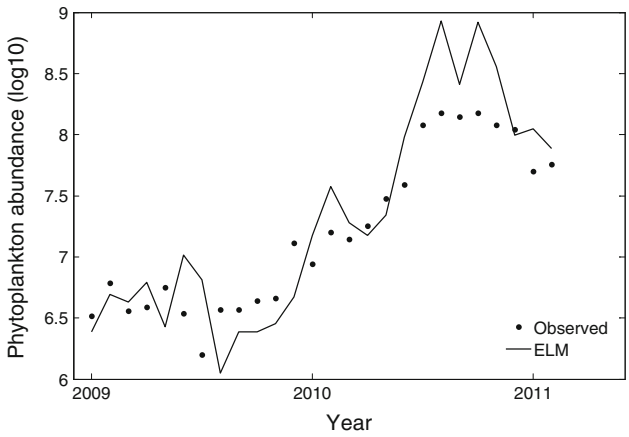

Fig. 6 Observed and predicted phytoplankton level for the testing data set of the forecast model 1 that based on the previous 1st, 2nd and 3rd months data

phytoplankton abundance changes over time are listed in Figs. 2, 3, 5, 6, 8 and 9. These results confirmed that ELM can handle well the nonlinear relationship between water parameters and phytoplankton abundance.

4 Conclusions

The ELM-based prediction and forecast models for phytoplankton abundance in MSR are proposed in this study. Fifteen water parameters with the correlation coefficients against phytoplankton abundance greater than 0.3 were selected, with 8-year (2001–2008) data for training and cross-validation, and the most recent 3 years (2009–2011) for testing. The results showed that the forecast model have better performance with the R^2 of up to 0.9 than prediction model with the R^2 of 0.83, implying that the algal bloom problem is a complicated nonlinear dynamic system that is

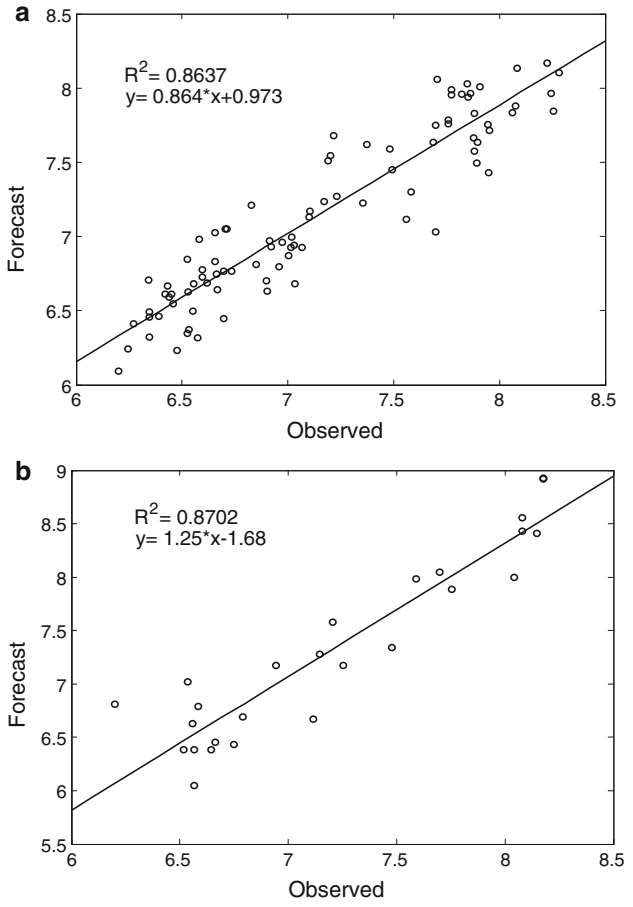

Fig. 7 ELM results for the training and validation (**a**) and testing (**b**) data set of the forecast model 1 that based on the previous 1st, 2nd and 3rd months data

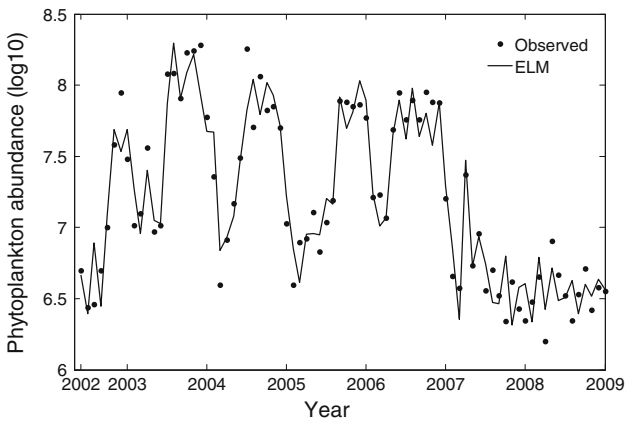

Fig. 8 Observed and forecasted phytoplankton lever for the training and validation data set of the forecast model 2 that based on the previous 1st, 2nd, 3rd and 12th months data

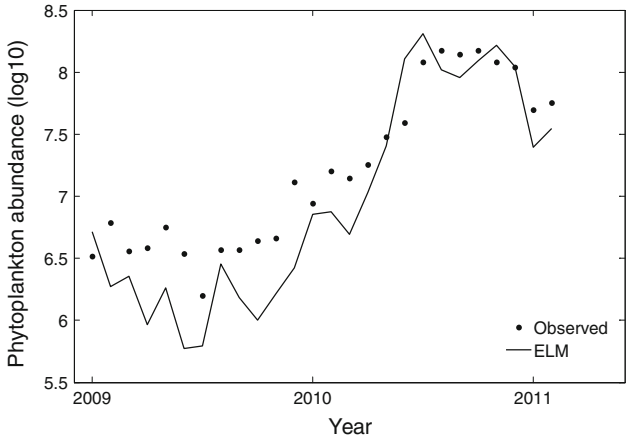

Fig. 9 Observed and predicted phytoplankton level for the testing data set of the forecast model 2 that based on the previous 1st, 2nd, 3rd and 12th months data

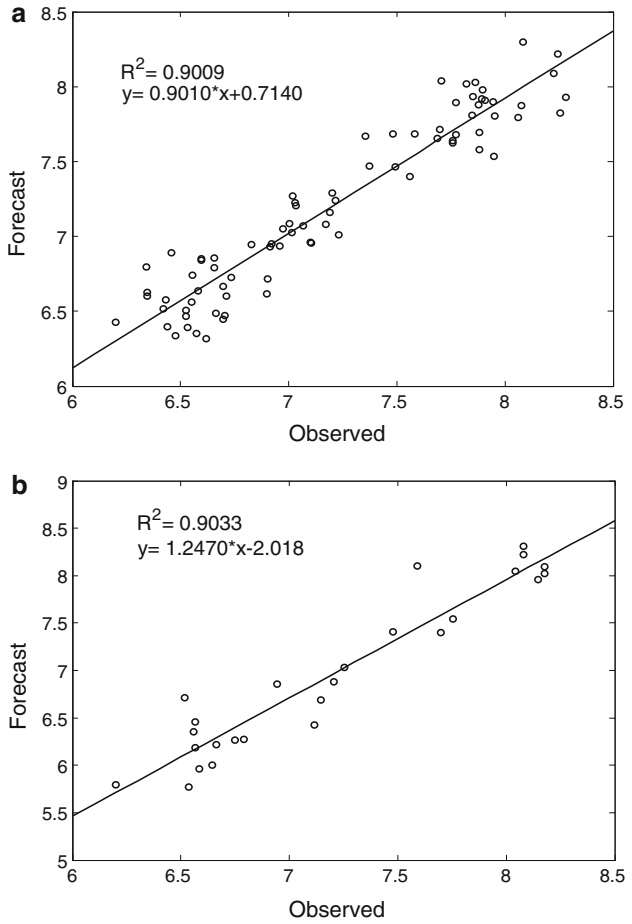

Fig. 10 ELM results for the training and validation (**a**) and testing (**b**) data set of the forecast model 2 that based on the previous 1st, 2nd, 3rd and 12th months data

affected not only by the water variables in current month but also by those in a couple of previous months. In addition, including the previous 12th month data in the forecast model, ELM in the study showed superior forecast power and root mean square errors, indicting that the historical water parameters and phytoplankton abundance

have impact on the phytoplankton dynamics of the reservoir. These results will provide an effective way for water quality monitoring and management of drinking water storage reservoirs. In addition, additional numerical approaches and optimization algorithms can be applied to enhance the performance [20–23].

Acknowledgments We thank Macao Water Co. Ltd. for providing historical data of water quality parameters and phytoplankton abundances. The financial support from the Fundo para o Desenvolvimento das Ciências e da Tecnologia (FDCT) (Grant # FDCT/016/2011/A) and Research Committee at University of Macau are gratefully acknowledged.

References

1. Selman Z, Greenhalgh S, Diaz R (2008) Eutrophication and hypoxia in coastal areas: a global assessment of the state of knowledge. World Resources Institute, Washington, DC
2. Pallant J, Chorus I, Bartram J (2007) Toxic cyanobacteria in water, SPSS Survival Manual
3. Hecht-Nielsen R (1987) Kolmogorov's mapping neural network existence theorem. In: Proceedings of 1st IEEE international joint conference of neural networks, New York
4. Huang G-B, Zhu Q-Y, Siew C-K (2004) Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE international conference on neural networks—conference proceedings, 2, 985–990
5. Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489–501
6. Huang G-B, Chen L, Siew C-K (2006) Universal approximation using incremental constructive feedforward networks with random hidden nodes. IEEE Trans Neural Netw 17(4):879–892
7. Wong KI, Wong PK, Cheung CS, Vong CM Modeling and optimization of biodiesel engine performance using advanced machine learning methods. Energy. doi:10.1016/j.energy.2013.03.057
8. Deng WY, Zheng QH, Chen L, Xu XB (2010) Power utility nontechnical loss analysis with extreme learning machine method. Chin J Comput 33(2):280–287
9. Nizar AH, Dong ZY, Wang Y (2008) Power research on extreme learning of neural networks. IEEE Trans Power Syst 23(3):946–955
10. Xu Y, Dong ZY, Meng K, Zhang R, Wong KP (2011) Real-time transient stability assessment model using extreme learning machine. IET Gener Transm Distrib 5(3):314–322
11. Su ZL, Ng KM, Soszyńska-Budny J, Habibullah MS (2011) Application of the LP-ELM model on transportation system lifetime optimization. IEEE Trans Intell Transp Syst 12(4):1484–1494
12. Rogers LL, Dowla FU (1994) Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resour Res 30:457–481
13. APHA (2002) Standard methods for the examination of water and wastewater. A. W. W. A. a. W. E. F. American Public Health Association
14. Huang G-B, Chen L (2007) Convex incremental extreme learning machine. Neurocomputing 70(16–18):3056–3062
15. Huang G-B, Chen L (2008) Enhanced random search based incremental extreme learning machine. Neurocomputing 71(16–18):3460–3468
16. Huang G-B, Wang DH, Lan Y (2011) Extreme learning machines: a survey. Int J Mach Learn Cybern 2(2):107–122
17. Penrose R (1955) A generalized inverse for matrices. Proc Camb Philos Soc 51(3):406–413
18. Rao CR, Mitra SK (1971) Generalized inverse of matrices and its applications. Wiley, New York
19. Huang G-B, Zhou H, Ding X, Zhang R (2012) Extreme learning machine for regression and multiclass classification. IEEE Trans Syst Man Cybern B Cybern 42(2):513–529
20. Cattani C, Chen S, Aldashev G (2012) Information and modeling in complexity. Math Probl Engineering 2012, AID 868413
21. Chen S, Zheng Y, Cattani C, Wang W (2012) Modeling of biological intelligence for SCM system optimization. Comput Math Methods Med 2012, AID 769702
22. Lu P, Chen S, Zheng Y (2013) Artificial intelligence in civil engineering. Math Probl Eng 2013, AID 145974
23. Xie Z, Lou I, Ung WK, Mok KM (2012) Freshwater algal bloom prediction by support vector machine in Macau Storage Reservoirs. Math Probl Eng 2012, Article ID 397473, pp 12