Available online at www.ijournalse.org

Emerging Science Journal

Vol. 4, No. 2, April, 2020

Artificial Neural Network Model to Prediction of Eutrophication and *Microcystis Aeruginosa* Bloom

Pawalee Srisuksomwong [a*], Jeeraporn Pekkoh [b]

[a] *Faculty of Science and Technology, Phuket Rajabhat University, Phuket, Thailand*

[b] *Department of Biology, Faculty of Science, Chiang Mai University, Chiangmai, Thailand*

Abstract

Maekuang reservoir is one of the water resources which provides water supply, livestock, and recreational in Chiangmai city, Thailand. The water quality and *Microcystis aeruginosa* are a severe problem in many reservoirs. *M. aeruginosa* is the most widespread toxic cyanobacteria in Thailand. Difficulty prediction for planning protects Maekuang reservoirs, the artificial Neural Network (ANN) model is a powerful tool that can be used to machine learning and prediction by observation data. ANN is able to learn from previous data and has been used to predict the value in the future. ANN consists of three layers as input, hidden, and output layer. Water quality data is collected biweekly at Maekuang reservoir (1999-2000). Input data for training, including nutrients (ammonium, nitrate, and phosphorus), Secchi depth, BOD, temperature, conductivity, pH, and output data for testing as Chlorophyll *a* and *M. aeruginosa* cells. The model was evaluated using four performances, namely; mean squared error (MSE), root mean square error (RMSE), sum of square error (SSE), and percentage error. It was found that the model prediction agreed with experimental data. C01-C08 scenarios focused on *M. aeruginosa* bloom prediction, and ANN tested for prediction of Chlorophyll *a* bloom shown on M01-M09 scenarios. The findings showed, this model has been validated for prediction of Chlorophyll a and shows strong agreement for nitrate, Log cell, and Chlorophyll a. Results indicate that the ANN can be predicted eutrophication indicators during the summer season, and ANN has efficient for providing the new data set and predict the behavior of *M. aeruginosa* bloom process.

Keywords:

Artificial Neural Network Model;
Eutrophication;
Toxic Cyanobacteria;
Microcystis Aeruginosa.

Article History:

Received:	19	November	2019
Accepted:	23	February	2020
Published:	01	April	2020

1- Introduction

Reservoirs provide essential ecosystem services such as water for irrigation, drinking, food supply for many people around the world, and sites for recreation and tourism. Eutrophication and algal blooms have been recognized as a severe environmental [1]. These phenomena are caused by green algae, euglenoids, and dinoflagellate but are mostly caused by cyanobacteria, and this competitive dominance of cyanobacteria is fostered by their resistance to grazing, buoyancy regulation and massive accumulation of nutrients [2-3]. Algal blooms are the process whereby water resources become enriched by nutrients from external (e.g., municipal and industrial effluent, livestock processing, and agricultural runoff from fertilized topsoil) and internal source. Algal blooms can occasionally grow so dense, which they cause, not only often lead to water discoloration but also invertebrate and fish mortality due to oxygen depletion. Many cyanobacterial species are able to produce toxins and off-flavors [4-5]. *Microcystis aeruginosa* is one of the cyanobacteria, can produce a family of hepatotoxins called microcystins, which are the most frequently encountered cyanotoxins in freshwater [6-8].

* **CONTACT**: Pawalee.s@pkru.ac.th

DOI: http://dx.doi.org/10.28991/esj-2020-01217

In the present time, research on preventive technologies, control, and monitoring methods of algal bloom have received worldwide attention. The artificial neural network model (ANN) is the most useful for perdition in many problems [9-12], and the artificial neural network is one of the soft computing techniques, have been applied in the prediction of algal bloom [13-16]. Tian Wenchong [17] emphasized species of algae, chlorophyll dynamics model by using an artificial neural network for algal bloom forecasting. The artificial neural networks are easy to set up and can provide quick response and thus are suitable for real-time operation. Also, artificial neural networks can model dynamic, non-linear, and noisy data [18]. In this paper, the most common computation algorithm, back-propagation, was used in the ANN model to determine the nonlinear relationship between each input data. Finally, the prediction of eutrophication and *M. aeruginosa* bloom.

2- Theory

Artificial neural network (ANN) is predicted for species abundance and succession of algae by feedforward architecture characteristic with back-propagation for training. Back-propagation neural network is suitable for predicting because of training to learn the relationship with data. The model consisted of three layers, input, hidden, and output layer. The number of the hidden layers was generated by iteration, where consider a mean squared error value and a correlation coefficient. The sum of weighted each input ($x_1,\ldots,x_k$) and output is obtained by Equation 1.

$$y(k) = F\left(\sum_{i=1}^{m} w_i(k)\cdot x_i(k) + b\right) \tag{1}$$

Where $x_i(k)$ is the input variable, $w_i(k)$ is weight value, b is biased, F is a transfer function, and $y_i(k)$ is output value. The sigmoid function or activation function is the limits the amplitude of the output shows as:

$$f(z) = \frac{1}{1+\exp(-z)} \tag{2}$$

The multilayer feed-forward networks (Multilayer perceptrons) have a different number of nodes and different activation functions. In the feed-forward network, information flows along the connecting pathway, from the input layer vis the hidden layers to the output layer. The ANN has input (n_I), hidden neuron (n_H), and output neuron (n_o), the hidden layer ($h_1,\ldots,h_{nI}$) was calculated from the input layer ($x_1,\ldots,x_{nI}$). Finally, the output $y_1,\ldots,y_{n_O}$ was calculated from the hidden layer shown processing on Equations 3 and 4.

$$h_j(x) = F\left(\sum_{i=1}^{n_j} w_{jl}^{I} x_i + b_j^{I}\right) \tag{3}$$

$$y_j(x) = F\left(\sum_{i=1}^{n_0} w_{jl}^{0} x_i + b_j^{0}\right) \tag{4}$$

The artificial neural network was calculated by error measure. Given N case are available to evaluate the model, where y is the actual output and is the output from ANN.

The sum of squared errors is defined $$SSE = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{5}$$

And the mean sum of square error (MSE) $$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{6}$$

The root mean sum of square error (RMSE) $$RMSE = \sqrt{MSE} \tag{7}$$

The percentage error is a good indicator of the performance of the model as $$\%Error = \frac{RMSE}{width} \times 100$$

3- Methodology

3-1- Study Area

Maekuang Reservoir is the lake in the northern part of Thailand. The Maekuang Dam, which was built in 1977, is located at latitude 18°55'35.4"N and longitude 99°07'31.4"E (Figure 1). The dam is one of the water resources which provides water supply, livestock, and recreational in some parts of Chiangmai and Lamphun province, Thailand. At average water elevation of 390 m above sea level, the dam withholds a reservoir of 15 km^3.

The water quality data were used in this study were collected biweekly from April 1999 to September 2000, consist of nutrients (ammonium, nitrate, soluble reactive phosphorus (SRP), Secchi depth, biochemical oxygen demand (BOD), temperature, conductivity pH, chlorophyll *a* and *M. aeruginosa* cells.

Microcystis blooms in Maekuang Reservoir have been reported every summer. They observed that concentrations of toxic microcystin released by *Microcystis* cells varied from month to month and year to year depending on the composition of *Microcystis* species. Concentrations of both intra- and extra-cellular microcystin were measured complementary to water quality conditions and phytoplankton abundances two to three times per month. These data were linearly interpolated to produce daily values required for modeling by ANN.

In this paper, the Neural Network Toolbook in the MATLAB (License, 40512555, MathtWorks, Inc., 2018) is making an artificial neural network.

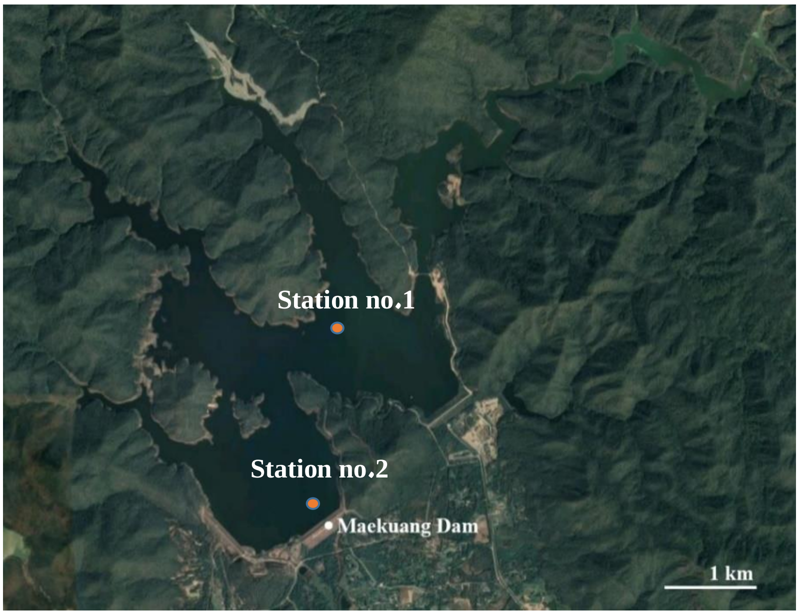

Figure 1. Map of Maekuang Dam.

3-2- The parameters of ANNs

The number of parameters in the artificial neural network. First, the input including ammonium (mg/L), nitrate (mg/L), SRP (mg/L), Secchi depth (m), BOD (mg/L), temperature (ºC), conductivity (µs/cm), pH, Chl-*a* (µg/L), Log cell (cell/mL), respectively. The change of *M. aeruginosa* at the same time point was presented as the output of the model. The total amount of Chlorophyll *a* was presented as the output (Figure 2). The initial weight value was assigned randomly based on input data. All the networks had ten nodes in the input layer and two nodes in the output layer.

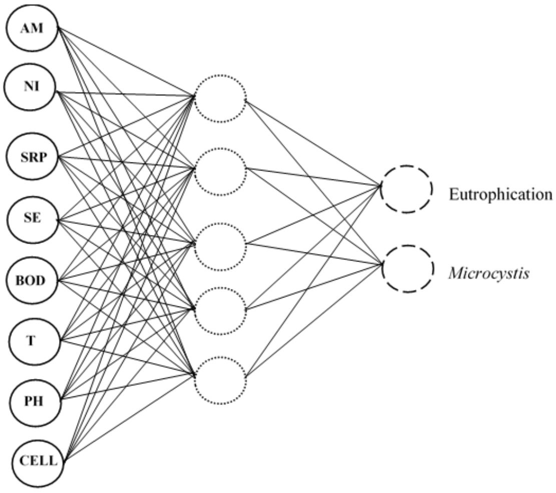

Figure 2. An ANN model between input data and output data in Maekuang Reservoir. ( ) represent the input layer; ( ) represent the hidden layer; ( ) represent the output layer.

4- Results and Discussion

The minimum, mean, maximum, and standard deviation (SD) values of the input layer (ammonium, nitrate, SRP, Secchi depth, BOD, temperature, conductivity, pH, Chl-*a*, and Log cell) showed in Table 1. The ANN model was trained by using the scenarios. However, two outputs (*M. aeruginosa* and Chlorophyll *a*) in Maekuang Reservoir was considered and shown in Tables 2 and 3.

Table 1. Maximum, minimum, average and Standard deviation (SD) of each parameter measured in 17 months period (biweekly).

Variables	Maximum	Minimum	Mean	SD
Ammonium (mg/L)	0.24	0.00	0.06	0.08
Nitrate (mg/L)	1.60	0.30	0.76	0.31
SRP (mg/L)	0.30	0.01	0.13	0.08
Secchi depth (m)	3.59	1.10	2.34	0.54
BOD (mg/L)	5.20	0.10	1.31	1.05
Temperature (ºC)	31.30	22.20	28.40	2.67
Conductivity (µs/cm)	149.70	68.00	83.24	16.23
pH	9.43	6.61	8.00	0.64
Chl-*a* (µg/L)	23.09	0.01	4.37	5.30
Log cell (cell/mL)	4.68	0.00	1.98	1.57

* SRP: Soluble reactive phosphorus; BOD: Biochemical oxygen demand; Chl-a: Chlorophyll *a*; Log cell: Log *M. aeruginosa* cell

For each neural network, the correlation coefficient for the model prediction during training and testing and root mean square error (RMSE) of *M. aeruginosa* blooms in Maekuang Reservoir are given in Table 2. It can be seen that scenario M04-M07 shows similar performance during training and testing. Also, the neural networks shown in Table 3 that provided scenario M08 are suitable for prediction of Chlorophyll *a*. It is challenging to decide which neural network is excellent. However, there is the advantage of the network consist of nitrate, Log cell, and Chlorophyll *a*. Thailand was located in the tropical area, divided the climate to three seasons as a rainy season (mid-May to mid-October), winter season (mid-October to mid-February), and summer season (mid-February to mid-May). According to Thailand Meteorology Department (TMD) data, the northern part of Thailand has to maximum temperature as winter (31.1C^{o}), summer (36.2C^{o}), and rainy (27.8C^{o}). 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, temperature, and other physical and chemical factors. Water temperature and nutrients levels, including total nitrogen (TN) and total phosphorus (TP) were reported to be positively correlated with the abundance of total *Microcystis* and cyanobacterial blooms [19-22]. Consequently, these environmental factors can be used to predict the formation of *Microcystis* and cyanobacterial blooms.

Table 2. Different tested neural networks for prediction of *M. aeruginosa* blooms in Maekuang Reservoir.

Scenario	Inputs	Root mean square error (RMSE)	Correlation coefficient	
		Output	Training	Testing
M01	Ammonium, Nitrate, SRP, Secchi depth, BOD, Temperature, Conductivity, pH, Log cell	2.1594	1.0000	0.9537
M02	Ammonium, Nitrate, SRP, Secchi depth, BOD, Temperature, Conductivity, Log cell	2.1399	1.0000	0.9756
M03	Ammonium, Nitrate, SRP, Secchi depth, BOD, Temperature, Log cell	2.0962	0.9941	0.9708
M04	Ammonium, Nitrate, SRP, Secchi depth, Temperature, Log cell	2.1136	0.9265	0.9957
M05	Ammonium, Nitrate, SRP, Temperature, Log cell	2.1254	0.9941	0.9914
M06	Ammonium, Nitrate, SRP, Log cell	2.1722	0.9977	0.9982
M07	Ammonium, Nitrate, Log cell	2.0821	0.9833	0.9935
M08	Nitrate, Log cell	2.0788	0.9499	0.9637
M09	Log cell	2.1189	0.9360	0.9905

Table 3. Different tested neural networks for prediction of Chlorophyll *a* bloom in Maekuang Reservoir.

Scenario	Inputs	Root mean square error (RMSE)	Correlation coefficient	
		Output	Training	Testing
C01	Ammonium, Nitrate, SRP, Secchi depth, BOD, Temperature, Conductivity, pH, Chl-*a*	1.2104	0.9959	0.9695
C02	Ammonium, Nitrate, SRP, Secchi depth, BOD, Temperature, Conductivity, Chl-*a*	1.1362	0.9913	0.9720
C03	Ammonium, Nitrate, SRP, Secchi depth, BOD, Temperature, Chl-*a*	1.3590	0.9850	0.9466
C04	Ammonium, Nitrate, SRP, Secchi depth, Temperature, Chl-*a*	1.1984	0.9734	0.9380
C05	Ammonium, Nitrate, SRP, Temperature, Chl-*a*	1.8101	0.9762	0.6826
C06	Ammonium, Nitrate, SRP, Chl-*a*	2.0765	0.9322	0.9661
C07	Ammonium, Nitrate, Chl-*a*	1.5222	0.9451	0.6312
C08	Nitrate, Chl-*a*	2.1608	0.8599	0.9981

Figure 3 and Figure 4 represented a comparison between ANN predicted and observation data during 17 months period (biweekly). Figure 3 showed that the *M. aeruginosa* almost highly from April to May, which is the summer season due to more sunlight available, lower water, temperature, and limited nutrient concentration. Moreover, the summer of 1999 is higher than in summer 2000 because of the Thailand drought in 1999. However, the scenario M05 – M09 provided the good of ANN predicted when compared with actual data. Figure 4 shown the ANN neural networks for the prediction of Chlorophyll *a* in Maekuang Reservoir. The scenario C08 is suitable for predicting. It is more error in October 1999, January 2000, and July 2000 because of the value of data closed to zero, and then prediction has an error.

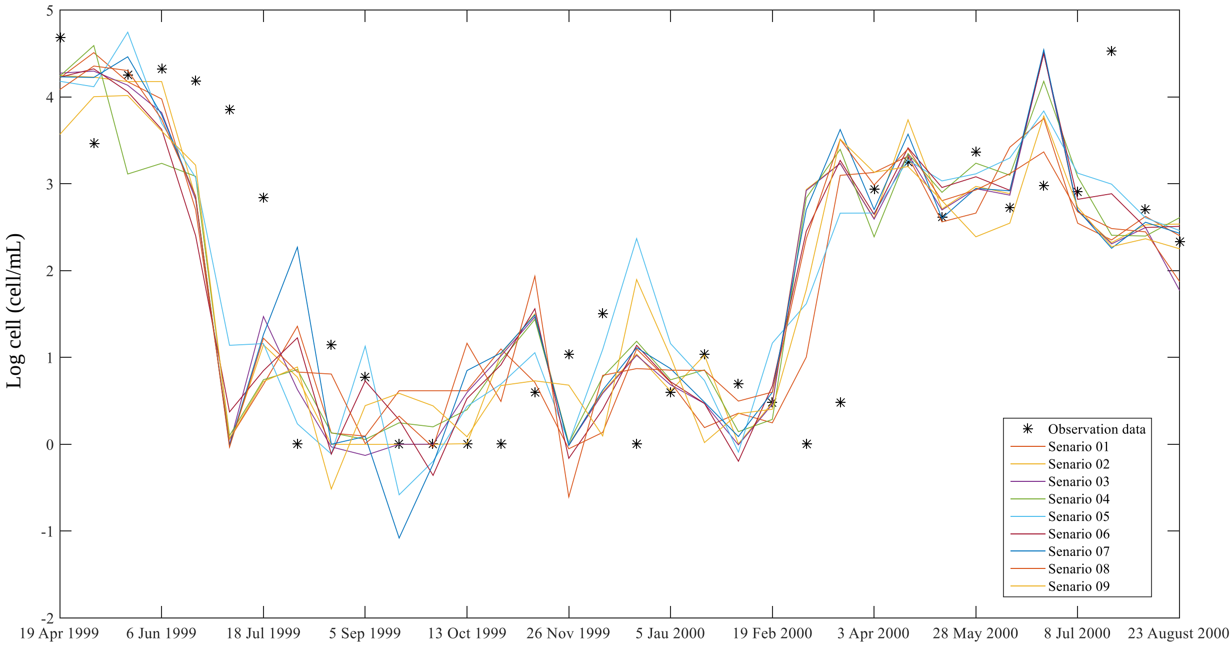

Figure 3. The ANN neural networks for prediction of *M. aeruginosa* blooms in Maekuang Reservoir.

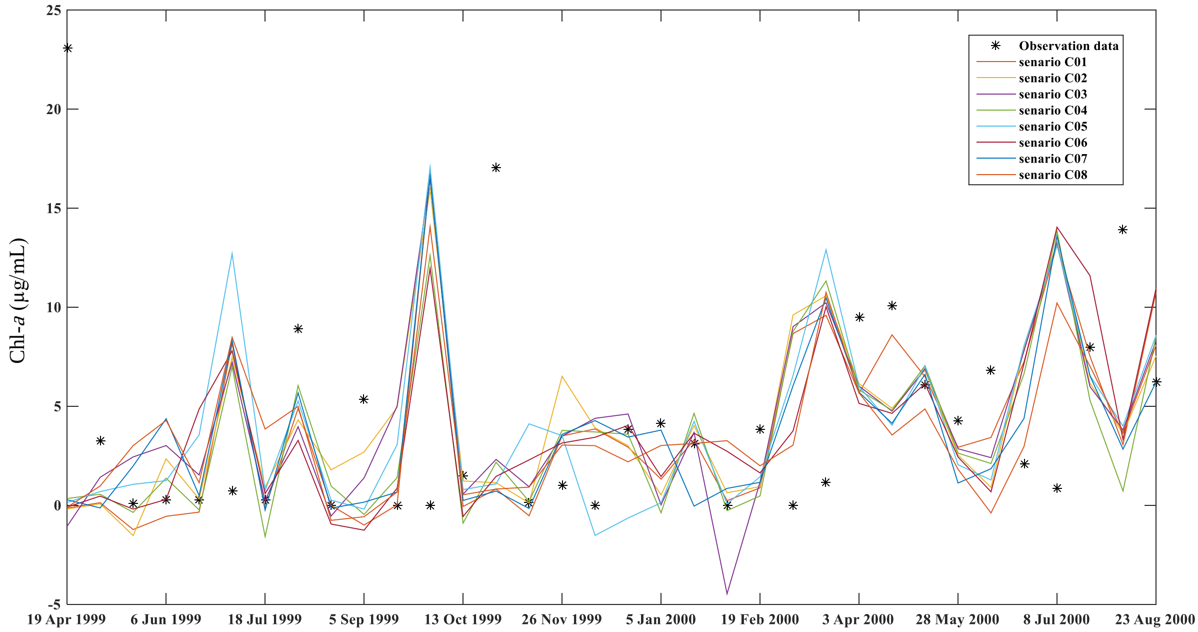

Figure 4. The ANN neural networks for prediction of Chlorophyll *a* bloom in Maekuang Reservoir.

5- Conclusion

The artificial neural network is excellent machine learning. It was learned from input data. The multilayer feed-forward networks were discussed. Water quality data is collected biweekly at the Maekuang reservoir (1999-2000) in the 17 months. Input data for training, including nutrients (ammonium, nitrate, and phosphorus), Secchi depth, BOD, temperature, conductivity, pH, and output data for testing as Chlorophyll *a* and *M. aeruginosa* cells. The statistical is the sum of square error, the mean sum of square error, the root mean square error and the percentage error. In 17 months period, shown maximum temperature 31.3 ºC, maximum Chlorophyll *a* 23.09 µg/L, and maximum Nitrate 1.60 mg/L. Results show that the scenario M05 – M09 provided the good of ANN prediction data, and M08 is suitable for prediction Chlorophyll *a*. The ANN has the relationship between nitrate and log cell with Chlorophyll *a*. The nitrate and log cell is an indicator of water quality in Maekuang reservoirs. The *M. aeruginosa* almost highly in the summer season. Thus, finding the relation between nitrate and Chlorophyll with the ANN, the scenario C08 is suitable for predicting of Chlorophyll *a* in Maekuang Reservoir. *M. aeruginosa* and Chlorophyll *a* almost highly from April to May, which is the summer season due to more sunlight available, lower water, temperature, and limited nutrient concentration. Moreover, the summer of 1999 is higher than in summer 2000 because of the Thailand drought in 1999

6- Funding and Acknowledgments

The author would acknowledge Phuket Rajabhat University, Thailand, and Chiang Mai University.

7- Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

8- References

[1] Sanseverino, Isabella, Diana Conduto, Luca Pozzoli, Srdan Dobricic, and Teresa Lettieri. "Algal bloom and its economic impact." European Commission, Joint Research Centre Institute for Environment and Sustainability (2016).

[2] Ahn, Chi-Yong, Myung-Hwan Park, Seung-Hyun Joung, Hee-Sik Kim, Kam-Yong Jang, and Hee-Mock Oh. "Growth Inhibition of Cyanobacteria by Ultrasonic Radiation: Laboratory and Enclosure Studies." Environmental Science & Technology 37, no. 13 (July 2003): 3031–3037. doi:10.1021/es034048z.

[3] Rodgers, John H. Jr. "Algal toxins in pond aquaculture." SRAC Publication 4605, (November 2008): 1-8.

[4] Zimba, Paul V, and Casey C Grimm. "A Synoptic Survey of Musty/muddy Odor Metabolites and Microcystin Toxin Occurrence and Concentration in Southeastern USA Channel Catfish (Ictalurus Punctatus Ralfinesque) Production Ponds." Aquaculture 218, no. 1–4 (March 2003): 81–87. doi:10.1016/s0044-8486(02)00519-7.

[5] Smith, Juliette L., Greg L. Boyer, and Paul V. Zimba. “A Review of Cyanobacterial Odorous and Bioactive Metabolites: Impacts and Management Alternatives in Aquaculture.” Aquaculture 280, no. 1–4 (August 2008): 5–20. doi:10.1016/j.aquaculture.2008.05.007.

[6] Falconer, Ian R. “Potential Impact on Human Health of Toxic Cyanobacteria.” Phycologia 35, no. sup6 (November 1996): 6–11. doi:10.2216/i0031-8884-35-6s-6.1.

[7] Codd, G. A., C. J. Ward, and S. G. Bell. “Cyanobacterial Toxins: Occurrence, Modes of Action, Health Effects and Exposure Routes.” Applied Toxicology: Approaches Through Basic Science (1997): 399–410. doi:10.1007/978-3-642-60682-3_38.

[8] Chorus, Ingrid, and Jamie Bartram, eds. “Toxic Cyanobacteria in Water” (1999). doi:10.4324/9780203478073.

[9] Sultan, Abdullah. “New Artificial Neural Network Model for Predicting the TOC from Well Logs.” SPE Middle East Oil and Gas Show and Conference (2019). doi:10.2118/194716-ms.

[10] Sinha, Ankita, and Atul Bhargav. “An Artificial Neural Network Model for Predicting Characteristic Input Parameters for Physics Based Modelling of Drying Process” (November 2, 2019). doi:10.31224/osf.io/8sxcd.

[11] Agwu, Okorie E., Julius U. Akpabio, and Adewale Dosunmu. “Artificial Neural Network Model for Predicting Drill Cuttings Settling Velocity.” Petroleum (December 2019). doi:10.1016/j.petlm.2019.12.003.

[12] Zheng, Chunlei, and Rong Xu. “Predicting Cancer Origins with a DNA Methylation-Based Deep Neural Network Model” (November 29, 2019). doi:10.1101/860171.

[13] Wei, Bin, Norio Sugiura, and Takaaki Maekawa. “Use of Artificial Neural Network in the Prediction of Algal Blooms.” Water Research 35, no. 8 (June 2001): 2022–2028. doi:10.1016/s0043-1354(00)00464-4.

[14] Lee, Joseph H.W., Yan Huang, Mike Dickman, and A.W. Jayawardena. “Neural Network Modelling of Coastal Algal Blooms.” Ecological Modelling 159, no. 2–3 (January 2003): 179–201. doi:10.1016/s0304-3800(02)00281-8.

[15] Kuo, Jan-Tai, Ming-Han Hsieh, Wu-Seng Lung, and Nian She. “Using Artificial Neural Network for Reservoir Eutrophication Prediction.” Ecological Modelling 200, no. 1–2 (January 2007): 171–177. doi:10.1016/j.ecolmodel.2006.06.018.

[16] Raee, Mohammad, and Mahsa Jahangiri-Rad. “Artificial Neural Network Approaches to the Prediction of Eutrophication and Algal Blooms in Aras Dam Iran.” Iranian Journal of Health Science 3, no.1 (February 2015): 25-32. doi:10.7508/ijhs.2015.01.004.

[17] Tian, Wenchong, Zhenliang Liao, and Jin Zhang. “An Optimization of Artificial Neural Network Model for Predicting Chlorophyll Dynamics.” Ecological Modelling 364 (November 2017): 42–52. doi:10.1016/j.ecolmodel.2017.09.013.

[18] Lee, Joseph H.W., Yan Huang, Mike Dickman, and A.W. Jayawardena. “Neural Network Modelling of Coastal Algal Blooms.” Ecological Modelling 159, no. 2–3 (January 2003): 179–201. doi:10.1016/s0304-3800(02)00281-8.

[19] Wang, Li, Xiaoyi Wang, Xuebo Jin, Jiping Xu, Huiyan Zhang, Jiabin Yu, Qian Sun, Chong Gao, and Lingbin Wang. “Analysis of Algae Growth Mechanism and Water Bloom Prediction Under the Effect of Multi-Affecting Factor.” Saudi Journal of Biological Sciences 24, no. 3 (March 2017): 556–562. doi:10.1016/j.sjbs.2017.01.026.

[20] Bui, Manh-Ha, Thanh-Luu Pham, and Thanh-Son Dao. “Prediction of Cyanobacterial Blooms in the Dau Tieng Reservoir Using an Artificial Neural Network.” Marine and Freshwater Research 68, no. 11 (2017): 2070. doi:10.1071/mf16327.

[21] Luo, Wenhuai, Huirong Chen, Anping Lei, Jun Lu, and Zhangli Hu. “Estimating Cyanobacteria Community Dynamics and Its Relationship with Environmental Factors.” International Journal of Environmental Research and Public Health 11, no. 1 (January 20, 2014): 1141–1160. doi:10.3390/ijerph110101141.

[22] Mowe, Maxine A.D., Simon M. Mitrovic, Richard P. Lim, Ambrose Furey, and Darren C.J. Yeo. “Tropical Cyanobacterial Blooms: a Review of Prevalence, Problem Taxa, Toxins and Influencing Environmental Factors.” Journal of Limnology 73, no. AoP (December 30, 2014). doi:10.4081/jlimnol.2014.1005.