# Assessment on the water quality characteristics of River Ganga at Kolkata Region, India using Water Quality Index and ANN simulation method

Papita Das Saha[1], R. Sengupta[1], Jhuma Saha[2], P.K. Banerjee[2]
[1]*Biotechnology Department, National Institute of Technology- Durgapur 713209, West Bengal, India*
[2]*Chemical Engineering Department, Jadavpur University, 188, Raja S. C. Mullick Road, Kolkata 700032, India*
*To whom correspondences should be addressed*
*E-mail: papitasaha@gmail.com*

*Received October 16, 2011, Revised manuscript received February 01, 2012, Accepted February 06, 2012*

## Abstract

A rapid disappearance of forests, coastal mangrove forests and wetlands is increasingly lacking in natural purification of polluted waters. The present study of Water Quality Index (*WQI*) is based on the analysis of samples of water collected from Ganga River from various locations in and around Kolkata city, India. The water quality parameters are analysed using linear correlation coefficient statistical technique and artificial neural network is used to model Water Quality Index. The results show that Water Quality Index mainly changes depending on the location and seasonal variation and artificial neural network able to simulate the experimental Water Quality Index satisfactorily.

**Keywords**: River Ganga, Water Quality, Contamination, Water Quality Index, Artificial neural network, Correlation Coefficient

## 1. Introduction

The Ganga is a holy river in India. It traverses a distance of 507 km in the state of West Bengal and falls into Bay of Bengal. The River Ganga flows through more than 700 cities and about 120 million liters of waste water added daily by treated and partially treated effluents discharged into River Ganga directly and indirectly from different industries like pesticides, tanneries, paper and pulp mills, petrochemicals and fertilizer complexes, rubber factories and host of others use river. Also a large number of municipal wastes are also being discharged and polluted the water of River Ganga [1-3].

With ever growing population and increasing demand of water for food production, industrial and domestic activities increases every year. No doubt, water is essential for life but 99% of the water in the world has been polluted by man-made resources mainly due to rapid urbanization, industrialization and increasing population [4]. Water resources and the quality of water may be insufficient to meet the long-term requirements of consumption, agriculture and other uses unless their judicious and economic uses are ensured. Making the best use of the available water resources involves not only the prevention and control of its depletion and degradation but also managing it in view of the present and future needs is necessary.

It is revealed that the socio-economic development of Calcutta, the most potential economic zone in India situated on the east bank of Hugli River, a tributary of Ganga, has had a significant impact on the water quality of this major river. The deterioration of water quality is directly related to nonfunctioning and malfunctioning of wastewater treatment plants and lack of environmental planning and coordination [4].

Chemicals in the form of fertilizers, insecticides used in agricultural fields or industries find their destination into the rivers thus posing a threat to the existing flora and fauna native to these aquatic bodies as well as human health. The release of domestic and industrial wastes into the Ganga River makes water unsuitable for consumptive, semi consumptive or non consumptive purposes.

In order to monitor the quality of water throughout the year under different weather condition, six strategic locations were chosen for the sampling of water from Ganga (from downstream of the river, West Bengal, India) to verify the changes of water quality. The locations were Khardah, Agarpara, Dakshineswar, Bagbazar, Babughat and Diamond Harbour. The continuity of Ganga was from North 24-parganas - Kolkata - South 24-Parganas. The DO, pH, COD, BOD, oil and grease, turbidity total suspended solids and total dissolved solids were the parameters based on which the analysis was done. Two schemes were devised to determine the Water Quality Index, the aggregative Water Quality Index and the multiplicative Water Quality Index. The analysis was done for three seasons: pre Monsoon, Monsoon and post Monsoon at high and low tide.

Water Quality Index [5-20] is basically a mathematical means of calculating a single value from multiple test results. It represents the level of water quality in a given water basin, such as a lake, river, or stream important to monitor water quality over a period of time in order to detect changes in the water's ecosystem. It gives an indication of the health of the watershed at various points and can be used to keep track of and analyze changes over time. The *WQI* can be used to monitor water quality changes in a particular water supply over time, or it can be used to compare a water supply's quality with other water supplies in the region or from around the world.

Regression analysis describes the relation between two variable measured simultaneously where one variable is dependent and other is independent variables. Correlation coefficient ($r$) describes the strength of relationship between two inter dependent variables and signifies the ability to prediction of variable from the other. The correlation coefficient index [21-22] can be calculated using the following equation:

$$r=\frac{n.\sum x.y-\sum x.\sum y}{\sqrt{((n\sum x^2-(\sum x)^2)\ (n\sum y^2-(\sum y)^2)}} \quad (1)$$

where $x$ and $y$ are variables and $n$ is the number of variables. $r$ may be positively or negatively correlated with each other. If the value of $r$ is greater than 0.6, it indicate that the variables are dependent with each other.

Artificial neural network (ANN) is the computational methods and it is formulated by the empirical correlation of the black box. This modeling process is information computing techniques and developed by the biological neural processing system [23-24]. The process represents the complex nonlinear relationship without knowing the original procedure and makes the procedure more promising alternatives in many applications. It can be used to develop the complex relation between the input and output process variables.

## 2. Materials and Methods

The parameters like dissolved oxygen, pH, Chemical Oxygen demand, Biochemical Oxygen demand, Oil & Grease, Turbidity, Total suspended solid, total dissolved solid etc were determined using standard procedure [25].

After the experimental determination of the parameters, it was being asked to rank according to their significance as contributor to overall quality by different experts. The rating was done on a scale of 1 (highest) to 5 (lowest), based on the polluting effect of the parameter relative to other parameters. Each of the parameters represents only a part of the overall quality, thus parameters of lesser importance even cannot be discarded, since they were still part of the overall quality [5-20].

Then to convert the rating into weights, a temporary weight of 1.0 was assigned to the parameter which received the highest significance rating. All other temporary weights were obtained by dividing the highest rating by the corresponding individual mean rating of the parameters. Each temporary weight was then divided by the sum of all temporary weights to deduce the final weights, which must sum up to one. A total weight of 1.0 was thus distributed among the parameters to reflect the relative importance of the parameters. The weightage thus assigned to a parameter was an indication of the degree to which water quality may be affected by that particular parameter.

The next step was the transformation of parameters to a common quality scale referred commonly as the quality rating score. The quality rating score was assigned to a particular parameter depending on an individual judgment or a consensus opinion of experts based on the water quality standards. It reflected the magnitude of violation of set of standards. The quality rating was done on a scale of 0 to 100 (i.e. highest to lowest polluting).

Finally, an overall quality rating was derived by multiplying the final weights ($w_1$) of each individual parameters with the corresponding quality rating ($q_1$), the sum of which gave the required single number *WQI*.

The systematic opinion research technique, as attempted by Robert M. Brown [5] incorporated the judgment of a large and diversed panel of experts in water quality management.

The Procedure used in formulating the Water Quality Index in this study was called as DELPHI process.

Aggregative and Multiplicative methods had been followed to determine *WQI* value of water bodies.

2.1. Aggregative Method

According to the method Water Quality Index was calculated as follows [19-20]:

$$WQI_a = \sum_{i=1}^{n} q_i w_i \qquad (2)$$

Where, $WQI_a$ = aggregative Water Quality Index, a number between 0 and 100.
$q_i$ = quality of $i^{th}$ parameter, a number between 0 and 100.
$w_i$ = weight of $i^{th}$ parameter a number between 0 and 1.
$n$ = total number of parameter.

As per the method water resource classification is done as below:

2.2. Multiplicative Method

According to the method Water Quality Index was calculated as follows [19-20]:

$$n\, WQI_m = \prod_{i=1}^{n} (q_i)^{w_i} \qquad (3)$$

Where, $WQI_m$ = multiplicative Water Quality Index, a number between 0 and 100.
$q_i$ = quality of $i^{th}$ parameter, a number between 0 and 100.
$w_i$ = weight of $i^{th}$ parameter a number between 0 and 1.
$n$ = total number of parameter.

As per the method water resource classification was done (Table 1).

**Table 1**. Water Quality Index for Aggregative and Multiplicative methods

| Class | Aggregative methods | | Multiplicative methods | |
|---|---|---|---|---|
| | $WQI_a$ value | Description | $WQI_m$ value | Description |
| A | 63-100 | good to excellent | 81-100 | very good |
| B | 50-63 | good to moderate | 51-80 | good |
| C | 38-50 | Bad | 21-50 | medium |
| D | below 38 | bad to very bad | 0-20 | bad |

A computer programming using MATLAB version R2009b had been written for training and the testing of the ANN. After a number of training trials, the best neural network model was generated. The maximum number of epochs (training cycles) was chosen by a trial and error approach. Performance goal and minimum performance gradient were set so as to ensure a model with good performance. Trial and error method was used to find the most suitable network model for the *WQI* analysis using the different eight parameters [23-24] and the results of *WQI* was validated in respect of experimental and ANN model deviation.

## 3. Results and Discussions

The quality rating curves using published water quality standards and guidelines related to specific water indicated the rating values on a scale of zero to one hundred for any individual parameter (Tables 1 and 2). The rating 100 signified the best water quality condition and 0 showed the worst water condition. The water samples were collected at different low tide and high tide period to understand the water quality parameters of the River Ganga at different tiding period also with seasonal variation (Fig. 1).

**Table 2**. Values of weightage (WI)

| Parameters | Weightage |
|---|---|
| pH | 0. 10 |
| DO | 0.15 |
| Turbidity | 0.05 |
| BOD | 0.25 |
| COD | 0.25 |
| Oil & Grease | 0.1 |
| TSS | 0.05 |
| TDS | 0.05 |

3.1. Water Quality Index for Pre-Monsoon (High Tide)

Based on the value of the aggregative Water Quality Index all the six stations had good to excellent water quality. Based on the multiplicative Water Quality Index all the six stations had also good water quality (Fig. 2).

**Figure 1**. Locations of sample collected from River Ganga at Kolkata

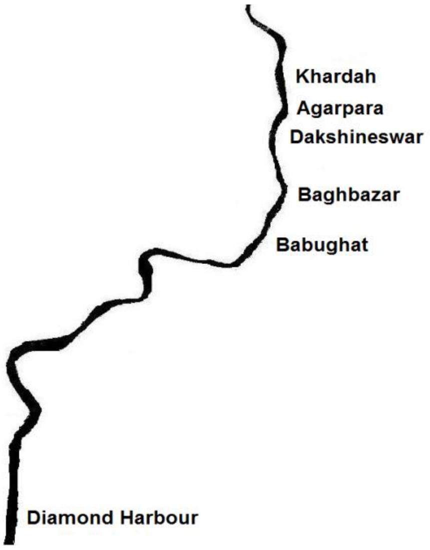
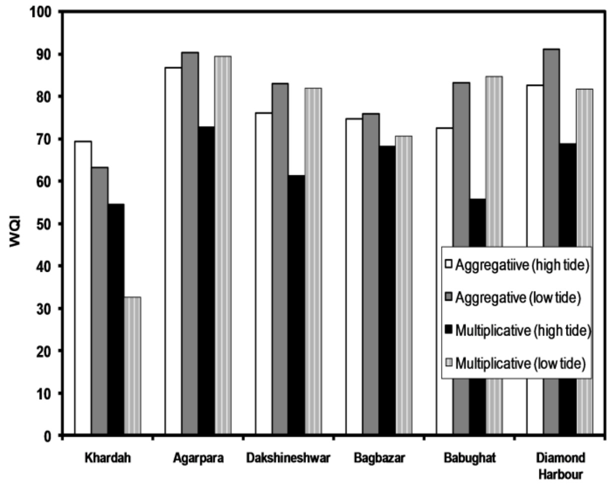



**Figure 2**. Water Quality Index for pre-Monsoon time using *WQI* method



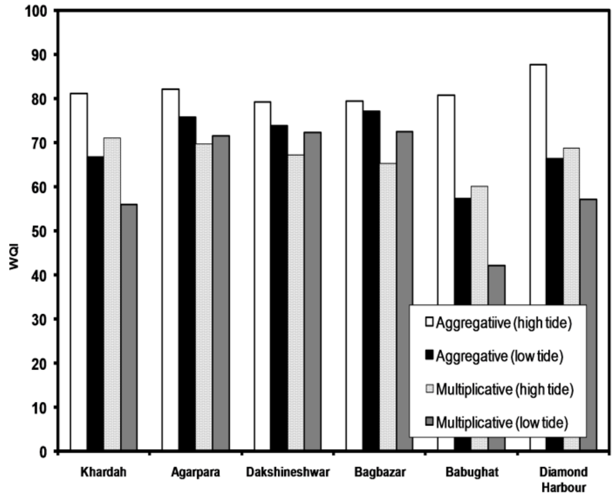

3.2. Water Quality Index for Pre-Monsoon (Low Tide)

For Khardah and Bagbazar, the water quality was good on the basis of aggregative Water Quality Index, while it was very good for Dakshineshwar and Babughat and excellent for Diamond Harbour and Agarpara (Fig. 2). However, on the basis of the multiplicative index the water quality of Khardah was medium, Bagbazar was good and all the other four stations had very good water quality indices.

3.3. Water Quality Index for Monsoon (High Tide)

Based on the value of the aggregative and multiplicative Water Quality Index it was observed that all the six stations had a very good water quality (Fig. 3).

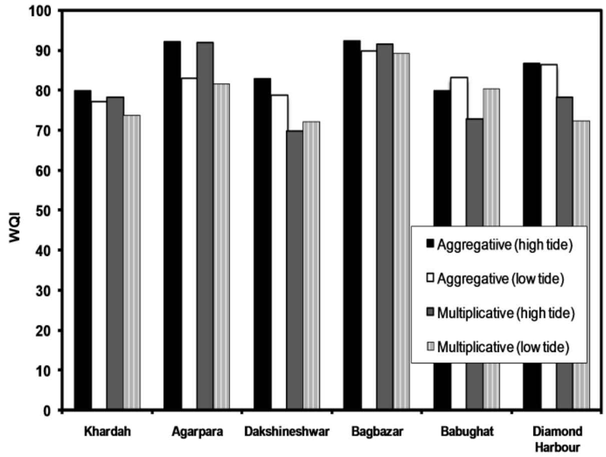
**Figure 3**. Water Quality Index of Monsoon sample of Ganga River using *WQI* Method




3.4. Water Quality Index for Monsoon (Low Tide)

On the basis of aggregative Water Quality Index Babughat had a moderate Water Quality Index, Khardah and Diamond Harbour had a good water quality and Agarpara, Dakshineshwar and Bagbazar had a very good water quality (Fig. 3).

**Figure 4**. Water Quality Index of Post - Monsoon sample of Ganga River using *WQI* Method




3.5. Water Quality Index for Post Monsoon (High Tide)

Khardah, Dakshineshwar and Babughat had a very good water quality while Agarpara, Bagbazar and Diamond Harbour had excellent water quality given by aggregative Water Quality Index (Fig. 4). As per the multiplicative Water Quality Index, the stations

Dakshineshwar, Khardah, Diamond Harbour and Babughat had a good Water Quality Index; and Agarpara and Bagbazar had very good Water Quality Index.

### 3.6. Water Quality Index for Post Monsoon (Low Tide)

Khardah, Agarpara, Dakshineshwar and Babughat had very good water quality and Diamond Harbour and Bagbazar had excellent water quality from the values of the aggregative water quality indices. Khardah, Dakshineshwar, Diamond Harbour and Babughat had good water quality; and Agarpara and Bagbazar had a very good water quality apparent from the multiplicative water quality indices (Fig. 4).

The reasons of the different Water Quality Index were due to the location of the sample collection and due to the weather at different time. The effects of location on Water Quality Index were:

At Khardah, there were mainly jute mills and chemical industries producing Superphosphate and Sulphuric Acid. The jute mills were responsible for the high BOD and COD values and were also accountable for the elevation in the TSS and TDS values. Futher, the jute mills released effluents which were alkaline in nature. These effluents resulted from the cooking and desizing of the jute fibres.

Agarpara had jute mills, metal processing industries pharmaceutical industries and cement manufacturing industries. The jute mills resulted in increase levels of BOD, COD, TDS and TSS. The metal processing industries resulted in the lowering of the pH of Ganga water in this area. They also produced scrap and slag materials, polluting the water. The effluents released by the pharmaceutical industries also affected the water quality. The chief raw materials used for cement production were limestone, dolomite, china-clay, gypsum and coal. These substances, when discharged into the water, contributed to the increase in TSS and TDS and turbidity.

Near Dakshineshwar there were several heavy chemical industries producing mainly alkalis like caustic soda and soda ash, which when allowed traversing into the river without being neutralized, may lower the pH of the water. Moreover, huge amounts of organic wastes find their way into the river due to the discharge of dead bodies, excreta and refused. This affected the water quality by altering mainly Dissolved Oxygen, Biochemical Oxygen Demand and Chemical Oxygen Demand of the river water.

In Bagbazar, there were many glass factories. The raw materials for manufacture of glass were silica sand, calcium, oxide, soda and magnesium. These materials usually tended to decrease the pH of water and produced several solid wastes, much of which finded its way into the meager flow of the Ganga.

Babughat located in the eastern part of the metropolitan megacity Calcutta (140 km upstream from seaface), Diamond Harbor (70 km upstream from sea face). Physicochemical characteristics of this partially mixed estuary were largely influenced by the interaction of seawater and discharge of riverine freshwater, annual precipitation and surface runoff. The levels of total dissolved solids and total suspended solids showed a decrease in the Monsoons leading to a higher q value. There was a high biochemical oxygen demand (BOD) values in Babughat, attributed mainly due to huge organic load of untreated sewage from the twin city Howrah and Calcutta situated in the east and west of the river. Elevated levels of dissolved Hg and Pb were previously also recorded in Babughat, with values ranging from 0.16 to 0.95 μg /mL and 0.017 to 0.076 μg/ mL, respectively. These high values for Hg can be attributed to the discharge from pulp and paper manufacturing units and to atmospheric input and runoff of automobile emission for Pb. Spatiotemporal distribution of heavy metals revealed a wide range of variations reflecting input of huge anthropogenic inputs associated with a number of physical and chemical processes. Levels of metals registered a seasonal pattern, with an increase during late Monsoon months (September–October), resulting in relatively low pH of the water.

At Diamond Harbour, there were mainly agro and food processing industries that affect the water quality. They usually increased the organic load of the water and also alter the pH.

From the significant coefficient correlation regression analysis (Table 3), it was observed that only two parameters are positively dependent on each other: BOD and COD ($r$ > 0.6), the linear relationship between this two parameters were: BOD = 0.2951 COD + 2.89; and $r^2$ = 0.4364.

**Table 3**. Correlation coefficient (*r*) values among the various water quality parameters of Ganga River (Pre – Monsoon)

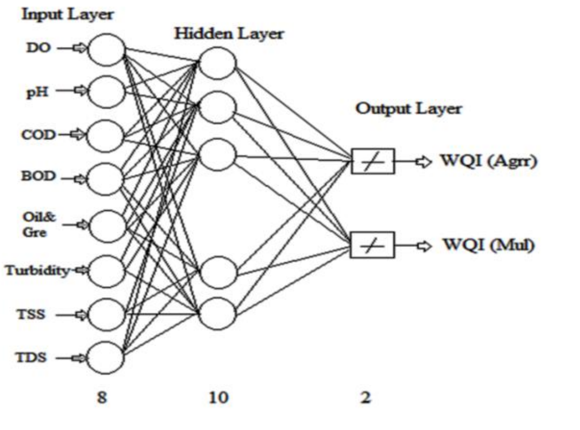
| | DO | pH | COD | BOD | Oil & Grease | Turbidity | TSS | TDS |
|---|---|---|---|---|---|---|---|---|
| DO | 1 | | | | | | | |
| pH | 0.105284 | 1 | | | | | | |
| COD | -0.35154 | 0.351519 | 1 | | | | | |
| BOD | -0.26405 | 0.358535 | 0.640855 | 1 | | | | |
| Oil & Grease | -0.02486 | -0.04932 | -0.15311 | 0.132528 | 1 | | | |
| Turbidity | 0.300052 | 0.198395 | -0.35681 | -0.00628 | 0.163251 | 1 | | |
| TSS | 0.042265 | -0.37323 | -0.30743 | -0.35239 | 0.405437 | -0.04869 | 1 | |
| TDS | 0.097014 | -0.09319 | -0.35624 | -0.26727 | 0.278046 | 0.096389 | 0.173791 | 1 |

All the models were tested with 10 neurons in artificial neural network method (Fig. 5). In case of *WQI* model development, "poslin" transfer function with resilient backpropagation algorithm gave the satisfactory results (Figs. 6a and 6b). It was found from artificial neural network analysis that 10-12 neurons produce the maximum *r* value (0.985 for multiplicative method and 0.968 for aggregative method). All the models were tested with 10 neurons. The % deviation (for validation of ANN) between the experimental *WQI* and the simulated ANN *WQI* results (Table 4) were very low (less that 1.3%). It signified that ANN results can also be used to predict the *WQI* at various season.

**Table 4**. % Deviation of experimental results with the ANN simulated results

| Sl No | *WQI* (Aggregative) | | | *WQI* (Multiplicative) | | |
|---|---|---|---|---|---|---|
| | Experiment | ANN Simulation | % Difference | Experiment | ANN Simulation | % Difference |
| 1 | 69.31 | 69.80 | -0.712 | 54.53 | 54.82 | -0.544 |
| 2 | 86.83 | 86.202 | 0.722 | 72.61 | 72.454 | 0.213 |
| 3 | 76.13 | 76.187 | -0.075 | 61.16 | 61.29 | -0.214 |
| 4 | 74.68 | 74.830 | -0.20 | 68.06 | 68.018 | 0.060 |
| 5 | 72.51 | 72.799 | -0.399 | 55.63 | 55.899 | -0.48 |
| 6 | 82.65 | 82.29 | 0.4350 | 68.81 | 68.749 | 0.087 |
| 7 | 63.15 | 64.038 | -1.40 | 32.52 | 33.367 | -2.60 |
| 8 | 88.25 | 89.40 | -1.307 | 89.4 | 88.825 | 0.643 |
| 9 | 82.98 | 82.599 | 0.458 | 81.81 | 81.424 | 0.470 |
| 10 | 75.94 | 76.009 | -0.091 | 70.6 | 70.49 | 0.148 |
| 11 | 83.175 | 82.78 | 0.472 | 84.76 | 84.301 | 0.541 |
| 12 | 91.18 | 90.27 | 0.993 | 81.63 | 81.249 | 0.466 |

**Figure 5.** Artificial Neural network structure for calculating *WQI*



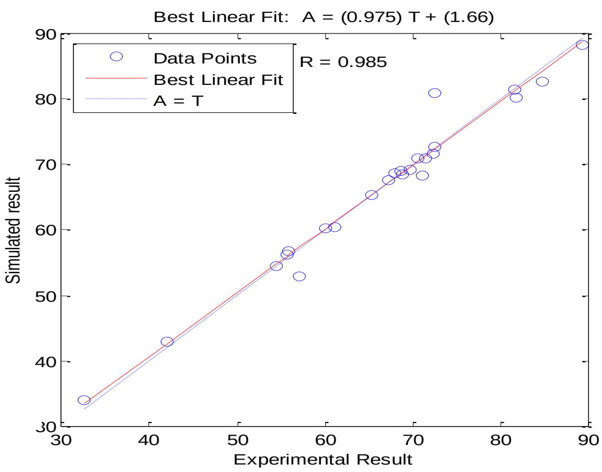

## 4. Conclusion

From the above analysis, it can be concluded that the water quality revealed an enhancement during the post Monsoon or winter period. The water of the river at these stations was the worst during the preMonsoon period with a high tide. During the preMonsoon period and high tide condition Agarpara and Diamond Harbour had the best water quality while Khardah and Babughat had the worst water quality. During preMonsoon and low tide conditions both the water quality indices were higher, indicating betterment in water quality. Agarpara and

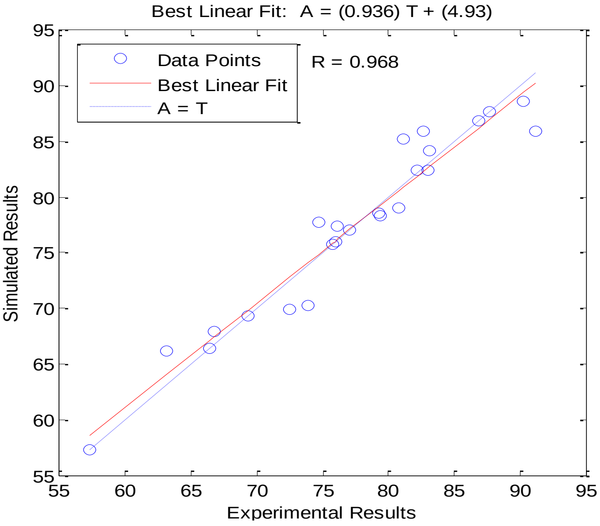
Diamond Harbour had the best water quality while Khardah and Babughat had the worst water quality. In the Monsoon period Agarpara had better Water Quality Index than the other stations whereas Khardah and Babughat had the lowest water quality indices. In the post Monsoon period, the high tide conditions exhibits better water quality than the low tide conditions. In the high tide status Agarpara had the best and Babughat has the worst water quality. In the low tide situation Diamond Harbour had the best water quality while Khardah has the most dire water quality. Thus, it was apparent that among the six sites Agarpara had the best water quality followed by Diamond Harbour While Babughat and Khardah had the worst water quality. Bagbazar and Dakshineshwar had a moderate water quality. The program created during the work can be used to determine the quality of the water in different water bodies and can be used to provide the values of Water Quality Index online through the internet. Once the water quality indices and the quality of the water were determined, the various reasons for the deterioration of the water quality had been studied and can be further analyzed in the future. Further, measures can be taken to reduce the water pollution where the Water Quality Index decreased at an alarming rate. From correlation coefficient regression analysis, it was observed that BOD and COD were dependent with each other and simulation of the parameters showed that the ANN results can describe the *WQI* of the water bodies of Ganga at any season.

**Figure 6**. Artificial Neural Network analysis of *WQI* using Multiplicative & Aggregative Method

(a)




(b)




## References

[1] Singh N. (2010) Physicochemical properties of polluted water of River Ganga at Varanasi, I. J. Energy Env.;1(5):823-832.

[2] Tiwary RK, Rajak, GP, Mondal MR. (2005) Water quality assessment of Ganga River in Bihar Region, India, J. Environ. Sci. Engg.;47(4):326-335.

[3] Trivedia RC. (2010) Water quality of the Ganga River - An overview, Aquat. Ecosys. Health Manage.;13(4):347-351.

[4] Singh L, Choudhary SK, Singh PK. (2011) Organochlorine and Organophosphorous pesticides residues in water of River Ganga at Bhagalpur, Bihar, India, Int. J. Environ. Pollut.;1(1):77-84.

[5] Brown RM, Mcclelland NI, Deinnings RA, Tagore RG. (1970) A water quality index, do we dare? Water Sewage Works;11:339-343.

[6] Prati L, Pavanello R, Pesarin F. (1971) Assessment of surface water quality by a single index of pollution, Water Res.;5(9):741-751.

[7] Singh RK, Anandh H. (1996) Water quality index of some Indian Rivers, Ind J. Environ.;38(1):21-34.

[8] Alexander SA, Jeyakar DA, Chellaraj J, Merlin P, Rajendran A. (1999) Formulation of a new water quality index – $HWQI_2$. Indian J. Environ. Prot.;19(11):842-845.

[9] Swamee PK, Tyagi A. (2000) Describing water quality with Aggregate Index. J. Environ. Eng.-ASCE;126(5):451-455.

[10] Pesce SF, Wunderlin D.A. (2000) Use of water quality indices to verify the impact of

Cordoba City (Argentina) on Suquia river, Water Res.;34(11):2915-2926

[11] http://www.ncsec.org/team26/wqi.htm last modified on July 7, 2001

[12] http://bcn.boulder.co.us/basin/watershed/wqi_nsf.html last modified on Jun 15, 2002

[13] http://www.cce.iastate.edu/research/lutz/dmrwqn/water_quality_index_calc.htm last modified on December 16, 2003.

[14] http://www.nsfconsumer.org/environment/wqi.asp last modified on December 2003

[15] http://www.indiana.edu/bradwood/eagles/wqi.htm last modified on December 2003

[16] http://bcn.boulder.co.us/basin/watershed/wqi_info.html last modified June 14, 2003

[17] Bordalo AA, Teixeira R, Wiebe WJ. (2006) A water quality index applied to an international shared river basin: the case of the Douro River, Environ. Manage.;38(6):910-920.

[18] Sanchez E, Colmenarejo MF, Vicente J, Rubio A, Garcia MG, Travieso L, Borja R. (2007) Use of the water quality index and dissolved oxygen deficit as simple indicators of watersheds pollution, Ecol. Indic.;7(2):315-328.

[19] Saha P, Banerjee PK, Datta S. (2007) Assessment of water quality of Tolly's Nullah by water quality index (WQI), J. Inst. Engg.;88:3-9.

[20] Saha P. (2010) Assessment of water quality of River Damodar by water quality index method, Indian Inst. Chem. Engg. J.;52:145-154.

[21] Rao KS, Prasad NVVS, Rambabu C, Kishore M, Ravi M, Vani KNK, (2004) Water management study of A. Konduru and Reddigudem Mandala of Krishna District, Indian J. Environ. Prot. 24(12):908-916.

[22] Saha P, Datta S, Banerjee PK. (2010) Study on the water quality properties of water bodies-Tolly's Nullah", ISBN: 978-3-8383-7459-8, LAP LAMBERT Academic Publishing.

[23] Khataee AR, Zarei M, Dehghan G, Ebadi E, Pourhassan M. (2011) Biotreatment of a triphenylmethane dye solution using a Xanthophyta alga: Modeling of key factors by neural network, J. Taiwan Inst. Chem. Eng.;42:380-386.

[24] Shetty KV, Nandnnavar S, Srinikethan G. (2008) Artificial neural networks model for the prediction of steady state phenol biodegradation in a pulsed plate bioreactor, J. Chem. Technol. Biot.;83:1181-1189.

[25] APHA – AWWA – WPCF. (1976) Standard Methods for the Examination of Water & Waste Water.' 14th Edition, pp 513-554.