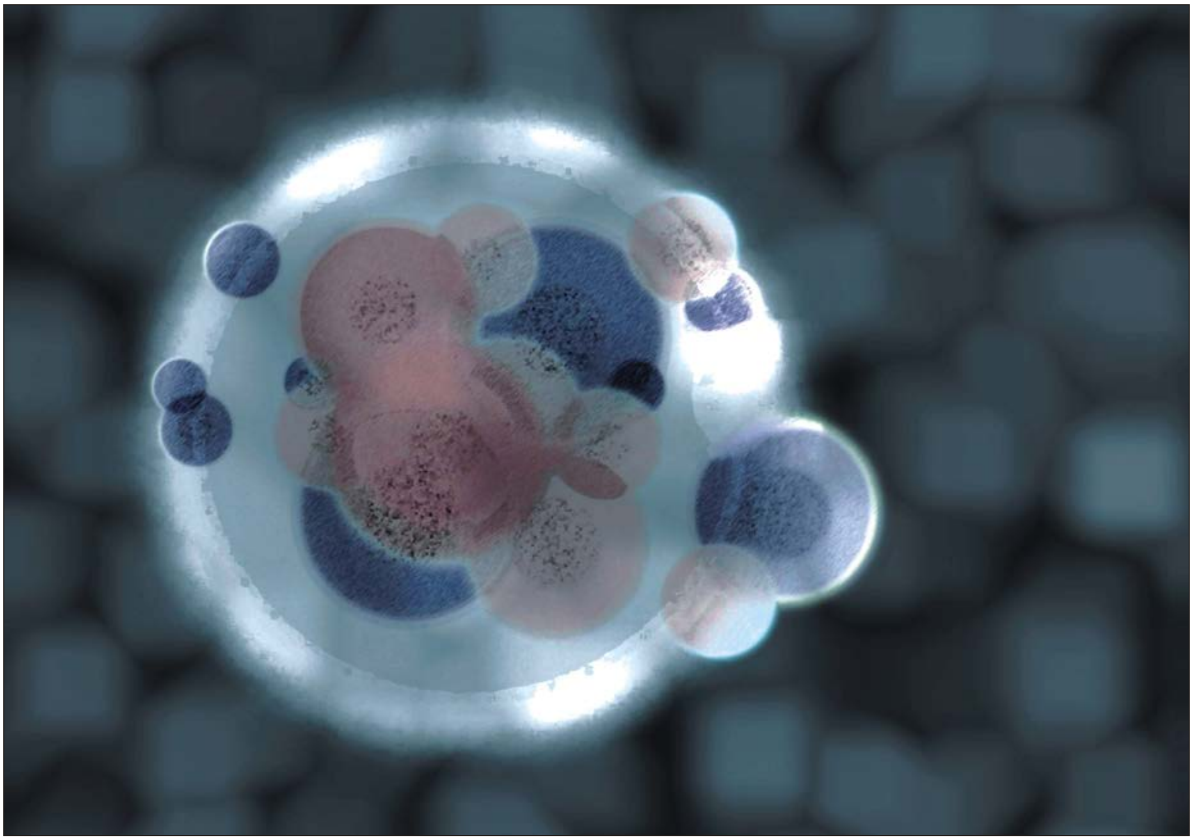

Journal of Biological Sciences

ISSN 1727-3048

Journal of Biological Sciences 6 (3): 581-585, 2006
ISSN 1727-3048

Water Quality Analysis and River Chemical Mass Transport Simulation for Shigenobu River in Shikoku Island, Japan

[1]Bin He, [2]Osamu Kaino, [1]Yi Wang and [2]Keiji Takase
[1]United Graduate School of Agricultural Sciences, Ehime University, 3-5-7 Tarumi, Matsuyama-City, Ehime, 790-8566 Japan
[2]Faculty of Agriculture, Ehime University, 3-5-7 Tarumi, Matsuyama-City, Ehime, 790-8566 Japan

Abstract: To analyze and evaluate the environmental self-purification function of river water, the chemical and hydro-meteorological data have been collected over the past 6 years in the Shigenobu river of Shikoku island, Japan. Generally, the problems of an adequate understanding and description of complex processes of pollutant circulation and water quality dynamics in rivers can be effectively solved by applying the water quality mathematical models. This approach allows taking into account, analyzing and ranking numerous interacting hydrological, meteorological and biological factors, impact of natural and anthropogenic sources of pollution. But in some cases the hydrological, meteorological and biological information are unavailable or incomplete. Thus this study presents the results of a study that examined the application of Artificial Neural Network (ANN) model to simulate the river chemical mass transport using the indicator of dissolved oxygen. Artificial Neural Network model is proved to behave well with the measured data. This study demonstrated that it was feasible to assemble and deploy ANN model to predict the dissolved oxygen under the condition of incomplete information.

Key words: Water quality, river chemical mass, dissolved oxygen; artificial neural network

INTRODUCTION

Dissolved oxygen (DO) is a measure of the free available oxygen within a body of water. Fish, invertebrates, plants and aerobic bacteria all require oxygen for respiration and adequate dissolved oxygen is necessary for good water quality (Bennett and Rathbunm, 1972). The concentration of DO in a water body directly affects the health of the water system and as such, is of primary concern for regulators. Depletion of oxygen may also result in changes in biogeochemical cycling (Naqvi *et al*., 2002). Dissolved oxygen is a direct measurement of the water quality of the river and it is one of the most important indicators of the quality of water for aquatic life. In previous study, the mass balance model was employed to analyze the DO transport phenomenon (Kaino and Ohashi, 1997; Kaino *et al*., 1998). This model enables us to establish the contaminant loading equilibrium at the inlets and outlets of a given river system. The mass balance equation for DO in a body of water must account for all material entering and leaving through direct and diffuse loading, adjective and dispersive transport and physical, chemical or biological transformation. Once the temporal variability has been determined, the spatial variability of the water body must be considered. Generally, the spatial characteristics must be homogeneous within a segment. In some cases, this restriction can be relaxed by judicious averaging over width, depth and length. Analysis of the problem should dictate the spatial and temporal scales for the modeling analysis. And the detailed database must be built from the measurement.

The development, calibration and application of a dynamic mass balance model for DO in rivers are documented with DO budget calculations in the Laboratory of Hydrology for Environment Science (LHES), Ehime University, Japan. Dissolved oxygen development for the Shigenobu River requires the application of sophisticated modeling tools to link changes in loadings to the system through management strategies to improve water quality. Because of the complexities of transport and the interaction with water quality, simplistic modeling approaches could not be used to develop a DO transport model, nor be used as management tools for rivers. A preliminary extension of the DO mass balance model is demonstrated to be successful in simulating the persistence of DO from upstream to downstream in LHES. This approach allows

Corresponding Author: Bin He, The United Graduate School of Agricultural Sciences, Ehime University, 3-5-7 Tarumi, Matsuyama-Shi, Ehime, 790-8566 Japan Tel: (+81) 80-6392-1840

taking into account, analyzing and ranking numerous interacting hydrological, meteorological and biological factors, impact of natural and anthropogenic sources of pollution. But in some cases, the hydrological, meteorological and biological information are unavailable or incomplete. On the other hand the collection of water quality data is often time consuming and costly. The ideal way is to use available observation data that is not costly for a calibrated and validated model. The aforementioned reasons have all contributed to the motivation for introducing an artificial neural network (ANN) model in the analysis of pollutant simulation.

ANN has extracted considerable research interests of engineers and scientists in a wide variety of study areas and various types of networks have been successfully applied in various fields including water resources. It is particularly useful as its simplicity, ability to capture any a prior unknown non-linear relationships among variables and having less assumption in its applications (Hsu *et al*., 1995). The major advantage of an ANN is its ability to represent the non-linearity by means of smaller number of parameters and to learn from examples (Heggen, 1995; Pal and Srimani, 1996). The use of ANN modeling has been increasing in various aspects of sciences and engineering (Chow and Leung, 1996; Duan *et al*., 1992; Lamedica *et al*., 1996; Rogers and Dowla, 1994). ANN has also been used in many hydrological problems such as forecasting reservoir inflow and river flow prediction (Coulibaly *et al*., 2000; Karunanithi *et al*., 1994), determining aquifer parameters (Aziz and Wong, 1992), water quality parameters (Maier and Dandy, 1996), etc. The ANN modeling can be seen as a sophisticated data-oriented modeling technique to find the relationship between the input and output patterns without using the detailed process knowledge. As stated earlier, the ANN has the ability to represent the non-linearity by means of smaller number of parameters and least requirement of information regarding the process to be modeled. But until now the application of ANN in river chemical mass transportation is very few. Thus in this study an ANN model is assessed to simulate the time series river chemical mass transportation for a river in Shikoku island, Japan. The study site, Shigenobu river, is one of the largest rivers in the Shikoku island and it discharges to the southern part of the Seto Inland Sea. In this decade there has been a concern about the DO conditions in the Shigenobu River estuary and adjacent coastal waters. In this study, an ANN prediction model was used to simulate the water quality in the Shigenobu river by using an indicator of DO. The simulated results will be compared with the measured DO.

MATERIALS AND METHODS

Site description: The Shigenobu River is located at Dogo plain in West Shikoku Island of Japan (Fig. 1). The study area extended from about 4.5 km upstream of Shigenobu River to the downstream near the Seto bay. Shigenobu River's water quality is highly influenced by its natural geographic location, weather patterns of the surrounding watershed and human uses. There is a significant gap in water quality data in Shigenobu River for calibration of a water quality model. The problems of an adequate understanding and description of complex processes of circulation and water quality dynamics in shallow density stratified estuaries can be effectively solved by applying the hydrodynamic and water quality mathematical models. Changing and diversifying land use practices that include forestry, agriculture, industry and urban development have placed increasing pressures on estuaries and coastal habitats to accommodate anthropogenic inputs.

The present study thus aims to assess the self-purification function of water in Shigenobu River. Physical/chemical properties (temperature, conductivity, DO and pH) of the water column were sampled at mid-river stations by LHES from year 2000 to date. Sampling was conducted using YSI-Model 6000. Stream flow is needed to compute constituent loads. Stream flow measurements are made at all sampling sites and are related to the surface water elevation. A rating curve of this relation is developed for each site. By recording gage height during each sampling event, the stream flow is determined and instantaneous loads for measured constituents are computed. The data of DO includes historical data from November 2000 to date.

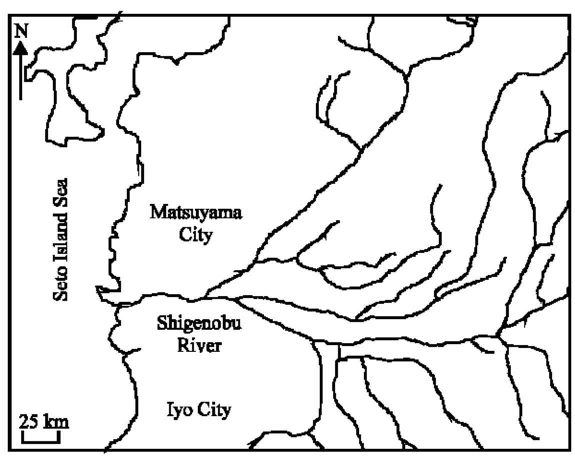

Fig. 1: Location of study site showing the Shigenobu River-the main river of Matsuyama City

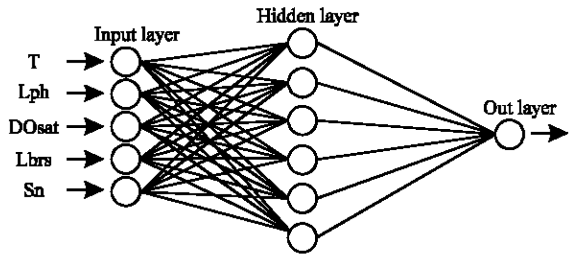

Fig. 2: ANN three-layer architecture used in this research. The input layer is the known datum of downstream and upstream gauges and the output layer is the unknown or predicted datum of DO T(Temperature), L_{ph} (Oxygen supplied by photosynthesis), Do_{sat} (Saturated density of DO), L_{brs} (Oxygen consumed by Algae breath) and S_n (Sunshine)

ANN model: Artificial neural network is an information processing system which is composed of simple processing elements linked by weighted connections. The ANN technique mimics the cognitive response of the human brain. It is a flexible mathematical structure patterned after the biological functioning of the nervous system and considered to be a versatile tool for approximating complex functions that are difficult to model mathematically or evaluate numerically. It is particularly useful for its simplicity, ability to capture any a priori unknown non-linear relationships among variables and having less assumption in its applications (Hsu *et al*., 1995). The ANN modeling can be seen as a sophisticated data oriented modeling technique to find relationships between input and output patterns without using detailed process knowledge. The ANN has the ability to represent non-linearity by means of a smaller number of parameters and least requirement of information regarding the process to be modeled. ANN method can be quite useful in many practical hydrological studies, as most of the hydrological processes are complex and entail heavy empiricism.

In this study, a typical three-layer network with an input layer (I), hidden layer (H) and an output layer (O) (Fig. 2) is adopted. The input layer includes the measured data of T, L_{ph}, DO_{sat}, L_{brs} and S_n in the upstream of the Shigenobu River. The hidden layer could be decided by calibration process and includes six neurons in this study. The output layer is the predicted DO in the downstream of the Shigenobu River. Each layer consists of several neurons; layers are interconnected by correlation weight sets. Ideally, the training set should contain significant observations of all possible ranges of the variables. The whole ANN model is just similar to a black box model which will give the output data after obtaining the input data and model parameters. It simplified the sophistic physical process and will be more useful with few or incomplete available data.

Data preparation and standardization: The type of data can include the measurement data of DO, meteorological data, precipitation data, discharge data, etc. In this study, the input layer includes the measured data of T, L_{ph}, DO_{sat}, L_{brs} and S_n in the upstream of the Shigenobu River. After obtaining the measurement data as input layer, all are standardized as data preprocess. The sigmoid function was used to result in a better normalized mean squared error (NMSE) for all the randomized data sets. Due to the nature of the sigmoid function, it was necessary to standardize the data in a range between 0 and 1. The ANN would require extremely small weighting factors causing computational inaccuracies due to floating point calculations, sluggish training and the near-zero gradient of sigmoid function at extreme values (Dawson and Wilby, 1998). Therefore, input values in the present study were standardized with respect to the maximum and minimum values in the range, as this provided better model predictions than other standardization.

$$\overline{X_i} = \left(X_i - X_{i\min}\right) \Big/ \left(X_{i\max} - X_{i\min}\right) \quad (1)$$

Where X_i is the respective standardized value for the node; X_i is the actual value of node I; $X_{i\min}$ and $X_{I\max}$ are the maximum and minimum of all values applied to the node, respectively.

ANN training and evaluation: ANN is trained with a set of known input and output data and is similar to the calibration of conceptual models. The training process was repeated with different sets of shuffled data. The NMSE was noted for each analysis and cross validation was done to estimate the R^2 values. The learning process was terminated when an optimum prediction statistic was obtained in relation to epoch size and cross-validation results. In this study, the sigmoid transfer function was used:

$$f(x) = 1 \Big/ \left(1 + e^{-x}\right) \quad (2)$$

The model efficiency factor (E), coefficient of determination (R^2) and absolute average deviation (AAD) between the observed and the predicted values were estimated for different prediction object. The efficiency factor was estimated for all the validation sets using the normalized mean squared error (NMSE) which is determined from:

$$E = \frac{\sum_{n=1}^{N}\sum_{p=1}^{P}\left(D_{np} - Z_{np}\right)^2}{\sum_{n=1}^{N}\sum_{p=1}^{P}\left(D_{np} - \overline{D_{np}}\right)^2} \quad (3)$$

Where P is the total number of input elements; N is the total number of output elements; D_{np} is the target outputs in transformed values; Z_{np} is the network's outputs in transformed values. The NMSE is the sum of squared errors normalized by the number of patterns over all output nodes and the estimated variance of the data. If subtracting the NMSE by unity, the result would be a statistic similar to the coefficient of determination (R^2) of the ANN model. Thus, from the above, once the training process was satisfactorily completed, the network was saved. The test and validation data sets were recalled and the model-predicted values were compared with observed values.

RESULTS AND DISCUSSION

The simulated DO concentrations showed no significant sensitivity to the changes in model's input rate kinetics. In Fig. 3, the calculated and measured DO was shown in different color and line type. Results showed that the ANN model forecasts the magnitude and timing of both big peaks and smaller peaks quite well. Also the scatter plots showing the observed historical DO on the X-axis against the forecasted DO from the ANN model on the Y-axis are displayed in Fig. 4 for the correlation comparison. The graphical representation of the observed and calculated DO values (Fig. 4) shows that the calculated values are closely correlated to the observed values, with few outlying points for ANN modeling. According to Fig. 4, the correlation coefficient is 0.935, which is reasonably good.

The Frequency Distribution (FD) of hourly difference between hourly values of DO predicted by ANN model and measured DO is shown in Fig. 5. It illustrates that differences greater than 0.1 g m^{-3} were found in less than 21% of the total observations; differences greater than 0.7 g m^{-3} were found in less than 2.3% of the total observations. This analysis indicates that hourly agreements between hourly values of DO predicted by ANN model and measured DO was good in spite of the departures when the difference was greater than 0.1 g m^{-3}.

The simulation result demonstrates the compatibility of the proposed ANN model in the learning and supplement of DO. The prediction of the present model overall agreed well with the observation of DO. It can be concluded that ANN is reliably accurate at forecasting the DO values. Thus the reliable DO value

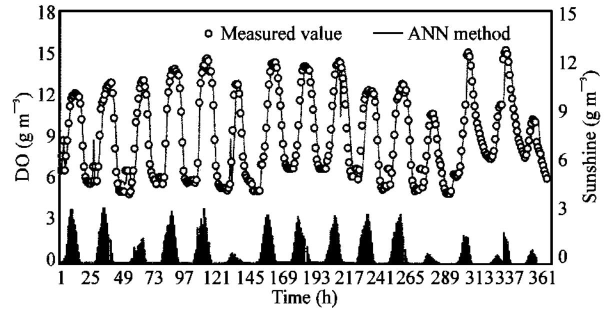

Fig. 3: Comparison of the observed DO and predicted DO. Prediction of DO was based on the ANN model. Column lines in the above graph correspond to the observed sunshine data

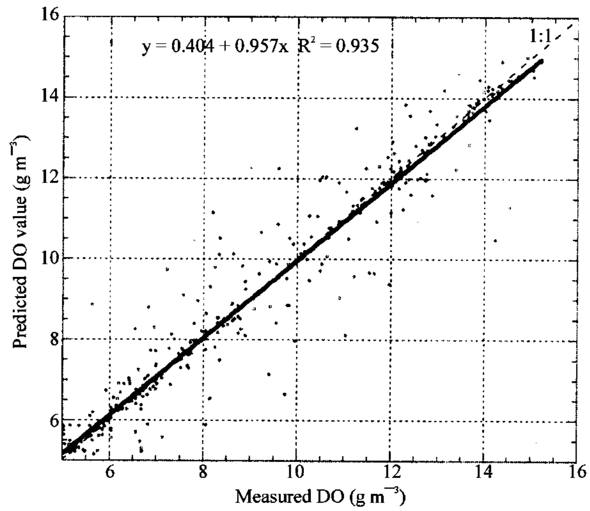

Fig. 4: Comparison of the observed DO and predicted DO. Scatter figure between the measured and calculated dissolved oxygen data using ANN Model

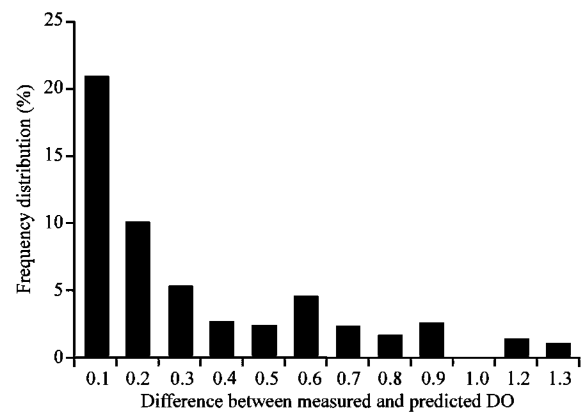

Fig. 5: Frequency distribution (*FD*) of hourly difference between hourly values of predicted DO by ANN model and measured DO (Unit: g m^{-3})

would be predicted by using a relative fewer input database comparing with the complex mass balance model. In that case, the ANN model would save us more time and produce more promising prediction results if the limited measurement database is only available.

In this study, the DO of the Shigenobu River was measured by using the meteorological conditions and hydro-chemical conditions. The estimation of DO which was presented in this study was based on the few available data of the Shigenobu River by using the proposed ANN model. It was dependent on the variable parameters for each data type of the input database. The results of the ANN model illustrated that the ANN is accurate at forecasting the DO values. It can be seen that the ANN model is capable to model the outputs of a river chemical mass transport model with a high accuracy. The ANN model proved an excellent, indeed improved method for studying the water chemistry and mass transport for river system. The Shigenobu River case study gave us one more demonstration of successful applicability of ANN modeling tool to analyze the complex hydrological-ecological processes in the river system.

ACKNOWLEDGMENTS

This research has been carried out under the financial support by the Grants-in-Aid for Scientific Research (Grant No. 09660260) of the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT). The authors are grateful for their supports. Thanks will also be given to the graduate and undergraduate students of Laboratory of Hydrology for Environment Science at Ehime University for their assistance with the hydrological field works, data analysis and data management.

REFERENCES

Aziz, A.R.A. and K.F.V. Wong, 1992. A neural network approach to the determination of aquifer parameters. Groundwater, 30: 164-166.

Bennett, J.P. and R.E. Rathbun, 1972. Reaeration in Open-Channel Flow. U.S.G.S., Professional Paper, 73.

Chow, T.W.S. and C.T. Leung, 1996. Neural network based short-term load forecasting using weather compensation. IEEE Trans. Power Sys., 11: 1736-1742.

Coulibaly, P.P., F. Anctil and B. Bobee, 2000. A recurrent neural networks approach using indices of low-frequency climiatic variability to forecast regional annual runoff. CGHU Special Issue of Hydrological Process, 15: 2755-2777.

Dawson, C.W. and R. Wilby, 1998. An artificial neural network approach to rainfall-runoff modeling. J. Hydrol. Sci., 43: 47-66.

Duan, O., S. Sorooshian and V.K. Gupta, 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res., 28: 1015-1031.

Heggen, R.J., 1995. Discussion on Neural Networks for River Flow Prediction By Karunanidhi *et al.*, J. Comp. In Civ. Eng. ASCE., 9: 293.

Hsu, K., H.V. Gupta and S. Sorooshian, 1995. Artificial neural network modeling of the rainfall-runoff process. Water Resour. Res., 31: 2517-2530.

Kaino, O. and G. Ohashi, 1997. Study on the mitigation of pollutant mass for the shallow marine. Ann. Proc. JSIDRE., pp: 192-193.

Kaino, O., G. Ohashi and M. Fujihara, 1998. Study on the mitigation of pollutant mass for the shallow marine. Ann. Proceeding of JSIDRE., pp: 290-293.

Karunanithi, N., W.J. Grenney, D. Whitley and K. Bovee, 1994. Neural networks for river flow prediction. J. Comp. Technol. Civil Eng. ASCE., 2: 201-220.

Lamedica, R., A. Prudenzi, M. Sforna, M. Caciota and V. Orsolini, 1996. A neural network based technique for short-term forecasting of anomoous load periods. IEEE Trans. Power Sys., 11: 1749-1756.

Maier, H.R. and G.C. Dandy, 1996. The use of artificial neural networks for the prediction of water quality parameters. Water Resour. Res., 32: 1013-1022.

Naqvi, S.W.A., D.A. Jayakumar, P.V. Narvekar, H. Naik, V.V.S.S. Sarma, W. DeSouza, S. Joseph and M.D. George, 2000. Increased marine production of N_2O due to intensifying anoxia on the Indian continental shelf. Nature, 408: 346-349.

Pal, K.S. and P.K. Srimani, 1996. Neurocomputing- guest editors. Introduction, Computer, 29: 24-28.

Rogers, L.L. and F.U. Dowla, 1994. Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resour. Res., 30: 457-481.

Wool, T.A., S.R. Davie and H.N. Rodriguez, 2003. Development of three-dimensional hydrodynamic and water quality models to support total maximum daily load decision process to the Neuse River Estuary, North Carolina. J. Water Resour. Plan. Manage., 129: 295-306.