

Water Temperature Forecasting in Sea Cucumber Aquaculture Ponds by RBF Neural Network Model

Shuangyin Liu^{1,2,3}, Longqin Xu¹, Ji Chen^{2,3}, Daoliang Li^{2,3,*}, Haijiang Tai^{2,3}, Lihua Zeng^{2,3,4}

¹College of Information, Guangdong Ocean University, Zhanjiang Guangdong 524025, China;

²China-EU Center for ICT in agriculture, China Agricultural University, Beijing 100083, China;

³Beijing Engineering Research Center for Agricultural Internet of Things, China Agricultural University, Beijing 100083, China;

⁴College of Mechanical and Electrical Engineering, Agricultural University of Hebei, Banding 071001, China

hdlsyxlq@126.com

Abstract. Water temperature is considered to be the most important parameter which can largely determine the aquaculture production of sea cucumbers, so it is extremely important to monitor and forecast the water temperature at different water depths. As the change of water temperature is a complex process which can not be exactly described with a certain formula, the artificial neural network characterized by non-linearity, adaptivity, generalization, and model independence is a proper choice. This paper presents a RBF neural network model based on nearest neighbor clustering algorithm and puts forward four improved methods, then integrates them into an optimization model and verifies it on matlab platform. Finally, a comparison between the optimized RBF model and the original RBF model is made to confirm the excellent forecasting performance of the optimized RBF neural network model. This paper provides a relatively impeccable learning algorithm to complete the choice of radial basis clustering center in the process of RBF network design, and obtains a high forecasting precision so that the demand of water temperature forecasting in sea cucumber aquaculture ponds can be satisfied.

Keywords. RBF neural network; nearest neighbor clustering algorithm; sea cucumber; water temperature

Introductions

Sea cucumber is a kind of traditional Chinese seafood with high edible and medicinal value. In the process of rearing sea cucumber, water temperature is the most important

*Corresponding author, Tel:+86-10-62736764, Fax:+86-10-62737741, Email: dliangl@cau.edu.cn

parameter of the water quality. On the one hand, water temperature influences the growth rate and development of sea cucumbers as well as their distribution within the pond environment. Sea cucumber has a specific range of water temperature that it can tolerate and high water temperatures can adversely affect it by limiting its habitat or can even result in sea cucumber mortality[1],[2]. On the other hand, water temperature can directly or indirectly influence the other water quality factors, so that it can largely determine the productive capacity of aquatic ecosystems. For example, when water temperature rises, dissolved oxygen will reduce, salinity will rise, PH value will reduce and ammonia nitrogen content will rise, the balance among the water quality factors will be seriously disrupted when the water temperature is excessively high, which can result in the death and rot of sea cucumbers. Therefore, it is extremely important to monitor and forecast the water temperature at different depths in order to control the water temperature stratification during the process of sea cucumber breeding[3].

So far, the major models of forecasting the water temperature generally include stochastic models, regression models, deterministic model and Empirical models. Because water temperature is influenced by water depth, weather, aquatic activities and many other factors, the change process of water temperature can not be exactly described with a certain formula[4]. In a specific application, each type of model has advantages and drawbacks. Stochastic and regression models use statistical techniques where water temperature is related to relevant input parameters (e.g., Solar radiation, air temperature, water depth)[5],[6]. Benyahya et al. provided a review of statistical models[7]. One limitation of this modeling approach is that it explains very little of the underlying physical processes. On the contrary, deterministic models are used to predict river water temperature using a mathematical representation of the underlying physics of heat exchange between the river and the surrounding environment[8]. However, a significant drawback of these deterministic models is the amount of data required (information on hydrology and meteorology) and these data is difficult to acquisition.

Empirical models such as artificial neural networks (ANNs) have been used as a viable alternative approach to physical models[9]. It has specific features such as non-linearity, self-adaptivity (i.e., learning from inputs parameters), self-generalization, associated memory, and model independence (no a priori model needed)[10]. ANNs can learn from patterns and capture hidden functional relationships in a given data even if the functional relationships are not known or difficult to identify [10]. Using the training methods, an ANN can be trained to identify the underlying correlation between the inputs and outputs[11].

Radial basis function neural networks (RBF-NNs), in particular, have been extensively studied by researchers in nonlinear identification and water quality forecasting areas such as Dissolved Oxygen Content forecasting [12],[13] and water temperature forecasting also in financial applications such as option pricing and exchange rates forecasting [13]. In this particular type of NN, a RBF takes the role of the activation functions of the network. It has been proved that a RBF-NN can approximate arbitrarily well any multivariate continuous function on a compact domain if a sufficient number of radial basis function units are given. Guo et al. built two artificial neural network (ANN) models, a feed-forward back-propagation (BP) model and a radial basis function (RBF) model, to simulate the water quality of the

Yangtze and Jialing Rivers in reaches crossing the city of Chongqing, P. R. China, the experimental results showed that RBF neural network has characteristics such as small network size, fast learning speed, high precision and so on, which is particularly well suited for problems in which datasets contain complicated nonlinear relations among different inputs, and thus more suitable for exactly describing the forecasting process of water temperature [10],[14],[15],[16],[17].

Although it has excellent features, RBF neural network is limited in academic research and industrial applications, for the difficult points of RBF neural network design include the determination of hidden nodes number and radial basis clustering center[18],[19]. At present, the biggest deficiency of the RBF network is that there is no relatively impeccable learning algorithm to complete the choice of radial basis clustering center, which leads to a deviation on the forecasting results. To overcome the defect, we put forward an improved nearest neighbor clustering learning algorithm which four optimization schemes are joined in the existing algorithm, and a well test result is obtained.

In conclusion, the neural network model based on RBF algorithm is an advisable choice in the field of water temperature dealing with various complex physical processes. This paper first develops an adaptive RBF neural network model, then quests the optimization method and compares the results obtained by the optimized RBF model and the original RBF model, finally provides guidelines of the use of optimized method for future water temperature forecasting purposes. Wherein, the test results show that the optimized RBF algorithm can obtain the smallest forecast error with a shorter training time.

1 Materials and methods

1.1 Study area and data source

The data used in this study were produced by a Digital Wireless Monitoring System for Aquaculture Water Quality. The system has been installed at China Agricultural University-Dongying Aquaculture Digital System Research Center in Shandong province, where is near the Yellow Rive estuary. In the study, three sea cucumber ponds and a settling pond are selected as monitor objects, wherein the area of each experimental pond is about 1300m², and the water level is about 1-2m.

Because the water temperature is mainly affected by the weather factors, we develop the system to monitor water temperature, dissolved oxygen, salinity, PH value and water level of the sea cucumber ponds as well as air temperature, solar radiation, wind speed and rainfall capacity around the sea cucumber ponds. The structure diagram of the Digital Wireless Monitoring System is showed as Fig. 1.

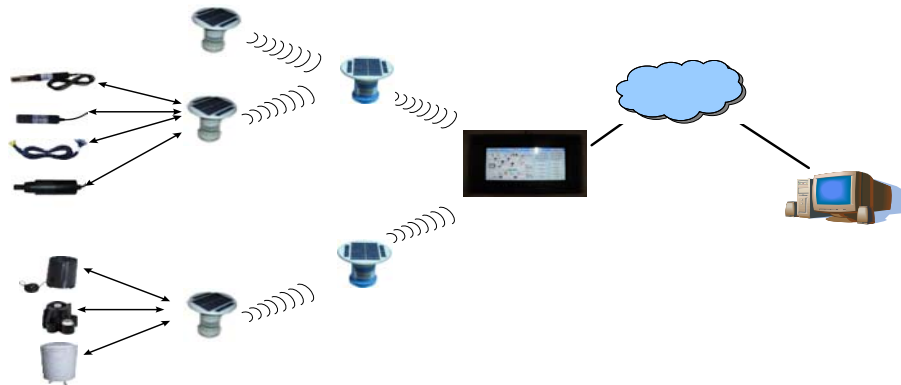


Fig. 1. Structure diagram of Digital Wireless Monitoring System

The system comprises four parts: data collection nodes, routing nodes, on-site monitoring center and remote monitoring center. Each data collection node assembles a number of sensors and a RF modem, and fixed in sea cucumber aquaculture ponds by buoys. On-site monitoring center includes on-site TPC, WSN unit and GPRS unit, remote monitoring center is a computer with fixed IP address.

The working process of the system is as follows, data collection nodes send data to the on-site monitoring center through routing nodes. On-site monitoring center collects all the data and then communicates with remote monitoring center through GPRS. From both on-site monitoring center and remote monitoring center, users can get the water quality data and weather data collected by the data collection nodes.

The data used in the paper include water temperature data at 4 different depths and real-time meteorological data in Dongying from August 7 to August 13, 2009. Data collection time interval is 10 minutes, and we use data of first 6 days for neural network training and the data of latter 1 day for test.

1.2 RBF neural network

ANN uses a multilayered approach that approximates complex mathematical functions to process data. An ANN is arranged into discrete layers each layer consisting of at least one neuron. Each node of a layer is connected to nodes of preceding and/or succeeding layers but not to nodes of the same layer with a connection weight. Thus, as the number of layers and nodes in each layer increases, the process becomes more complex demanding more computational effort. In general hydrologic and environmental problems are complex and require a complex ANN structure for prediction purposes. The number of layers and nodes in each layer is problem specific and needs to be optimized [20],[21].

Nearest neighbor clustering algorithm is an online adaptive dynamic clustering algorithm without confirming the number of hidden units in advance, which can obtain a optimal RBF network with a short learning time and a small amount of computation.

Routing node 1

⋮

Routing node n

1.3 Network optimization

By studying and analyzing the nearest neighbor clustering learning algorithm, we have found 4 deficiencies:

(1) It is not reasonable in some cases that only input information is used to determine whether a sample belongs to a clustering .

(2) It is not appropriate with a fixed clustering radius when there is a big difference in sample distribution density.

(3) In most cases, taking $C_j = \frac{1}{s} \sum_{j=1}^k x_j^k$ as a clustering is more reasonable, in

which s is the sample number of subset k.

(4) It doesn't take the learning errors as performance index to do the iterate in learning process, which will be limited on the occasions with high precision requirement.

Aiming at the defects of original algorithm, we have made corresponding improvements, the specific steps are as follows:

Step (1) Compute the distance d_{ij} (i , j =1, 2, 3, ..., N, N is the sample total) between each sample and their average \bar{d} :

$$d_{ij} = \sqrt{\|x^i - x^j\|^2 + \|y^i - y^j\|^2} \quad (1)$$

$$\bar{d} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_{ij} \quad (2)$$

Find out the nearest t samples to sample i, then set the distance as d_j respectively and work out their average distances.

$$\bar{d}_i = \frac{1}{t} \sum_{p=1}^t d_i^p \quad (3)$$

Step (2) Beginning from the first sample data set (x^1, y^1) , create a clustering center on x^1 expressed as \bar{x}^1 , then set $A(1)=y^1$, $B(1)=1$ and select a initial clustering radius r.

Step (3) Suppose the clustering center number of (x^k, y^k) is m and the center is $\bar{x}^1, \bar{x}^2, \dots, \bar{x}^m$ respectively, which means that there are m hidden units in the network as above-mentioned. Then find out the distance to these clustering centers respectively which expressed as $|x^k - \bar{x}^i|, i=1, 2, \dots, m$.

Set $|x^k - \bar{x}^j|$ as the smallest distance, the clustering j is the nearest neighbor clustering of sample x^k . Then make

$$r_k^* = \frac{r \bullet \bar{d}_k}{d} \quad (4)$$

$$\text{If } \sqrt{\|x^k - \bar{x}^j\|^2 + \|y^k - \bar{y}^j\|^2} \leq r_k^* \quad (5)$$

take the sample (x^k, y^k) into clustering j, and

$$A(j) = A(j) + y^k, B(j) = B(j) + 1, \bar{x}_{mp}^j = \bar{x}_{mp}^j + x^k \quad (6)$$

\bar{x}_{mp}^j is clustering j and the sum of inputs belong to it. Otherwise, take the sample (x^k, y^k) as a new clustering center, and set

$$\bar{x}^{m+1} = x^k, m = m + 1, A(m) = y^k, B(m) = 1 \quad (7)$$

Step (4) After determining the clustering, set $\bar{x}^i = \bar{x}_{mp}^i / B(i)$ and put learning samples into network, then get the fitting error sum of squares E,

$$E = \sum_{k=1}^N e_k^2, e_k = f^k - y^k \quad (8)$$

f^k is the output when input is x^k , at this time we can adjust the network parameters according to general gradient algorithm,

$$\sigma = \sigma - \eta \frac{\partial E}{\partial \sigma} \quad (9)$$

η is learning rate.

Step (5) The network output after learning is

$$f(x^k) = \frac{\sum_{j=1}^m A(j) \prod_{i=1}^n \exp(-((x_i - \bar{x}_i^j) / \sigma)^2)}{\sum_{j=1}^m B(j) \prod_{i=1}^n \exp(-((x_i - \bar{x}_i^j) / \sigma)^2)} \quad (10)$$

2 Implementation, results and discussion

This paper constructs a RBF neural network using five factors: air temperature, wind speed, solar radiation, water level and the previous water temperature value as network inputs to forecast the water temperature in sea cucumber aquaculture ponds.

The Gaussian function and the pure linear purelin type transformation function are separately used as hidden layer neuron and output layer neuron. And the result of this model is the water temperature value 0.5 hour ahead, the forecasting value should be in high coordination with the observed value.

In order to evaluate the performance of proposed algorithm effectively, we developed the original algorithm and the optimized algorithm on matlab platform respectively, and the forecasting results comparison diagrams are showed as follows:

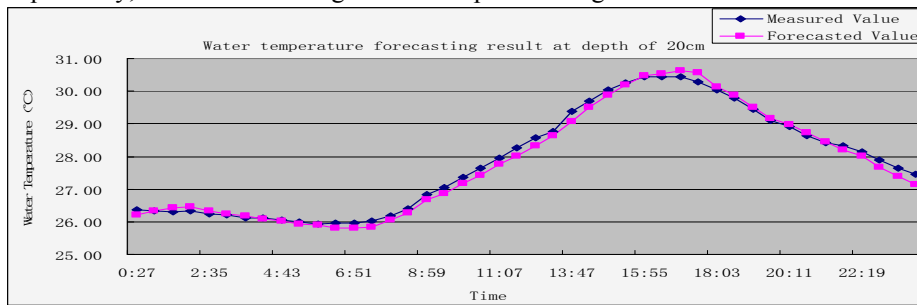


Fig. 2. Water temperature forecasting result diagram at depth of 20 cm

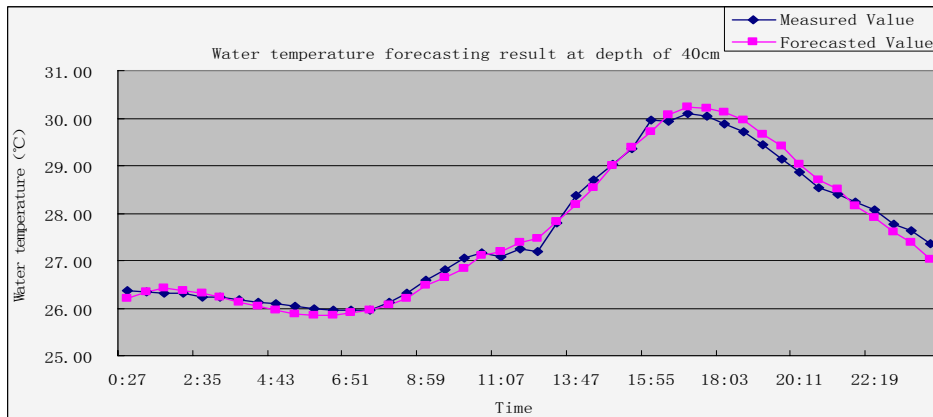


Fig. 3. Water temperature forecasting result diagram at depth of 40 cm

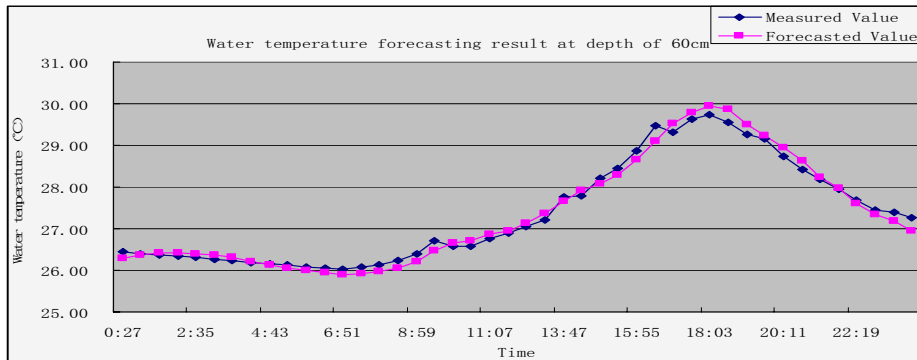


Fig. 4. Water temperature forecasting result diagram at depth of 60 cm

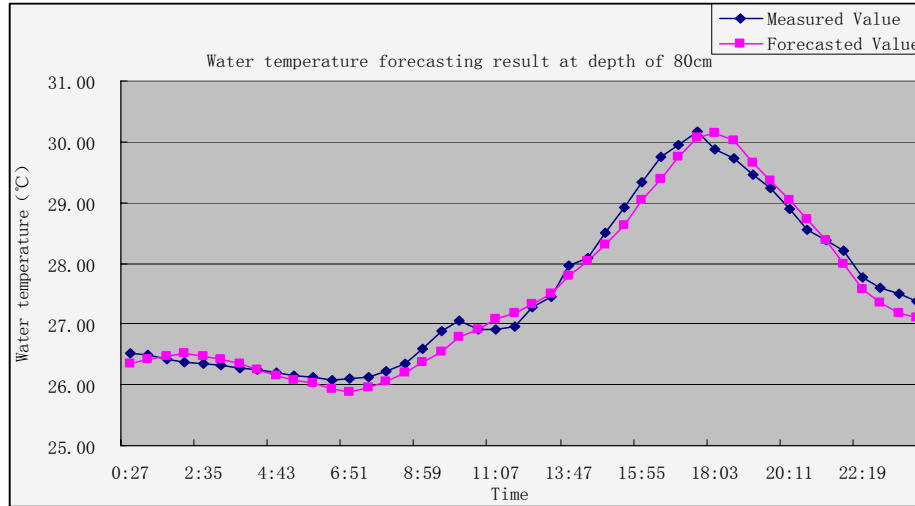


Fig. 5. Water temperature forecasting result diagram at depth of 80 cm

The forecasting performances of the optimized method is measured by using four different statistical efficiency criteria to evaluate the relative strength and weakness of the various models developed. These are mean relative error (MRE), minimal relative error, maximal relative error and root of mean square error (RMSE). Each term is estimated from the predicted and observed water temperature (targets). All of these efficiency terms are unbiased as they use error statistics relative to the observed values.

Table 1. Forecasting performance comparison of the optimized RBF network at different depths

Depth of water	MRE	Minimal relative error	Maximal relative error	RMSE
20cm	0.457%	0.029%	1.131%	0.154
40cm	0.464%	0.003%	1.168%	0.157
60cm	0.477%	0.061%	1.257%	0.160
80cm	0.580%	0.016%	1.241%	0.191

The four forecasting results comparison diagrams and the forecasting performance comparison table show that the RBF neural network can effectively forecast the change of water temperature 0.5 hour ahead, the mean square error of the forecasting results can be controlled in less than 0.2, and the maximal relative error can be controlled within 2%. Wherein the forecasting precision of RBF neural network has a slight decline with the increase of depth, this is because the input factors of neural network we used (weather factors mainly) have a gradually diminished influence on water temperature with the increase of depth. But fortunately, the tiny change does not have a big impact on forecasting result.

In addition, in order to verify the effectiveness of the optimization method, we compare the forecasting results between the original model and the optimized model at the depth of 20 cm, at the same time, the results obtained by each optimization method are also compared.

Table 2. Comparative analysis of forecasting performance efficiency of each optimization method at depth of 20 cm

<i>Method</i>	<i>MRE</i>	<i>Training time</i>
original RBF network	0.638%	7.6s
optimized RBF network by adding output information	0.597%	7.8s
optimized RBF network by adding adaptive adjustments	0.481%	6.2s
optimized RBF network by replacing the final clustering	0.574%	7.4s
optimized RBF network by gradient optimization	0.498%	7.1s
final optimized RBF network	0.457%	6.9s

Table 2 shows that the optimized network has an obviously better result than original network, these results are displayed not only in the forecasting precision, but also in training time. The final optimized RBF network makes the MRE reach a minimum with the shortening of the training time. And the optimization performances of each method are as follows:

(1) the optimized algorithm adds influence of output information in the process of judging whether a sample can form a new clustering (formula 5), which makes the determination of clustering more reasonable, and the result shows that it can improve the forecasting precision a little, but the training time is prolonged at the same time;

(2) adaptive adjustments of new clustering radius based on sample density (formula 4) can not only improve forecasting precision, but also shorten the training time greatly;

(3) we replace the final clustering input by clustering input and the input centers of samples belonging to it (formula 6 and step 4), so the clustering can reflect and replace the samples belong to it more accurately, and the forecasting precision and the training time can be simultaneously improved a little;

(4) we only adjust one-dimensional parameters in process of gradient optimization (formula 9), as a result, there is little effect on learning speed with a big forecasting increase.

(5) at last, this paper develops a forecasting network with optimal performance by integrating all the optimization modes (formula 10), and the final optimized RBF network has the minimum forecasting precision and a shorter training time.

3. Conclusion

Based on nearest neighbor clustering learning algorithm, a RBF neural network model of forecasting the water temperature in sea cucumber aquaculture ponds has been established. In order to maximise the forecasting precision and shorten learning time of network, four improved methods aiming at the defects of nearest neighbor clustering learning algorithm have been put forward. Then we integrated these methods reasonably after comparing the effects of each improved method and achieved an optimal optimization performance.

The results show that RBF neural network is a proper method for modeling a heat transfer problem due to the lack of information about internal process and boundary conditions. The optimized network has a better forecasting performance than the original network, wherein the mean square error of the forecasting results can be controlled in less than 0.2, and the maximal relative error can be controlled within 2%, thereby satisfying the demand of water temperature forecasting in sea cucumber aquaculture farms.

However, the forecasting precision of RBF neural network has a slight decline with the increase of depths, which owing to the input factors of neural network we used (weather factors mainly) having a gradually diminished influence on water temperature with the increase of depths, Fortunately, the tiny change does not have a big impact on forecasting results.

This paper proposed four improved methods, then contrasted and integrated them by experiments respectively, finally obtained an optimization algorithm with 0.181% decrease in MRE and 0.7s decrease in training time compared with the original algorithm. But restricted by the monitoring data, we can only estimate the water temperature stratification conditions by forecasted value of water temperature at the four different depths, the water temperature forecasting performance on the bottom water can not be verified, which needs to be improved in the future. Moreover, because the nearest neighbor clustering can not reflect the average value of samples, it is a feasible way to create a new kind of clustering algorithm aiming at this defect.

Acknowledgements

This research is financially supported by the National Key Technology R&D Program in the 12th Five year Plan of china (2011BAD21B01), Guangdong Science and Technology Program Project (2012A020200008) and the Zhanjiang Science and Technology Program Project (2010C3113011).

Reference

1. Tingting Ji, Yunwei Dong, Shuanglin Dong. Growth and physiological responses in the sea cucumber, *Apostichopus japonicus*, *Aquaculture* 283:180–187, (2008).

2. Longqin Xu, Shuangyin Liu. Water quality prediction model based on APSO-WLSSVR, *Journal of shangdong university(engineering science)*,42(5):80-86,(2012)
3. Fangyu Wang, Hongsheng Yang, Fei Gao, et al.. Effects of acute temperature or salinity stress on the immune response in sea cucumber *Apostichopus japonicus*, *Comparative Biochemistry and Physiology Part A* 151:491-498, (2008)
4. Juan M, Hernández, Miguel León-Santana, et al.. The role of the water temperature in the optimal management of marine aquaculture. *European Journal of Operational Research* 181:872–886,(2007)
5. Ahmadi-Nedushan, B., St-Hilaire, A., Ouarda, T.B.M.J., et al.. Predicting river water temperatures using stochastic models: case study of the Moisie River (Québec, Canada). *Hydrological Processes* 21 (1):21–34, (2007)
6. Chenard, J.F., Caissie, D..Stream temperature modeling using artificial neural networks: application on Catamaran Brook, New-Brunswick, Canada *Hydrological Processes*22(17):3361-3372, (2008)
7. Benyahya, L., St-Hilaire, A., Ouarda, T.B.M.J., et al.. Modeling of water temperatures based on stochastic approaches: case study of Deschutes River (Oregon, USA). *Journal of Environmental Engineering and Science* 6:437-448, (2007)
8. Caissie, D., Satish, M.G., El-Jabi, N.. Predicting water temperatures using a deterministic model: application on Miramichi River catchments (New Brunswick, Canada), *Journal of Hydrology* 336(3-4): 303-315, (2007)
9. Andre O. Falca, Thibault Langlois, Andreas Wichert. Flexible kernels for RBF networks, *Neurocomputing* 69:2356-2359.(2006)
10. G.B. Sahoo, S.G. Schladow, J.E. Reuter. Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models, *Journal of Hydrology* 378:325-342. (2009)
11. A.K. Palit, D. Popovic, in: *Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications*, 1st edition, Springer, (2005)
12. R. Zemouri, D. Racoceanu, N. Zerhouni. Recurrent radial basis function network for time-series prediction, *Engineering Applications of Artificial Intelligence* 16 : 453-463, (2003)
13. Honggui Han, Qili Chen, Junfei Qiao. An efficient self-organizing RBF neural network for water quality prediction, *Neural Networks* 24: 717-725, (2011)
14. Jinsong Guo, Zhe Li. Artificial neural network modeling of water quality of the Yangtze River system: a case study in reaches crossing the city of Chongqing. *Chongqing Univ: Eng Ed* 8(1): 1-9, (2009)
15. Carlos E. Romero, Jiefeng Shan. Development of an artificial neural network-based software for prediction of power plant canal water discharge temperature, *Expert Systems with Applications* 29:831–838, (2005)
16. Paweł Malinowski, Maciej Sułowicz, Janusz Bujak. Neural model for forecasting temperature in a distribution network of cooling water supplied to systems producing petroleum products, *International Journal of Refrigeration*

- 34:968-979, (2011)
17. Huang, B., Langpap, C., Adams, R.M.. Using Instream Water Temperature Forecasts for Fisheries Management: an Application in the Pacific Northwest. *Journal of the American Water Resources Association* 47:861-876, (2011)
 18. Aiming Wu , William W. Hsieh , Benyang Tang. Neural network forecasts of the tropical Pacific sea surface temperatures. *Neural Networks* 19:145-154, (2006)
 19. Awad, M., Pomares, H., Rojas, I.. Prediction of Time Series Using RBF Neural Networks: A New Approach of Clustering. *International ARAB Journal of Information Technology* 6:138-143, (2009)
 20. JunFeng Zhang; Shousong Hu. Chaotic time series Prediction based on RBF neural networks with a new clustering algorithm, *Acta Physica Sinica* 56:713-719, (2007)
 21. Lee Chengming, Chia-NanKo. Time series prediction using RBF neural networks with a nonlinear time-varying evolution PSO algorithm, *Neurocomputing* 73:449-460, (2009)
 22. Juan Ignacio Mulero-Martnez. Analysis of the errors in the modelling of manipulators with Gaussian RBF neural networks, *Neurocomputing* 72:1969-1978, (2009)
 23. Chunfei Hsu, Chienjung Chiu, Jangzern Tsai. Indirect adaptive self-organizing RBF neural controller design with a dynamical training approach, *Expert Systems with Applications* 39: 564-573, (2012)