Journal of Engineering Science and Technology
Vol. 7, No. 4 (2012) 447 - 461

PREDICTION OF WATER QUALITY INDEX USING BACK PROPAGATION NETWORK ALGORITHM. CASE STUDY: GOMBAK RIVER

FARIS GORASHI[1,*], ALIAS ABDULLAH[2]

[1]School of Engineering and Technology Infra-Structure (SETI), Infra-Structure University Kuala Lumpur (IUKL), Malaysia
[2]Faculty of Architecture and Environmental Design, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia
*Corresponding Author: faris@iukl.edu.my

Abstract

The aim of this study is to enable prediction of water quality parameters with conjunction to land use attributes and to find a low-end alternative for water quality monitoring techniques, which are typically expensive and tedious. It also aims to ensure sustainable development, which is essentially has effects on water quality. The research approach followed in this study is via using artificial neural networks, and geographical information system to provide a reliable prediction model. Back propagation network algorithm was used for the purpose of this study. The proposed approach minimized most of anomalies associated with prediction methods and provided water quality prediction with precision. The study used 5 hidden nodes in this network. The network was optimized to complete 23145 cycles before it reaches the best error of 0.65. Stations 18 had shown the greatest fluctuation among the three stations as it reflects an area of on-going rapid development of Gombak river watershed. The results had shown a very close prediction with best error of 0.67 in a sensitivity test that was carried afterwards.

Keywords: Water quality index, Land-use, ANN, Back propagation.

1. Introduction

In Malaysia there are 189 river basins nationwide. Rivers in Peninsular Malaysia are highly diverse ecosystems and support extensive artisanal fisheries. More importantly, rivers are the main source of drinking water in the peninsular. Surface water in the form of streams and rivers with or without reservoirs / impoundments

Nomenclatures

b_1^j and b_2^k	Bias terms
$f_1(.)$ and $f_2(.)$	Activation functions
$G : \mathbb{R}^N \to \Omega \subseteq \mathbb{R}^N$	Nonlinear activation mapping
g_i	Activation function of neuron i
N	Number of neurons in the networks.
OPk	Output from the k^{th} node of the output layer of the network for the P^{th} vector
u_i	External input imposed on neuron i
$W = (w_{ij})_{N\times N}$	Synaptic weight matrix
w_{ij}	Synaptic connectivity value between neuron i and neuron j
w_{ij}^h	Connection weight between the i^{th} node of the input layer and the j^{th} node of the hidden layer
w_{jk}^o	Connection weight of the communication strand between the j^{th} node of the hidden layer and the k^{th} node of the output layer
xPi	Input to the network for P^{th} vector
$x = (x_1, x_{12}, \ldots.., x_N)$	Neuron states
$y = (y_1, y_{12}, \ldots.., y_N)$	Local fields

Greek Symbols

Ω	Convex subset of $\mathbb{R}^N$.

contribute about 97% of raw water supply sources where groundwater is not widely used due to its limited availability. Out of 189 river basins, 120 rivers are being monitored by the department of environment (DOE). There are 926 monitoring stations for these rivers. According to the department of environment 44.5% are clean, 48.4 % slightly polluted and 7.1% are polluted. Generally, stations located upstream are usually clean, while those located downstream are either polluted or slightly polluted.

In the Klang basin in Selangor all rivers (Ampang, Batu, Damansara, Jinjang, Kerayong, Keroh, Klang, Kuyoh and pencala) are polluted according to DOE, except Gombak River which is slightly polluted. The main cause of pollution of these rivers basin is overdevelopment on the rivers' catchment area.

1.1. Applications of neural network models for water quality

In recent years, artificial neural networks were successfully applied in the area of water quality modeling. The use of ANN model was to be better than other simulations and commonly used statistical models [1] due to the complex inter-related and non-linear relationships between multiple parameters. However, modeling applications for water quality response due to Land use attributes are generally more difficult due to the complexities in environmental distribution, mobility and number of point &non-point sources of waste discharge.

Junaidah et al. [2] concluded that model derived using MLR technique gave a better prediction than the model derived using ANN in a study on sediment prediction, However, this statement can be debatable depending on the complexity of the model itself. Water quality responds to myriad stimuli and reactions. Many chemical constituents are involved either naturally or synthetically. The model of which cannot be treated in a linear manner. ANN is intended to be used with problem of complexity. A later study by Stewart [3] revealed that a multiple linear regression may be viewed as a special case ANN model that uses linear transfer functions and no hidden layers. If the linear model performs as well as a more complex ANN, then using the nonlinear ANN may not be justified, however, the optimization of the ANN model revealed a markedly better prediction than the MLR model in a study to predict the concentration of dissolved oxygen in a river. In addition multiple linear regression models failed to capture the long term patterns, however, ANN model was successful in predicting those patterns [4].

Zou et al. [4] proposed a neural network embedded Monte Carlo approach to account for uncertainty in water quality modeling. This could be quite difficult to attain, as the behavior of artificial neural networks is unpredictable. However the rule behind the selected trained model could be extracted and used in a stochastic approach to similar models.

The traditional statistical methods are rarely used and considered in ANN model building processes. In neural network, usually model-building processes are described poorly which makes it difficult to assess the optimization of output results. Ripley [5] had recommended the inclusion of statistical principles in the ANN model building in order to improve the performance of the model.

A study made by Kamarul and Ruslan [6] concluded that ANN could become a useful modeling method, as alternative to actual data collection, thus is the best choice to the government to manage water resources issues. The study uses existing raw water quality data from official sources and related to eight land uses categories i.e. residential, industrial, commercial, public utilities, recreational, institutional, and forest. In this study the authors used the simplified Fuzzy adaptive resonance theory map [SFAM]. The SFAM logic output was according to classes of water quality. Although the findings were well correlated, nevertheless, the wide range of water quality classes will not give an accurate forecast in terms of a particular water quality parameter. Using a different approach such as back propagation could have given more précised findings.

1.2. Artificial neural network (ANN)

The concept of artificial neurons was first introduced in 1943. Artificial Neural Network is a network of interconnected elements. These elements were inspired from studies to simulate the biological brain. The purpose of Artificial Neural Network is to Learn to Recognize Patterns in ones data. Once the neural Network has been trained on samples of data, it can make Predictions by detecting similar patterns in future data Cormac [7], and Picton [8]. Mostly used neural networks are SFAM algorithm, where one attempts to predict the class or category for a given pattern (Fig. 1). This method is good for predicting water quality index and class based on Malaysian National Interim Water Quality. A Typical SFAM Architecture could be illustrated as follows:

Other widely used neural network is the Back Propagation Neural Networks and sometimes called Feedback Network (Fig. 2). This method can predict the variable quantity with high precision, for example the concentration of a certain parameter. The Algorithm makes its prediction as numeric values, not as class names. It is best suited for predicting continuous numerical values such as water quality data. Artificial neural network is mathematical structure to mimic the information processing functions of a network of neurons in the brain.

Stewart [3] mentioned that ANN is particularly well suited for problems in which large datasets contain complicated non-linear relations among many different inputs.

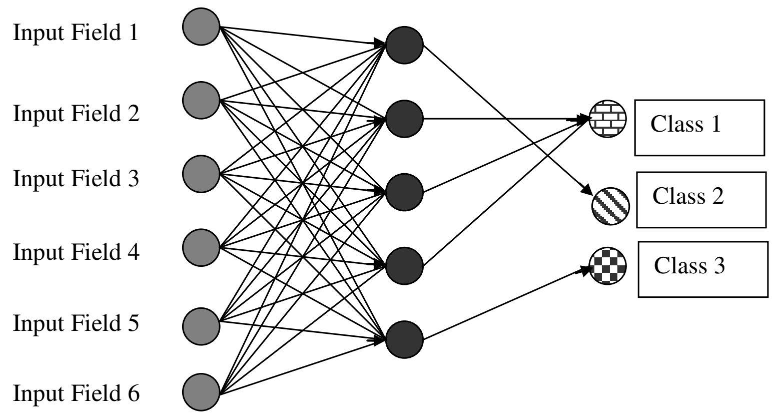

Fig. 1. Typical SFAM Architecture.

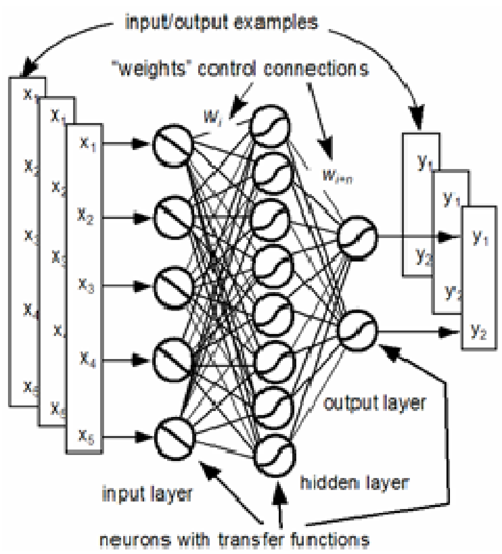

Fig. 2. Feedback Network Artificial Neural Network Architecture.

Training an ANN is a mathematical exercise that optimizes all of the ANN's weight and threshold values using some fractions of the available data. Neural networks serve to provide researchers with empirical models of complex system from which they can begin to unravel the underlying relationships and come to a more complete understanding of the environment.

Brion [9] expressed neural network mathematically as, a three-layer neural network with I Input nodes, J hidden nodes in a hidden layer and K output nodes can be expressed as shown in Eq. (1)

$$O_{Pk} = f_1\left\{\sum_{j=1}^{L} w_{jk}^{o} f_2\left(\sum_{i=1}^{N} w_{ij}^{h} x_{pi} + b_1^k\right) + b_2^k\right\} \quad \forall k = 1, 2, ..., k \tag{1}$$

The most commonly used activation function within the nodes is the logistic sigmoid function, which produces output in the range of 0–1 and introduces non-linearity into the network, which gives the power to capture non-linear relationships between input and output values. The logistic function as shown in Eq. (2) was used in this work in the form given below.

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

While many statistical and empirical models exist for water quality prediction, artificial neural network (ANN) models are increasingly being used for forecasting of water resources variables because ANNs are often capable of modeling complex systems for which behavioral rules are either unknown or difficult to simulate. Juahir et al. [10] mentioned "previous studies have shown that ANN models perform well in predicting short and long-term environmental data". Loke et al. [11] concluded that ANNs can deal with problems that are traditionally difficult for conventional modeling techniques to solve. Their advantages include good generalization abilities, high fault tolerance, high execution speed and the ability to adapt and learn.

2. Case Study

The study area of this research is Gombak River. Gombak River is a slow flowing river, which originates from many tributaries in the Gombak district. The river has several confluences with other streams such as Batu River, Untut River, and Kelang River in the Heart of Kuala Lumpur. Figure 3 shows map of Gombak River Catchment area.

Every development program is accompanied by impacts that can be, directly affective from the construction of a specific project, or indirectly so through later utilization. More understanding of the relationship between these programs and the impacts on its surrounding is a matter of greater concern nowadays.

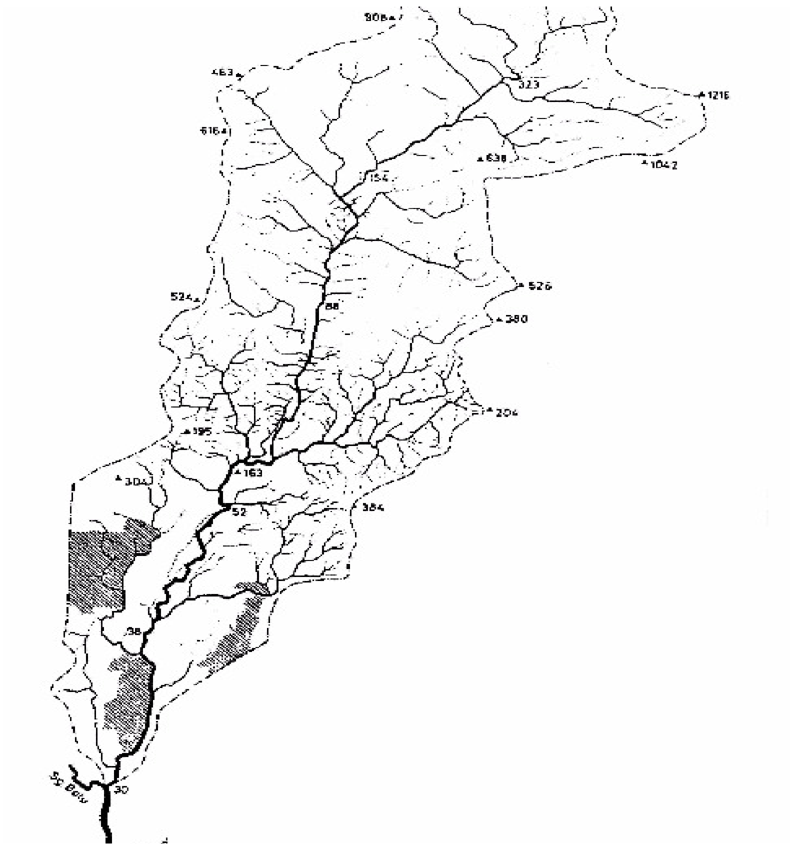

Fig. 3. Map of Gombak River Catchment Area.

2.1. Gombak River and its watershed

Gombak River is situated mainly in the Gombak District in Selangor state and its lower zone is situated in the Malaysian capital Kuala Lumpur. The catchment area within which the river passes through, has grown quite rapidly since early 1970 s and is expected to continue growing in the future. The topography of the watershed area, as it is surrounded by hilly mountains.

Gombak river watershed is in the upper part of Klang river basin. About 60% of the catchment is steep mountains rising to a height of 1220m. The Gombak River drains a narrow elongated watershed that runs slightly west of south from the steep-sloped main range mountains down through more gently sloping foothills to the alluvial plain in the vicinity of North Kuala Lumpur [12]. Sungai Keroh, Sungai Pusu, Sg. Rumput, Sg. Salak, Sg. Semampus and Sg. Blongkong feed Gombak River.

2.1.1. Characteristics of Gombak River watershed

The geologic formation of Gombak River is consisting of diverse lithology as shown on Table 1.

Table 1. Percentages of Catchment Area on Various Lithologies [12].

Types of Geologic Formation	Percentage %
Granite	68.1
Chert facies	0.9
Arenaceous facies	3.1
Lutaceous facies	8.6
Hawthornden schist	3.7
Dinding schist	0.8
Limestone	0.6
Limestone overlaid by quaternary alluvium	13.8
Quartz	0.5

The Gombak River traverses a vast spectrum of land use change within the Gombak river watershed area. The axial length of the drainage basin is 22.2 km, average width 5.5 km, and an area of 123.3 square km. The river confluence with Batu River is at 28.3 m altitude [13]. Gombak River and Kelang River meet at a confluence point in heart of Kuala Lumpur city. The watershed can be divided into three main sections. The upper zone, including the upper tributary sub-zone, takes in the undisturbed forest reserve areas of the watershed and terminates at the point where the river leaves the steep sloped hills and enters the gentler foothill section. The middle and lower zones are with gradients of 4.7% and 2.2% respectively.

Lai [13] showed steady deterioration of water quality with level of urban development in the Klang River. Unsteady deterioration, could only happen for a period of time that is construction period. Klang River around Kuala Lumpur is heavily polluted by industrial and domestic waste according to DOE reports. The activities within the river basin such as forest clearing, intensive and extensive agricultural practices, and urbanization alter the ambient chemistry of river water. All these factors contributed significantly to the increase of concentrations downstream as shown in Tables 2 and 3.

Table 2. WQI of Station 3116626.

Year	WQI			Class (Mean)	Ranking (Mean)
	Mean	Max	Min		
1983	61.7	70.6	40.1	III	Slightly polluted
1984	62.3	62.3	62.3	III	Slightly polluted
1985	66.1	79.7	52.0	III	Slightly polluted
1986	63.1	74.8	46.5	III	Slightly polluted
1994	68.8	74.5	63.4	III	Slightly polluted
1995	69.5	76.8	63.2	III	Slightly polluted
1996	63.1	74.8	46.5	III	Slightly polluted
1997	61.1	75.7	46.4	III	Slightly polluted

Table 3. WQI of Station 3217619.

Year	WQI			Class (Mean)	Ranking (Mean)
	Mean	Max	Min		
1983	89.0	93.0	40.86.3	II	Clean
1984	78.3	95.4	63.1	II	Slightly polluted
1985	81.5	86.6	76.2	II	Clean
1986	78.0	85.8	68.1	II	Slightly polluted
1994	69.2	74.8	59.0	III	Slightly polluted
1995	71.8	81.2	68.3	III	Slightly polluted
1996	66.7	75.9	43.0	III	Slightly polluted
1997	59.5	75.4	43.4	III	polluted

Station 3116626 is located at the lower stream of Gombak river. The overall water quality ranked as slightly polluted for these two periods.

The sources of pollution are expected to increase due to the increase in population and industrialization. Due to vast numbers of point and non-point sources, it is very difficult to carry out monitoring due to the expenses and manpower.

2.1.2. Development across Gombak River

Gombak River is a slow flowing river, which originates from many tributaries in the Gombak district. The river has several confluences with other streams such as Batu River, Untut River, and Kelang River in the Heart of Kuala Lumpur.

Gombak District falls under the jurisdiction of the state of Selangor Darul Ehsan in Malaysia. The utilization length of development plan of the current structure plan is from 1995 to 2020. The GIS map of Gombak district for current and future development as shown on Fig. 4 indicates the utilization of the district in heavy developmental schemes. Gombak District tops other districts in the country in terms of growth percentage.

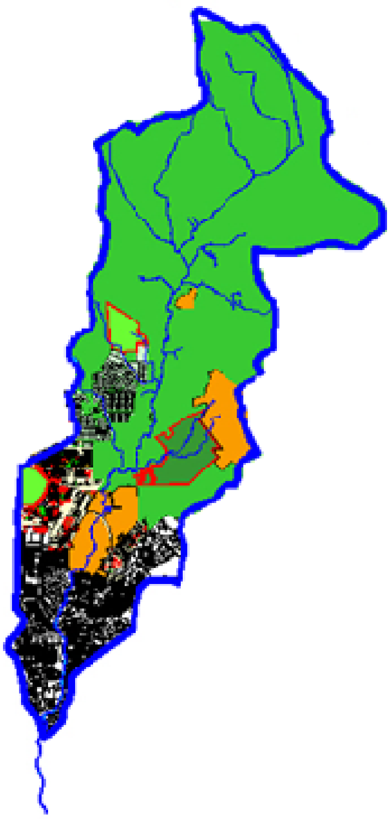

Fig. 4. Land Use Map of Gombak River Catchment Area.

3. Research Approach

In this study, the selection of study area was based on several criterions. The location of the river and its watershed, the variation of water quality along the river stream, land uses around the watershed and wastewater loads discharged to the river were the basis for selection. The secondary sources were obtained from governmental authorities. The study uses existing raw water quality data from official sources and related to eight land uses categories i.e. residential, industrial, commercial, public utilities, recreational, institutional, and forest.

The study uses 8 years of comprehensive and consistent water quality data in three different stations along the Gombak River. These water quality data will undergo an iteration process using a designed neural network in order to predict the water quality index (WQI) based on given data. The model will be able to produce very accurate results using a back propagation network. The results will be then optimized by a sensitivity test carried out with hidden data. These data will be kept away for validation process only and will not be part of the iteration process.

3.1. Structure of the neural network

The networks, as in their standard form shown in Eq. (4) and as a direct generalization of the well known back-propagation network [14], are modeled by

$$\tau \frac{dv_i}{dt} = -v_i + g_i\left(\sum_{j=1}^{N} w_{ij} v_i + \theta_i\right), \quad i = 1, 2, \ldots\ldots, N \tag{3}$$

Equation 4 shows the state of neuron *i* with

$$u_i = \sum_{j=1}^{N} w_{ij} v_i + \theta_i \quad i = 1, 2, \ldots\ldots, N \tag{4}$$

Being its local field, the activation function of neuron *i*; the external input imposed on neuron *i*; the synaptic connectivity value between neuron *i* and neuron *j*; and *N* the number of neurons in the networks. On the other hand, the famous Hopfield networks [15] as shown in Eq. 5 are examples of the second approach and can be described in terms of the local field state $u_i = 1, 2, \ldots, N$ of neurons as

Generic static neural network model

$$\tau \frac{dx}{dt} = -x + G(Wx + q), \quad x = x_o \in \mathbb{R}^N \tag{5}$$

where $x = (x_1, x, \ldots.., x_N)$ is the neuron states, $y = (y_1, y_{12}, \ldots.., y_N)$ is the local fields, $W = (w_{ij})_{N\times N}$ is the synaptic weight matrix and $G : \mathbb{R}^N \to \Omega \subseteq \mathbb{R}^N$ is the nonlinear activation mapping with Ω being a convex subset of $\mathbb{R}^N$.

3.2. Network configuration

After the network has been trained as shown in Table 4 and is able to produce reliable results with a fixed number of cycles, the results will be compared with land use data. Despite the heterogeneity of water quality data due to non-linear changes within the basin, the network was able to identify patterns of changes and minimizing most of associated anomalies. Only 149 record data sets out of 199 were used in the initial iteration Stage for three monitoring stations. The first monitoring station is station 17 which is represented by the water quality data sets from 1 to 71 in the neural network. The second station (station 18) is represented by data sets from 72 to 144, and the last monitoring station (station 24) is represented by 145 to 199 data sets.

Table 4. Neural Network Configuration for the Water Quality of Gombak River.

Property	Value	Property	Value
BestError	0.65	Momentum	0.5
Cycles Completed	23145	Navigate	4
InitMax	1	Process	15
JogMax	25	TableName	All stations combined
LayerJCount	5	TestHigh	149
LearnRate	0.5	TrainHigh	149

4. Results, Validation and Discussion

4.1. Results

Despite the heterogeneity of water quality data due to non-linear changes within the basin, the network was able to identify patterns of changes and minimizing most of associated anomalies. As shown on Table 4, only 149 record data was used in the initial iteration stage for three monitoring stations.

The number of hidden nodes is 5 in this network. The network had to complete 23145 cycles before it reaches the best error of 0.65 as shown on Table 4 and the field configuration in Table 5.

Table 5. Network Design of Gombak River Data Set.

Property	Value
BestError	0.65
CyclesCompleted	23145
InitMax	1
JogMax	25
LayerJCount	5
LearnRate	0.5
Momentum	0.5

Using the mentioned configuration the network was able to attain very precise prediction of water quality index. The first monitoring station is station 17 which is represented by the water quality data sets from 1 to 71 in the neural network. The second station (station 18) is represented by data sets from 72 to 144, and the last monitoring station (station 24). is represented by 145 to 199 data sets. The time series graph in Fig. 5 shows the close precision of water quality prediction as well as the pattern of water quality from station to the other.

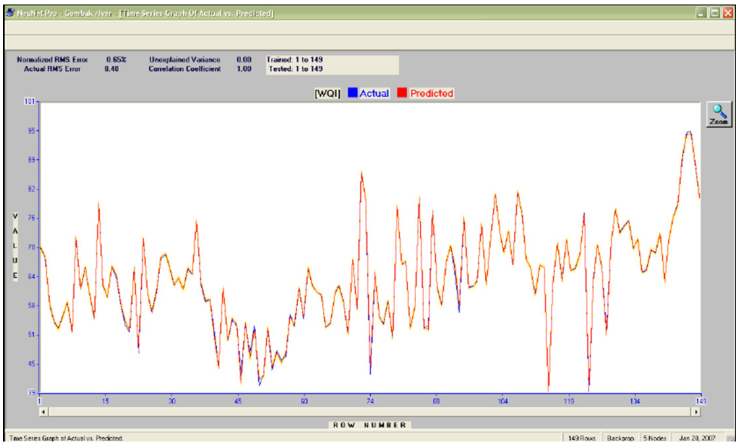

Fig. 5. Time Series Graph of Actual vs. Predicted.

As shown on Fig. 6, the scatter graph shows deterioration of water quality as stream moves from upstream to downstream.

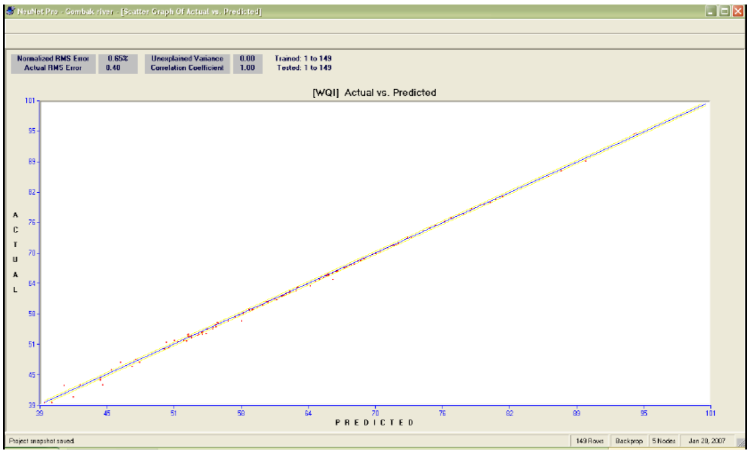

Fig. 6. Scatter Graph of Actual vs. Predicted.

The results from neural network were then compared with the land use of the watershed area shown in Tables 6, 7 and 8.

Table 6. Land Use of Gombak River Watershed Downstream Stations.

Gombak Current LandUse	Size (acre)	No. of Lot	%
Agriculture	1,757.140	552	6.249
Infrastructure and Utility	40.346	91	0.143
Open Sapce and Recreational Area	458.615	132	1.631
Residential	2,538.160	11311	9.026
Industrial	9.955	97	0.035
Commercial	36.148	830	0.129
Transportation	533.059	139	1.896
Institution & Public Facility	938.462	37	3.337
Water Body	1.291	1	0.005
Forest	21,807.720	1	77.550
Total	**28,120.89**		**100.000**

Table 7. Land Use of Gombak River Watershed Downstream Stations.

KL Current LandUse	Size (acre)	No. of Lot	%
Residential	2802.5	15867	38.656
Commercial	620.625	6190	8.560
River reserve and drainage	241.048	222	3.325
Utility	112.505	85	1.552
Education	196.873	113	2.716
Vacant Land	77.842	110	1.074
Cemetery	26.875	7	0.371
Road Reserve	1520.74	1950	20.976
Open Space and Recreational Area	471.386	236	6.502
Electrical Transmission Line reserve	68.915	72	0.951
Industrial	77.552	42	1.070
Institution	875.912	208	12.082
Religion	54.002	69	0.745
Parking Space	44.819	98	0.618
Railway reserve	35.267	13	0.486
Terminal	15.888	12	0.219
Forest	7.161	1	0.099
Total	**7,249.910**		**100.000**

Table 8. Land Use of Gombak River Watershed Downstream Stations.

KL Current LandUse	Size (acre)	No. of Lot	%
Residential	2802.5	15867	38.656
Commercial	620.625	6190	8.560
River reserve and drainage	241.048	222	3.325
Utility	112.505	85	1.552
Education	196.873	113	2.716
Vacant Land	77.842	110	1.074
Cemetery	26.875	7	0.371
Road Reserve	1520.74	1950	20.976
Open Space and Recreational Area	471.386	236	6.502
Electrical Transmission Line reserve	68.915	72	0.951
Industrial	77.552	42	1.070
Institution	875.912	208	12.082
Religion	54.002	69	0.745
Parking Space	44.819	98	0.618
Railway reserve	35.267	13	0.486
Terminal	15.888	12	0.219
Forest	7.161	1	0.099
Total	**7,249.910**		**100.000**

Land use statistics and patterns of changes with association to the neural network can be clearly seen from the time series graph (Fig. 7). From 1 until 72 represent the downstream stations. This part of the network has almost stable discharges as the development activities stabilized where the mid-stream fluctuation is due to the current development schemes that are taking place currently. The last station shows the best quality index due to very low development area. This area has the potential to be contaminated if proper procedures for point and non-point discharges were not carefully followed.

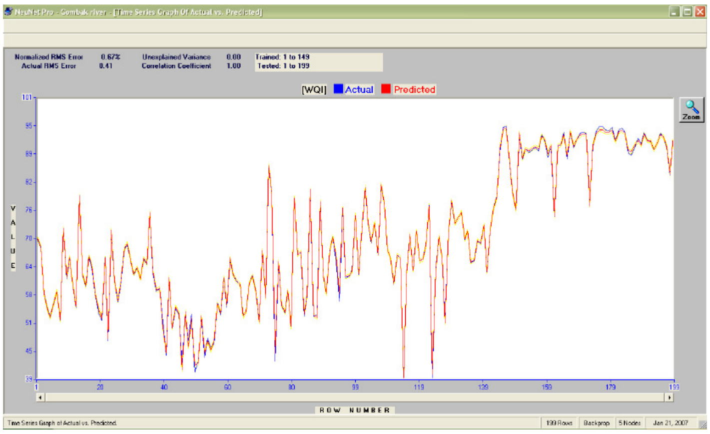

Fig. 7. Time Series Graph of Actual vs. Predicted Data for Downstream Stations.

4.2. Validation and Discussion

After the training phase, Validation process was carried out on an independent data set, which was not been used as part of the training. In this study 50 data sets were kept for validation. The number of completed cycles of iteration has been kept constant with the same field and network configuration as shown on Table 9.

Table 9. ANN Training Configuration of Station 24 in Gombak River.

Name	Max	Min	Field average	Field standard deviation	User Maximum	User Minimum
DO INDEX	100	0	73.3652	25.061	123.487	23.242
BOD INDEX	96.17	10.4134	77.372	15.359	108.09	46.652
COD INDEX	93.78	0.2544	62.863	14.29	91.452	34.274
AN INDEX	100	7.971E-02	44.842	37.760	120.363	30.67
SS INDEX	100	0	66.401	20.774	107.950	24.85
PH INDEX	99.64	85.83	98.0000	1.8642	101.728	94.2715
WQI	94.78	28.10	70.0097	15.4472	100.904	39.115

In this study 50 data sets were kept for validation. The number of completed cycles of iteration has been kept constant with the same field and network configuration. The aim of using a similar network design is to make sure that the prediction of water quality index follows the same pattern circumstances such as land uses and pollutants influences. The results had shown a very close prediction with best error of 0.67 as shown in Fig. 8.

A very good result of 99.39% was achieved during the sensitivity test. Stations 18 had shown the greatest fluctuation among the three. This indicates a rigorous development activity in the region.

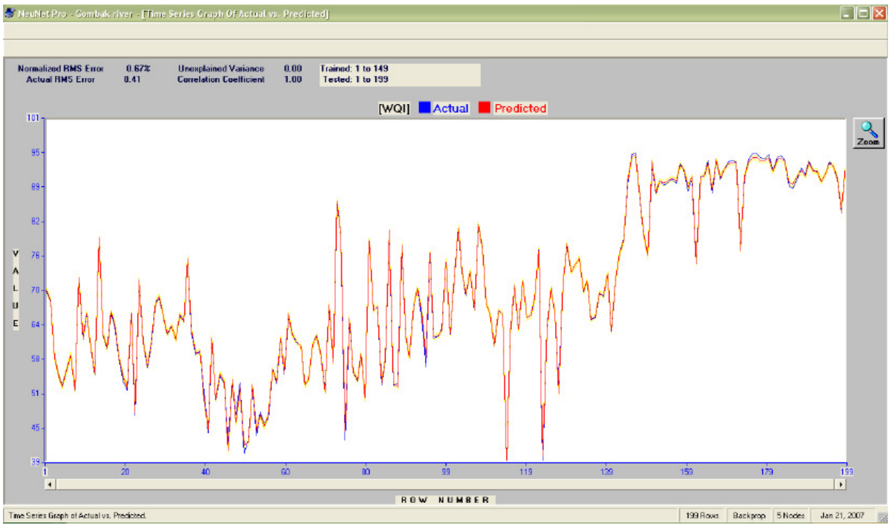

Fig. 8. Time Series Graph of Actual vs. Predicted Data using Hidden Data for Station 18.

5. Conclusions

This study has focused on finding a low-end alternative for water quality monitoring techniques. It uses the ANN approach to provide an effective prediction model that suites environment with high heterogeneity. The proposed approach minimizes most of anomalies associated with prediction methods and provides water quality forecasting and prediction with precision.

References

1. Mas, D.M.; and Ahlfeld, D.P. (2004). Use of artificial neural network models to predict indicator organisms concentrations in an urban watershed. *American Geophysical Union*, *Spring Meeting* 2004, *Abstract* #H53A-06.
2. Arrifin, J.; Ghani, A.A.; Zakaria, N.Z.; and Yahya, A.S. (2004). Sediment prediction using ANN and regression approach. *Proceedings of* 1st *International Conference on Managing Rivers in the 21*st *Century: Issues and Challenges, Malaysia.*, 168-174.
3. Rounds, S.A. (2002). Development of a neural network model for dissolved oxygen in the Tualatin River, Oregon. *Proceedings of the second Federal Interagency Hydrologic Modeling Conference*, Las Vegas, Nevada.
4. Zou, R.; Lung, W.-S.; and Guo, H. (2002). Neural network embedded Monte Carlo approach for water quality modeling under input information uncertainty. *Journal of Computing in Civil Engineering*, *ASCE*, 16(2), 135-142.
5. Ripley, B.D. (1994). Neural networks and related methods for classification. *Journal of Royal Statistical Society*, *Series B*, 56(3), 409-456.
6. Ismail, K.; and Rainis, R. 2004. *Modeling River Water Quality Index Using Artificial Neural Networks and Geographical Information system.* Editors

Alias Abdullah; Norio Okada; and Mohd Kamil Yusoff. Water Environmental Planning, Towards Integrated Planning and Management of Water Resources for Environmental Risks. Published by Bureau of Consultancy & Entrepreneurship, IIUM Malaysia, 99-306.

7. Cormac Technologies Inc. (1999). *Manual, NeuNet Pro*, *Revision 2.2*. Cormac Technologies Inc. http://www.cormactech.com/Neunet.
8. Picton, P.D. (1994). *Neural networks*. Palgrave Macmillan.
9. Brion, G.M.; and Lingireddy, S. (2003). Artificial neural network modeling: A summary of successful applications relative to microbial water quality. *Water Science and Technology*, 47(3), 235-240.
10. Juahir, H.; Zain, S.M.; Ahmad, Z.; Jaafar, N.M. 2004. *An Application of Second Order Neural Network Back Propagation Method in Modeling River Discharge*. Editors: Alias Abdullah; Norio Okada; and Mohd Kamil Yusoff. Water Environmental Planning, Towards Integrated Planning and Management of Water Resources for Environmental Risks. Published by Bureau of Consultancy & Entrepreneurship, IIUM Malaysia, 307-325.
11. Loke, E.; Warnaars E.A.; Jacobsen, P.; Nelen, F.; and Almeida, M.C. (1997). Artificial neural networks as a tool in urban storm drainage. *Journal of Water Science and Technology*, 36(8-9), 101-109.
12. Bishop, J.E. (1973). *Limnology of a small Malayan river Sungai*, *Gombak*. Volume 22. Junk.
13. Lai, F.S. (1983). Biochemical oxygen demand concentration of two river basins of Selangor. *Pertanika*, 6(3), 32-43.
14. Hertz, J.; Krogh, A.; and Palmer, R.G. (1991). *Introduction to the theory of neural computation*. Reading, MA: Addison-Wesley Pub. Co.
15. Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the USA National Academy of Sciences*, 79(8), 2554-2558.