

A DECISION SUPPORT SYSTEM FOR DO PREDICTION BASED ON FUZZY MODEL AND NEURAL NETWORK

Ruimei Wang¹, Qigen Liu², Youyuan He¹, Zetian Fu^{1,*}

¹ College of Economics & Management, China Agricultural University, Beijing, P. R. China, 100083

² College of Aqua-life science and technology, Shanghai Fisheries University, Shanghai, P. R. China, 200090

* Corresponding author, Address: College of Economics & Management, China Agricultural University, Beijing, P. R. China, 100083, Tel: +86-10-62736323, Fax: +86-10-62736323, Email: wangruimei@cau.edu.cn

Abstract: Dissolved oxygen (DO) concentration plays a very important role in fish life and aquaculture, but DO prediction is very difficult. So a decision support system for DO prediction based on fuzzy model and neural network was attempted. The paper was based on vast monitored data, every day detecting for two years, in aquaculture pond in North China for two years. This is a preliminary attempt towards a wider use of Artificial Neural Networks in the management of aquaculture water quality. It proposes a model to be used effectively in prediction of DO concentration in aquaculture. This is really a crucial task, especially during the long dry summer months. The prediction of potential risk due to low DO is also very important. This data volume was divided in the training subset comprising of 106 cases and in the testing subset containing 26 cases. The input parameters are sunlight, wind speed, temperature, water temperature, air pressure, pH value and NH-NH₃. Consequently three structural and seven dynamic factors are considered. After several and extended training-testing efforts a Modular Artificial Neural Network was determined to be the optimal one.

Keywords: DO, neural network, prediction, fuzzy model

1. INTRODUCTION

In recent years, many fish, shrimp and shellfish death and frequent emerging of all kinds of disease had reduced the fishermen income and brought out quality unsafety due that serious aging and eutrophicated aquaculture ponds and unbalance of ecosystem of aquaculture water (Boyd et al., 1982; Madenjian et al., 1987). So water quality management has becoming the most important element for constraining the output of the fish. DO is the most important factor in aquaculture (Madenjian, 1990; Fritz, 1985; Miller, 1990). If DO concentration sustains a low or a high value for a long time, it will lead to fish's die, to know DO concentration at anytime is very essential for fisheries, it is far from the detection of DO (Hjelmfelt et al., 1993; Jan, 1997), the key is to predict DO to prevent the risk.

Literature offers applications of different approaches to predict DO concentration in aquaculture. Some of approaches are simple and empirical in nature. Others involve solutions to equations that yield time and space dependent levels of DO. Dynamic models with hydrodynamic transport modeling components involve solution of differential equations, which formulate relevant physical, chemical and biological mechanisms and interactions as ecological formulations (Babovic et al., 1997).

The results obtained, utilizing simple empirical and statistical approaches, usually do not possess high precision. On the other hand, the calibration, validation and test steps of models that have hydrodynamic components require extensive field work and calculation efforts. Development of NN approaches with high precision of estimation was considered to be effective in eliminating the arduous task of field works and development of complex models that involve partial differential estimations. The main intentions were to minimize field work and to improve precision of estimations; most of these studies are related to primary productivity quantification in water bodies (Babovic et al., 1997; Soyupak et al., 2003).

From vast detected data, it is a complicated relation between DO and other factors in aquaculture. They have not evidently relation; neural network is feasible for these issues, but neural network is a black box, therefore, this paper combined fuzzy system and neural network (Stamou, et al., 1995; Lacroix et al., 1997; Yang et al., 1993; Guan,1996). Fuzzy system can express human experience, knowledge for computer to identify; it can also be founded some models to human feelings, language and actions; it can reasoning and deducing. Multidimensional characteristic, parallel distributed processing nature, adaptability, self-organization and self-learning can be used with fuzzy system to predict DO concentration (Wang et al., 2004; Loftus,1993; RAFI, 1992; Scott et al., 1990).

A traffic light decision system for marine finfish aquaculture siting (Hargrave, 2002); A decision support system POND which enables

definition of an entire aquaculture facility has been developed by John Bolte (2000); And Aqua Farm is a simulation and decision support for aquaculture facility design and management planning (Ernst, 2000); fish sorting (Zion et al., 2000), and automated feeding system (Papandroulakis et al., 2002). So, it can be seen that the DSS has developed a lot in aquaculture in recent years, but the development in the aspect of pond water quality management, which is the main factor for successful aquaculture, is of a small quantity.

2. DESIGN OF ARCHITECTURAL FRAMEWORK

For n input and single output fuzzy system, the fuzzy rule is:

IF : $x_1 = A_1^j$,and $x_2 = A_2^j$, \dots ,and $x_n = A_n^j$,**THEN** $y = B_j$;

$A_1^j, A_2^j, \dots, A_n^j, B_j$ are fuzzy subset. Fuzzy subset B_j was instead at the end, i. e. :**IF :** $x_1 = A_1^j$,and $x_2 = A_2^j$, \dots ,and $x_n = A_n^j$,**THEN** $y = f_j(x)$; $f_j(x)$ often take as the linear combination of input parameter, i.e. $f_j(x) = c_1^j x_1 + c_2^j x_2 + \dots + c_n^j x_n + c_j^{n+1}$

Instead of the division of input space, the below model was adopted.

IF: $X \in R_j$,**THEN** $y = f_j(X)$ (1)

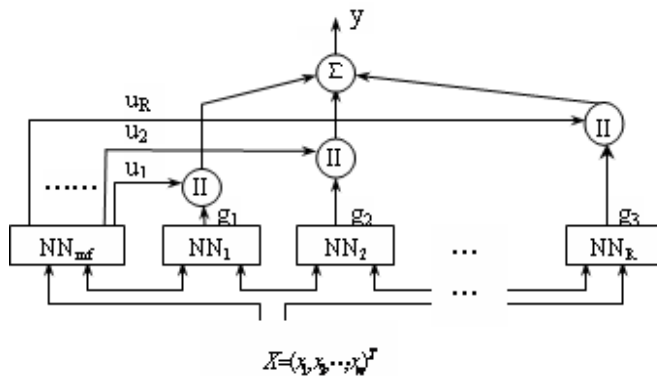


Fig.1: The architecture of fuzzy system based on neural network

Fig.1 is architecture of fuzzy system, $X = (x_1, x_2, \dots, x_n)^T$, R_j is the partial space after division. It can be given that the grade of membership of union the input parameter. At the same time, $f_j(X)$ can be describing by neural network. So the combination the neural network and the fuzzy system,

we can obtain the fuzzy system based on neural network (Meulenberg et al., 1996; Borchering et al., 1993).

In above fuzzy system, there are $R+1$ of neural network, $NN1\sim NNR$ describe $f_j(x)$ in the result of R rules, and $NNmf$ gives the rule applicability, and the output of fuzzy system was determined by follow formula.

$$y = \sum_{j=1}^R u_j \cdot g_j \quad (2)$$

3. STRUCTURE AND ARITHMETIC ANALYSIS

3.1 Structure of Neural Network Model

We choose BPNN to simulate. The network includes input layer, hidden layer and output layer. For a transportation network with n nodes and m links, the number of nodes of input layer is $n^2 - n$, and that of output layer is m , and that of hidden layer is $\sqrt{m(n^2 - n)}$. For simply transportation network, number of nodes of hidden layer should be increased appropriately. Activation function is shown as follows.

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \quad (3)$$

According to training cases value of input and out (x_i and t_i), they are real detection data. Then calculate output error in course of random weight and value, and transmit error converse to input layer to modify the weight and value to study. Suppose there are p group data for training, i.e. x_i^p and t_i^p , t_i^p is network expectation output, y_i^{p1} is the output value according to training, so to p group data, the average variance is shown as follows.

$$E = \sum_{p_1=1}^p \sum_{l=1}^{m-1} (t_l^{p_1} - y_l^{p_1})^2 \quad (4)$$

3.2 Arithmetic Analysis of the Neural Network

3.2.1 Initialization of the Neural Network

Let all joint weight to be a random value between -1 and 1. i.e. $w_{sq}(0) = \text{Random}(\bullet)$, sq is ij, jk, kl ; let all value is to be a random value between 0 and 1. Then input the samples to the network in fig.2.

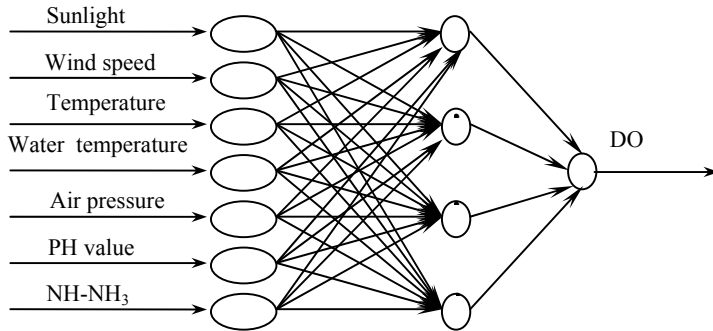


Fig.2: Neural network framework of DO prediction

3.2.2 Training of Neural Network

Provide cases $(\mathbf{x}^{p_1}, \mathbf{y}^{p_1})$, $p_1 = 1, 2, \dots, P$ for training.

Steps 1: calculate every layer value:

$$\begin{cases} y_l = f\left(\sum_{k=0}^{n_2-1} w''_{kl} x''_k - \theta_l\right) \\ f(x) = (1 - e^{-\lambda x})^{-1} \end{cases} \quad l = 0, 1, 2, \dots, m-1 \quad (5)$$

$$\begin{cases} x''_k = f\left(\sum_{j=0}^{n_1-1} w'_{jk} x'_j - \theta'_k\right) \\ f(x) = (1 - e^{-\lambda x})^{-1} \end{cases} \quad k = 0, 1, 2, \dots, n_2 - 1 \quad (6)$$

$$\begin{cases} x'_j = f\left(\sum_{i=0}^{n-1} w_{ij} x_i - \theta_j\right) \\ f(x) = (1 - e^{-\lambda x})^{-1} \end{cases} \quad j = 0, 1, 2, \dots, n_1 - 1 \quad (7)$$

In the paper, hidden layer is tow layers, the first layer has n_1 nerve cell, $x' \in \mathbf{R}^{n_1}$, $x' = (x'_0, x'_1, \dots, x'_{n_1-1})^T$; The second layer has n_2 nerver cell, $x'' \in \mathbf{R}^{n_2}$, $x'' = (x''_0, x''_1, \dots, x''_{n_2-1})^T$

Output layer $y \in \mathbf{R}^m$, there are m nerve cells, $y = (y_0, y_1, \dots, y_{m-1})^T$, the joint weight between input layer and the second layer is w_{ij} , threshold is θ_j , the joint weight between the second layer and the third layer is w'_{jk} , threshold is θ'_k , the joint weight between the third layer and output layer is w''_{kl} , threshold is θ''_l .

Step 2: calculate the error of every layer

$$\delta_{kl}^{P_1} = (t_l^{P_1} - y_l^{P_1}) y_l^{P_1} (1 - y_l^{P_1}) \quad (8)$$

$$\delta_{jk}^{P_1} = \sum_{l=0}^{m-1} \delta_{kl}^{P_1} w''_{kl} x_k^{P_1} (1 - x_k^{P_1}) \quad (9)$$

$$\delta_{ij}^{P_1} = \sum_{k=0}^{n_2} \delta_{jk}^{P_1} w'_{jk} x_j^{P_1} (1 - x_j^{P_1}) \quad (10)$$

Step 3: modify the weight and value of every layer

$$w''_{kl}(n+1) = w''_{kl}(n) + \alpha(n) \sum_{P_1=1}^P \delta_{kl}^{P_1} x_k^{P_1} \quad (11)$$

$$\theta''_l(n+1) = \theta''_l(n) + \beta(n) \sum_{P_1=1}^P \delta_{kl}^{P_1} \quad (12)$$

$$w'_{jk}(n+1) = w'_{jk}(n) + \alpha(n) \sum_{P_1=1}^P \delta_{jk}^{P_1} x_j^{P_1} \quad (13)$$

$$\theta'_k(n+1) = \theta'_k(n) + \beta(n) \sum_{P_1=1}^P \delta_{jk}^{P_1} \quad (14)$$

$$w_{ij}(n+1) = w_{ij}(n) + \alpha(n) \sum_{P_1=1}^P \delta_{ik}^{P_1} x_j^{P_1} \quad (15)$$

$$\theta_j(n+1) = \theta_j(n) + \beta(n) \sum_{P_1=1}^P \delta_{ik}^{P_1} \quad (16)$$

Step 4: according to the joint weight after modified, redo the step 1, step 2 and step 3, until for every P_1 and l , there is:

$$\left| t_l^{P_1} - y_l^{P_1} \right| < \varepsilon \quad (17)$$

Then training is over. In above formula: n —iterative time; $\alpha(n)$ and $\beta(n)$ —respectively are joint weight and study rate of threshold, commonly think them as ($0 < \alpha(n), \beta(n) > 0$); λ —shape function of Sigmoid

function ,in the paper, $\lambda =0.1$; y_i —the output value of output layer y network study; $t_i^{P_i}$ —the expectation output of the output layer y ; ε — according to the users’ precision require, choose a decimal fraction that no less than zero, in the paper, $\varepsilon =0.0028$ 。

4. FOUNDATION OF DO PREDICTION MODEL

4.1 Grouping for Data

For the veracity of the training network and validating prediction, the input data (x_i, y_i) was divided into two parts, one part is for training network, the other part is for evaluation the output result of the model. The paper adopts 106 group data as training data, and 43 group data as evaluation data.

4.2 Foundation of the Fuzzy Rule

The number of the fuzzy rule m was confirmed by fuzzy clustering, one rule will be adapt for one group data, suppose there are R group, then the number of rule is R . The precision and complexity must be considered, so the paper adopted distance square summation D in K-means, the number of group was confirmed by the decrease of D when a group increasing.

For the above 106 data, it would be standardized with the follow methods:

$$x = \frac{x' - x'_{\min}}{x'_{\max} - x'_{\min}} \tag{18}$$

x —data after standardization; x' —original data; x'_{\min} —the minimum of the original data; x'_{\max} —the maximum of the original data.

Now the data was in the scope of 0 and 1, it would be propitious to the training of neural network. And then the law absolute subtrahend was used for number grouping. The result of grouping is as [table1](#).

Table 1. The result of grouping

Group number	The number of samples for training			
1	1	2	3	47-106
2	4-46			

4.3 Training of NNmf network

NNmf was made up of n input variable and R output variable. The study sample was defined as the follows: suppose that x_i in the study sample was clustered to the S group, then:

$$w_j^i = \begin{cases} ON & j = S \\ OFF & j \neq S \end{cases} \quad j = 1, 2, \dots, R \quad (19)$$

In the fomula: On=0.9, Off=0.1

4.4 The Design of Input Layer and Output Layer and Other Parameters

The input layer is sunlight, wind speed, temperature, water temperature, air pressure, pH value and NH-NH₃. These factors are related to DO concentration according to the experts' research. Output layer is DO only. The paper chooses three layer backwards BP network model using for reference other researchers in this domain.

The cell number of hidden layer is a complex issue in neural network. In this paper, 14 cell numbers was chosen as double of ell number of the input layer, after training, if it was not successful, tow cell numbers would be added, repeat the above until it successes. Consequently three structural and seven dynamic factors are considered. Fig.2 is framework of DO prediction.

5. RESULTS

The pond simulation can be established to simulate the existing complicated pond ecosystem dynamically. And interactive modeling in real-time is the key advantage of this part.

From the systematic point of view, the pond ecosystem is comprised of a series of dynamic, nonlinear and variably interacted factors, which shows the characteristic of a fuzzy system. Accordingly, the theory of system dynamics and improved neural network were used to analysis and simulate the factors affected the pond ecosystem. Subsequently, the forecasting models, which can provide time series analysis and a substantially sophisticated view of pond ecosystem dynamics, were established to simulate DO.

Resulting from collecting real-time data from the databases and computing by the forecasting models, the trending of the simulated process will be displayed in trend curves on the computer screen. Therefore, the manager can acquire needed trending information of the key factors controlling the pond ecosystem.

The simulated results turned out that the average error was 3.7%, indicating that the models accorded with the prediction demand. Fig.3 shows the DO real-time simulated trend curves interface. Fig.4 shows the prediction interface of DO.

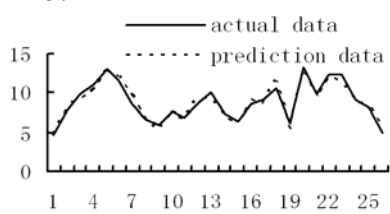


Fig.3: Comparison of real value and prediction value



Fig.4 Forecasting interface of DO

6. CONCLUSION

From Fig.3, it is known that prediction value is very close to actual value. Average error between actual data and prediction data is 0.538(mg/l), it is receivable for aquaculture. This can prove that DO prediction fuzzy system model based on neural network can be retrieved successfully.

REFERENCES

- B.T. Hargrave, 2002, A traffic light decision system for marine finfish aquaculture siting, Ocean & Coastal Management 45(2002):215 - 235.
- Babovic-V; Bartoli-N; Yang-SSY An application of artificial neural networks in computational hydraulics. Water for a changing global community. Proceedings of theme B-environmental and coastal hydraulics: protecting the aquatic habitat. Vol. 1. 27th

- Congress of the International Association for Hydraulic Research, San Francisco, California, USA, 10-15 August 1997. 1997, 865-870;
- Borcherding,-J.; Volpers,-M. The "Dreissena-monitor" - First results on the application of this biological early warning system in the continuous monitoring of water quality. International Conference on Rehabilitation of the River Rhine, Arnhem (Netherlands) 15-19 Mar 1993
- Boyd,Claude E.1982.Water Quality Management for Pond Fish Culture. Elsevier Scientific Publishing Company.pp318
- Douglas H. Ernst, John P. Bolte, 2000, AquaFarm: simulation and decision support for aquaculture facility design and management planning, *Aquacultural Engineering* 23(2000): 121-179
- Fritz,Jack J. 1985.Mathematical Models for Waste Stabilization Pond.'Mathematical Models in Biological Waste Water Treatment' Edited by Jorgensen,S. E. & Gromiec,M. J.
- Guan-BT; An artificial neural network with partitionable outputs. *Gertner-GZ. Computers-and-Electronics-in-Agriculture*. 1996, 16: 1, 39-46
- Hjelmfelt-AT Jr.; Wang-M; Heatwole-CD. Runoff hydrograph estimation using artificial neural networks. Application of advanced information technologies: effective management of natural resources. Proceedings of the 18-19 June 1993 conference, Spokane, Washington. 1993, 315-320; 5 ref
- Jan-JihnFa; Jan-JF. Artificial neural networks for classification of remote sensing data. *Quarterly-Journal-of-the-Experimental-Forest-of-National-Taiwan-University*. 1997, 11: 3, 79-89
- John Bolte, Shree Nath, Doug Ernst, 2000, Development of decision support tools for aquaculture: the POND experience, *Aquacultural Engineering* 23(2000)103-119.
- Lacroix-R; Salehi-F; Yang-XZ; Wade-KM. Effects of data processing on the performance of artificial neural networks for dairy yield prediction and cow culling classification. *Transactions-of-the-ASAE*. 1997, 40: 3, 839-846;
- Loftus-R.World Watch for domestic animal diversity released by FAO and UNEP provides "early warning system". *Diversity*. 1993, 9: 3, 34-36
- Madenjian,C.P.,Roger,G.L.& Fast,A.W.Predicting Night Time DO Loss in Prawn Ponds of Hawaii,Part II : A New Method. *Aquacultural Engineering*.6: 209-225,1987
- Madenjian,C.P.1990.Nighttime Pone Respiration Rate: Oxygen or Temperature Dependent? *Can. J. Fish. Aquat. Sci.* 47: 180-183
- Meulenberg,-E.; Stoks,-P. The application of immunochemical methods in monitoring and early warning systems for water quality control. IWSA Specialized Conference on New Developments in Modelling, Monitoring and Control of Water Supply Systems, Amsterdam (Netherlands) 24-25 Sep 1996
- Nikos Papandroulakis, Papaioannou Dimitris, Divanach Pascal, An automated feeding system for intensive hatcheries, *Aquacultural Engineering* 26 (2002) 13-26.
- Rural Advancement Fund International. Genetic engineering of pyrethrins: early warning for East African pyrethrum farmers. *RAFI-Communique*. 1992, June, 3
- Scott-MG; Hutchinson-TC; Piekarz-D. The use of lichen growth abnormalities as an early warning indicator of forest dieback. Special issue: Ecological indicators of the state of the environment. Papers presented at the workshop on ecological indicators of the state of the environment, held at the University of Toronto, Canada. *Environmental-Monitoring-and-Assessment*. 1990, 15: 3, 213-218
- Selcuk Soyupak, Feza Karaer, Hasan Gurbuz Ersin Kivra. A Neural Network-based Approach for Calculating DO Profiles in Reservoirs.*Neural Commput and Appliction* (2003) 12:166-172

- Stamou-GP; Stamou-GB; Straalen-NM-van; Krivolutsky-DA. Possible application of fuzzy system simulation models for biomonitoring soil pollution in urban areas. Bioindicator systems for soil pollution. Proceedings of the NATO Advanced Research Workshop on New Approaches to the Development of Bioindicator Systems for Soil Pollution, Moscow, Russia, 24-28 April 1995. 1996, 55-65
- W.T. Miller, et al. CMAC: An associative neural networks alternative to back propagation, Proc. Of IEEE, Vol.78, No.10, 1990
- Wang R.M., Fu Z.T., Fu LZ.: Evaluation of the Aquaculture Pond Water Quality. Land and Water Management Decision Tilos and Practices Vol-II-Proceedings of the 7th Inter Regional Conference on Environment and Water (2004)1305-1313
- Yang-ChunChieh; Prasher-SO; Lacroix-R; Sreekanth-S; Patni-NK; Masse-L; Yang-CC. Artificial neural network model for subsurface-drained farmlands. Journal-of-Irrigation-and-Drainage-Engineering. 1993, 119: 6, 947-963
- Zion, B., Shklyar, A., Karplus, I., 2000. In vivo fish sorting by computer vision. Aquacultural, Engineering 22, 165-179.