Indian J. Fish., 65(4): 75-80, 2018
DOI: 10.21077/ijf.2018.65.4.72559-09

Deep-learning based approach for forecast of water quality in intensive shrimp ponds

QI LIN, WEN YANG, CHENG ZHENG, KAIHONG LU, ZHONGMING ZHENG, JIANPING WANG* AND JINYONG ZHU
School of Marine Sciences, Ningbo University, Ningbo - 315 800, China
**Ningbo Academy of Ocean and Fisheries, Ningbo, China*
e-mail: zhujinyong@nbu.edu.cn

ABSTRACT

With the enormous development of aquaculture, reducing the impacts of effluent discharge and improving water quality has become a critical global environmental concern. It is important to assess and predict water quality in the environmental management process of shrimp mariculture. Meanwhile, accurate forecast of water quality in mariculture systems is still in the initial stages at present. In this study, deep belief networks (DBN) models were used to forecast water quality in intensive shrimp culture system. This method based on deep learning includes a five-layered structure to extract relationships between the quantitative characteristic of water bodies and water quality variables. The water quality was forecasted using the Canadian Water Quality Index (WQI) obtained from the output layer of simulated model. The results show that the DBN model has a great potential to predict water quality and the ability of generalisation and accuracy of model is satisfied.

Keywords: Aquaculture, Deep belief networks, Deep learning, Shrimp culture, Water quality prediction

Introduction

Shrimp farming is one of the most important commercial activities in the coastal zone of many tropical and subtropical countries around the world (Gillett, 2008). However, deteriorating water quality has caused disease, mortality and slow growth of shrimp (Thitamadee *et al*., 2015) and has turned into one of the major obstacles to increasing yield (Ferreira *et al*., 2011). In addition, effluent discharge of cultivation ponds usually enriched in nutrients and antimicrobial agents may severely affect ecosystems adjacent to shrimp farms (Anh *et al*., 2010). If the quantitative characteristic information of water bodies can be obtained in advance, potential dangerous situations in shrimp ponds and negative effect over the ecosystem could be avoided before they appear. Therefore, predicting the water quality in intensive shrimp culture ponds is an important goal during the culture period and before sewage discharges.

Several studies have used regression approaches to construct water quality predictive models (Lopes *et al*., 2005). These models are not only time consuming but also prone to prediction errors because they are simplified into a linear process (Singh *et al*., 2009). However, water quality in aquaculture ponds is a highly dynamic and complex phenomenon. Recent studies have applied data-driven models to address water quality prediction (Gutierrez-Estrada *et al*., 2004; Li *et al*., 2012; Liu *et al*., 2012; Sun *et al*., 2012; Ma *et al*., 2014). Although data-driven techniques can deal with nonlinear and complex phenomenon successfully, most of these models, such as the artificial neural networks (ANN) have some limitations in managing the uncertainty and highly nonlinear cases through conducting only one hidden layer of framework, which is called shallow learning (Zhang *et al*., 2016). As long as sufficient neural units are added, shallow learning techniques can also approximate an arbitrary nonlinear function (Schmidhuber, 2015). Nevertheless, it will result in an extremely slow learning speed and over-fitting (Panchal *et al*., 2011). The emergence of deep learning architectures provide effective tools for forecasting highly dynamic and complex phenomena (Hinton and Salakhutdinov, 2006).

Deep learning can build a hierarchical structure like the human brain, reproducing more detailed information by learning data set from lower to higher layers. This approach has a more superior ability of representation and generalisation than traditional techniques (Le Roux and Bengio, 2008). As a typical deep learning approach, deep belief networks (DBN) have attracted wide spread attention and has gained some achievement in the fields of pattern recognition (Nie *et al*., 2015; Liu *et al*., 2016) and complex phenomena prediction (Cheng *et al*., 2017). Some studies also demonstrated that DBN-based drought prediction model (Chen *et al*., 2012) and algal bloom forecast model (Zhang *et al*., 2016) have an advantage over the shallow learning approach in respect of generalisation and accuracy.

In a shrimp culture system, numerous complex interrelated processes affect water quality and water quality variables have complicated non-linear relationships with one another. To the best of our knowledge, relatively few researchers have used deep learning as basis in studying the prediction of water quality in culture systems. This study attempted to develop a water quality forecasting model based on DBN approach through the Water Quality Index (WQI) calculated from nine environmental parameters in intensive shrimp culture ponds and to demonstrate its application to complicated non-linear relationships between input and output water variables.

Materials and methods

Sampling

The data set used in this study was generated by continuously monitoring the water quality in twelve shrimp culture ponds in Ningbo, Zhejiang Province, eastern China (29°32′N; 121°31′E). These same size ponds (ca. 2,000 m^2, 1.5 m deep) are located in greenhouses and are identically managed in terms of seawater inputs, daily water exchange rate (5%), shrimp stocking density (360,000 nos. per pond), feed type and schedule. Bottom aeration is applied to maintain a suitable level of dissolved oxygen.

The surface water samples (25-50 cm) were collected weekly. To mirror the spatial variability within ponds, samples were chosen from four representative points (similar locations among the ponds) and combined to form a composite replicate sample representing a given pond and time point. In total, we collected 708 water samples (twelve ponds × fifty-nine time points). The samples were immediately transported (within 4 h) to the laboratory in icebox.

Water temperature (WT), pH, dissolved oxygen (DO) and salinity (Sal) were recorded *in situ* with a multi-parameter probe (YSI 6000, YSI Inc., Yellow Springs, USA). The levels of ammonia (NH_4^+), nitrate (NO_3^-), nitrite (NO_2^-) and orthophosphate (PO_4^{3-}) were analysed following standard methods using an automated spectrophotometer (Smart-Chem 200 Discrete Analyzer, Westco Scientific Instruments, Brookfield, USA). Chlorophyll-*a* concentration (Chl *a*) was measured using a multi-wavelength submersible fluorescence probe (FluoroProbe, Kiel, Germany).

Water Quality Index (WQI)

The Canadian Water Quality Index (CCME WQI) proposed by the Canadian Council of Ministers of the Environment (2001) was adopted to evaluate the degree of water quality in this study. The advantage of CCME WQI includes the ability to represent measurements of numerous variables as a single number, the ability to combine various measurements in different units of measurement into a single metric value and the facilitation to communicate the results. Practitioners are also free to select appropriate water quality parameters and guidelines for their purposes therefore accommodating the specific water body (Hurley *et al*., 2012). This index has been applied to characterise the water quality for several intended uses including drinking water (Khan *et al*., 2003), surface water (Lumb *et al*., 2006) and rearing water (Ferreira *et al*., 2011). The CWQI is a mathematical instrument that compares the measurements of water quality parameters with the guidelines to produce a single score ranging from 0 to 100. Zero represents the worst water quality, whereas a score close to 100 represents excellent water quality. Based on the reference values, the recommended values for the nine water quality parameters used in this paper are listed in Table 1.

The values of WQI were calculated using the following equation:

$$\mathrm{WQI} = 100 - \left(\frac{\sqrt{F_1^2+F_2^2+F_3^2}}{1.732} \right) \tag{1}$$

where F_1 (Scope) is the number of selected variables whose objectives are not met, F_2 (Frequency) is the frequency with which the objectives are not met, F_3 (Amplitude) is the amount by which the objectives are not met. Detailed information of functions and application are described in the User's Manual of CCME (CCME, 2001).

Deep Belief Networks

A typical DBN is usually based on sequence serial training with restricted Boltzmann machines (RBMs). An RBM contains a visible and a hidden layer, with weighted connection with each other (Fig. 1). A visible layer is composed of input nodes and output nodes and every hidden layer is composed of several hidden nodes. Nodes (units per neurons), *v* in the visible layer have no connections between them and are connected to all other units, *h* in the hidden layer. The connections between nodes with weight, *w* are bidirectional and symmetric. The RBM has been mainly used as learning and generative modules that are composed to form DBN.

Each RBM extracts potential relationships behind the input data and the output of low-level RBMs is taken as

Table 1. Guideline values of water quality variables for marine shrimp culture from literature and the values recorded during the present study

Parameters	Recommended range						
	Chen (1985)	CGWQSF (1989)	Van Wyk *et al.* (1999)	Ferreira *et al.* (2011)	Carbajal-Hernandez *et al.* (2012)	Ma *et al.* (2013)	Present study
Temperature (°C)	28-33°C		28-32	>24	20-30	27-30	27-30
DO (mg l^{-1})	>3.7	>3	5.0-9.0	>3	>5	-	>3
pH	8.0-8.5	7.0-8.5	7.0-8.3	>7	6.5-9.5	7.7-8.3	7.0-8.5
Salinity (‰)	15-25		0.5-35	>15	15-23	-	15-25
Ammonia (mg l^{-1})	0.1*	≤0.02**	≤0.03**	<0.2*	0.1-1.0*; <0.1**	<0.02**	<0.2*
Nitrate (mg l^{-1})	-		≤60	<0.2	<0.5	<1	<1
Nitrite (mg l^{-1})	-		≤1	<0.7	0.4-0.8	<0.38	<0.5
Orthophosphate (mg l^{-1})	-		-	<0.2	0.1-0.3	-	<0.2
Chlorophyll-*a* (mg l^{-1})	-		-	<0.01	0.05-0.07	<0.01	0.05-0.07

*$NH_{3,4}$ - Total ammonia; ** NH_3 - un-ionised ammonia (toxic form)

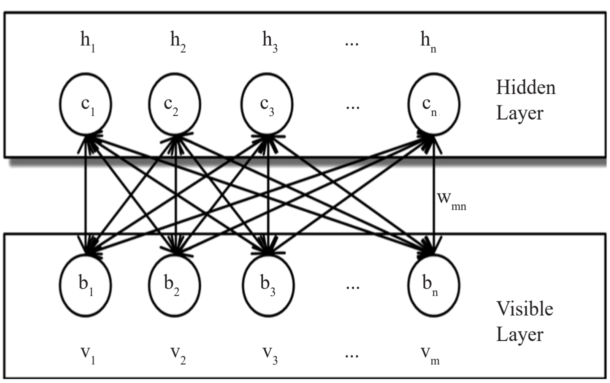

Fig. 1. Structure of a standard RBM

the input data of high-level RBMs. The greedy layer-wise unsupervised learning algorithm for DBNs which is based on sequence training with RBMs proposed by Hinton *et al*. (2006) was adopted in this study.

Prediction model based on deep learning

In our case, the water quality forecast model is based on a five-layered neural network which consisted one input layer, three hidden layers and one output layer (Fig. 2). The input layer is used for data import of water quality variables. In order to efficiently train the hidden units, a visible layer was added with three hidden layers together into three RBMs. The output layer is used for generating the WQI index, which can reflect the status of water quality. The main steps using DBN-based model for water quality forecast were as follows:

Step 1: Obtained the value of WQI by calculating the time series of water quality variables data

Step 2: Checked the statistical feature of data set. If the data were not normally distributed or did not show homogeneous variance, they were log transformed. Then the data set was divided into training data and testing data.

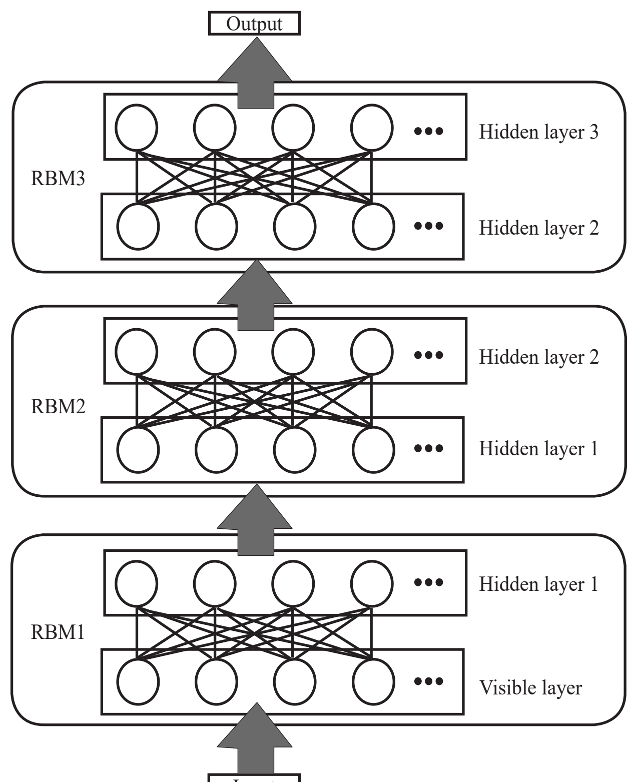

Fig. 2. Structure of the water quality forecast model based on deep learning

Step 3: Imported the training data and trained the DBN model by unsupervised deep learning. According to the criterion of smallest root mean square error (RMSE), the optimal network structure and parameters of the DBN were determined. Then a multi-layered architecture as the water quality forecast model was built.

Step 4: Applied testing data to test the proposed model. If the performance of model was able to meet the requirements, then returned to Step 3 and retrained.

Model performance criterion

In the present study, the root mean square error (RMSE) was adopted to assess the performance of the DBN model. The numbers of hidden nodes, iteration number and the learning rate of the RBM have significant effect on the performance of forecast model. Once the RMSE of the trial was the smallest, the model parameters were selected as the optimum combination.

Results and discussion

The number of input units was based on the number of water quality variables imported in the training. The proper number of nodes in hidden layers ranged from $\sqrt{n}$ to $(2n+1)$, where n was the number of input units. Thus, nineteen levels of the number of hidden units ranging from 3 to 21 were tested and 672 samples as the training data were taken to generate the proposed model. The learning rate and iteration number were fixed as 0.01 and 1000, which are the most common parameters of deep learning. When the number of units was 10 in the hidden layer, the RMSE was the smallest (Fig. 3). The numbers of hidden units were then fixed as 10 and the iteration number was fixed as 1000 and the RMSE of model in different learning rates was tested for determining the optimal learning rates. The performance of DBN was best when the learning rate was 0.01 (Fig. 4). When the hidden nodes and learning rate were fixed as 10 and 0.01, tests were done for obtaining the optimal iteration number. There was no significant fall in the value of RMSE with increasing iteration number from 1000 to 10,000, but the training time was stretched to some extent. Thus, the optimal structure of a deep neural network for water quality in this case was 9 units in a visible layer and 10 units in three hidden layers and one unit in the output layer. The learning rate and iteration number of the model training were determined to be 0.01 and 1000, respectively.

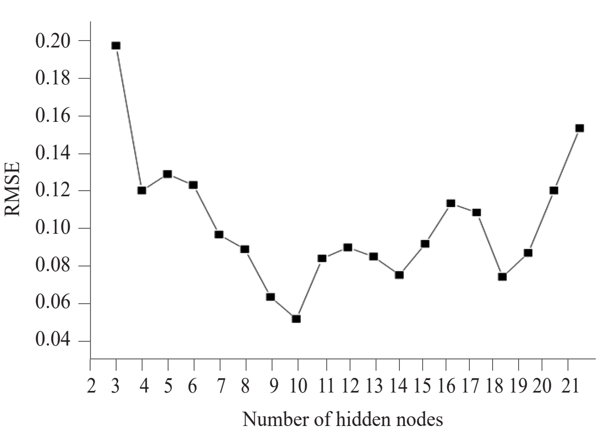

Fig. 3. RMSE of DBN in different number of hidden nodes.

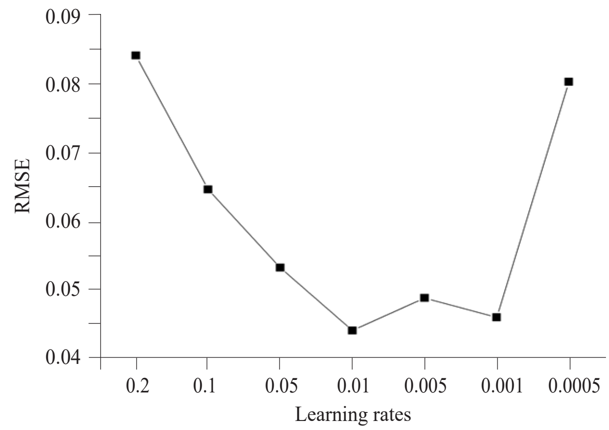

Fig. 4. RMSE of DBN in different learning rates

The comparison between measurements and forecasted value of WQI in shrimp pond are shown in Fig. 5. It is obvious that the forecast data were very close to the actual ones. The results show that the DBN model was suited to predict WQI and could obtain high precision.

In the present study, the prediction model based on DBN was used to predict the water quality of shrimp culture water. A sample of shrimp farms in eastern China was used as a test case. The predicted results showed that the DBN-based model was reliable and efficient for forecast of CWQI. The DBN model appears to be good enough in terms of generalisation and accuracy, which could meet the needs of the water quality forecast in shrimp ponds. This methodology can be useful in water quality management decisions. While the flexibility of deep belief network as a potential model for the purpose of predicting water quality in shrimp farms was demonstrated, the final aim is to prove the feasibility and practicality of application of DBN in predicting disease outbreak

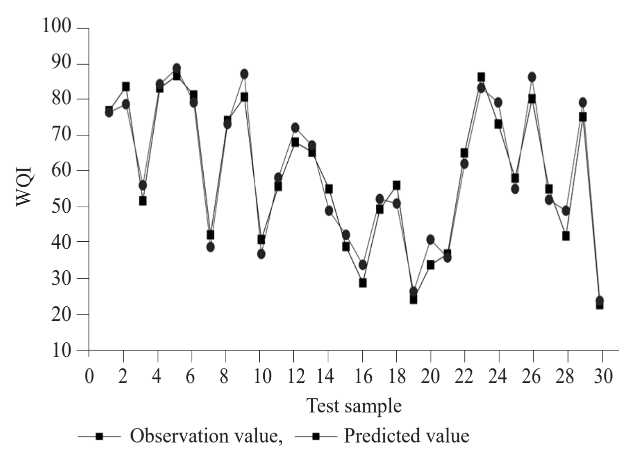

Fig. 5. Comparison of results between measurements and predicted values using DBN for different test samples

or/and shrimp production. In fact, given the complexity of miscellaneous factors affecting shrimp production, the prediction model based DBN appears to be not apt for the purpose of production forecasts. Therefore, presently we propose the deep belief network only for forecast of water quality in shrimp cultivation environment.

Acknowledgements

Our work was supported by Zhejiang Provincial Natural Science Foundation of China (No. LY17C190004) Zhejiang Provincial Undergraduate Scientific and Technological Innovation Project (No. 2017R405061), Zhejiang Province Public Welfare Technology Application Research Project (LGN19C200010) and also sponsored by the K. C. Wong Magna Fund in Ningbo University.

References

Anh, P. T., Kroeze, C., Bush, S. R. and Mol, A. P. J. 2010. Water pollution by intensive brackish shrimp farming in south-east Vietnam: causes and options for control. *Agr. Water Manage.,* 97(6): 872-882. DOI: 10.1016/j.agwat. 2010.01.018.

CCME 2001. Canadian water quality guidelines for the protection of aquatic life: CCME Water Quality Index 1.0, *User's Manual. Canadian Environment Quality Guidelines, Technical Report,* Canadian Council of Ministers of Environment, Winnipeg.

Carbajal-Hernandez, J. J., Sanchez-Fernandez, L. P., Carrasco-Ochoa, J. A. and Martinez-Trinidad, J. F. 2012. Immediate water quality assessment in shrimp culture using fuzzy inference systems. *Expert Syst. Appl.,* 39(12): 10571-10582. DOI: 10.1016/j.eswa.2012.02.141.

Chen, H. C. 1985. Water quality criteria for farming the grass shrimp, *Penaeus monodon*. In: *Proceedings of the First International Conference on the Culture of Penaeid Prawns/Shrimps*, 165 pp.

Chen, J., Jin, Q. and Chao, J. 2012. Design of deep belief networks for short-term prediction of drought index using data in the Huaihe River basin. *Math. Probl. Eng.,* 2012. http://dx.doi.org/10.1155/2012/235929.

Cheng, F., Fu, X. and Yang, Z. 2017. A short-term building cooling load prediction method using deep learning algorithms. *Appl. Energy,* 195: 222-233. DOI: 10.1016/j. apenergy.2017.03.064.

CGWQSF 1989. *Chinese National water quality standard for fisheries*. Compiling group of water quality standards for fisheries, Standards Press of China, Beijing.

Ferreira, N. C., Bonetti, C. and Seiffert, W. Q. 2011. Hydrological and Water Quality Indices as management tools in marine shrimp culture. *Aquaculture,* 318 (3-4): 425-433. DOI: 10. 1016/j.aquaculture.2011.05.045.

Gillett, R. 2008. Global study of shrimp fisheries. *FAO Fish. Tech. Paper,* 359 pp.

Gutierrez-Estrada, J. C., de Pedro-Sanz, E., Lopez-Luque, R. and Pulido-Calvo, I. 2004. Comparison between traditional methods and artificial neural networks for ammonia concentration forecasting in an eel (*Anguilla anguilla* L.) intensive rearing system. *Aquac. Eng.,* 31(3): 183-203. DOI: 10.1016/j.aquaeng.2004.03.001.

Hinton, G. E., Osindero, S. and Teh, Y. W. 2006. A fast learning algorithm for deep belief nets. *Neural Comput.,* 18(7): 1527-1554. DOI: 10.1162/neco.2006.18.7.1527.

Hinton, G. E and Salakhutdinov, R. R. 2006. Reducing the dimensionality of data with neural networks. *Science,* 313(5786): 504-507. DOI: 10.1126/science.1127647.

Hurley, T., Sadiq, R. and Mazumder, A. 2012. Adaptation and evaluation of the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) for use as an effective tool to characterise drinking source water quality. *Water Res.*, 46(11): 3544-3552. doi: 10.1016/j. watres.2012.03.061.

Khan, F., Husain, T. and Lumb, A. 2003. Water quality evaluation and trend analysis in selected watersheds of the Atlantic region of Canada. *Environ. Monit. Assess*., 88(1-3): 221-248.

Le Roux, N. and Bengio, Y. 2008. Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.,* 20(6): 1631-1649. DOI: 10.1162/neco. 2008.04-07-510.

Li, Z., Jiang, Y., Yue, J., Zhang, L. and Li, D. 2012. An improved gray model for aquaculture water quality prediction. *Intell. Autom. Soft Comput.,* 18(5): 557-567. doi.org/10.1080/107 98587.2012.10643265.

Liu, P., Choo, K. K. R., Wang, L. and Huang, F. 2016. SVM or deep learning? A comparative study on remote sensing image classification. *Soft Comput.,* 1-13.

Liu, S., Yan, M., Tai, H., Xu, L. and Li, D. 2012. Prediction of dissolved oxygen content in aquaculture of *Hyriopsis cumingii* using Elman Neural Network. *Computer and Computing Technologies in Agriculture V,* p. 508-518.

Lopes, J., Dias, J., Cardoso, A. and Silva, C. 2005. The water quality of the Ria de Aveiro Lagoon, Portugal. from the observations to the implementation of a numerical model. *Mar. Environ. Res.,* 60(5): 594-628. DOI: 10.1016/j. marenvres.2005.05.001.

Lumb, A., Halliwell, D. and Sharma, T. 2006. Application of CCME Water Quality Index to monitor water quality: A case study of the Mackenzie River Basin, Canada. *Environ. Monit. Assess.*, 113(1-3): 411-429. DOI: 10.1007/s10661-005-9092-6.

Ma, Z., Song, X., Wan, R. and Gao, L. 2013. A modified water quality index for intensive shrimp ponds of *Litopenaeus vannamei*. *Ecol. Indic.,* 24: 287-293. DOI: 10.1016/j. ecolind.2012.06.024.

Ma, Z., Song, X., Wan, R., Gao, L. and Jiang, D. 2014. Artificial neural network modeling of the water quality in intensive

Litopenaeus vannamei shrimp tanks. *Aquaculture,* 433: 307-312. doi.org/10.1016/j.aquaculture.2014.06.029.

Nie, S., Wang, Z. and Ji, Q. 2015. A generative restricted Boltzmann machine based method for high-dimensional motion data modeling. *Computer Vision and Image Understanding*, 136: 14-22. DOI:10.1016/j.cviu.2014.12.005.

Panchal, G., Ganatra, A., Kosta, Y. and Panchal, D. 2011. Behaviour analysis of multilayer perceptions with multiple hidden neurons and hidden layers. *Int. J. Computer Theory Eng.,* 3(2): 332. DOI: 10.7763/IJCTE.2011.V3.328.

Schmidhuber, J. 2015. Deep learning in neural networks: An overview. *Neural Netw.,* 61: 85-117. doi.org/10.1016/j.neunet.2014.09.003.

Singh, K. P., Basant, A., Malik, A. and Jain, G. 2009. Artificial neural network modeling of the river water quality - a case study. *Ecol. Model.,* 220(6): 888-895. DOI: 10.1016/j.ecolmodel.2009.01.004.

Sun, M., Chen, J. and Li, D. 2012. Water temperature prediction in sea cucumber aquaculture ponds by RBF neural network model. In: *Proceedings on the Conference on Water temperature prediction in sea cucumber aquaculture ponds by RBF neural network model,* p. 1154-1159.

Thitamadee, S., Prachumwat, A., Srisala, J., Jaroenlak, P., Salachan, P. V., Sritunyalucksana, K., Flegel, T. W. and Itsathitphaisarn, O. 2015. Review of current disease threats for cultivated penaeid shrimp in Asia. *Aquaculture,* 452: 69-87. doi.org/10.1016/j.aquaculture.2015.10.028.

Van Wyk, P., Davis-Hodgkins, M., Laramore, C., Main, K. L., Mountain, J. and Scarpa, J. 1999. Water quality requirements and management. In: Van Wyk, P., Davis Hodykins, M., Laramore, R., Main, K. L., Mountain, J. and Scarpa, J. (Eds.), *Farming marine shrimp in recirculating freshwater systems*. Harbour Branch Oceanographic Institution, Florida Department of Agriculture and Consumer Services, USA, p. 141-162.

Zhang, F., Wang, Y., Cao, M., Sun, X., Du, Z., Liu, R. and Ye, X. 2016. Deep-learning-based approach for prediction of algal blooms. *Sustainability*, 8(10): 1060. doi.org/10.3390/su8101060.

Date of Receipt : 26.07.2017
Date of Acceptance : 16.11.2018