International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072

PREDICTION OF WATER QUALITY PARAMETERS OF KALLAI RIVER USING ARTIFICIAL NEURAL NETWORK

Adharsh K C[1], Muhammad Fayis [2]

[1]M.Tech student, Environmental Engineering, Dept. of Civil Engineering, M-Dasan Institute of Technology, Kerala, India
[2]Assistant Professor, Dept. of Civil Engineering, M-Dasan Institute of Technology, Kerala, India

---***---

Abstract - *Human activities have adverse impacts on river water quality, whether by direct disposal or indirectly. The action of direct disposal polluted the river by illegal garbage dumping or chemical disposal from a nearby factory. In the present scenario Modelling of water quality parameters is one of the most prominent areas of study and various research papers are available in this area. This study focuses on the possibility of forecasting river water quality parameters using Artificial Neural Network. Predicted future river water quality parameters can be used as a baseline for identifying unprecedented changes in the river water. For this, two different stations of Kallai River in Kozhikode, Kerala have been taken as the study area. The two sampling points are Mampuzha Bridge and Kallai Bridge. River water quality parameter data's where collected from various authorities and found monthly data to be the preferred data for my work. Data were preprocessed and wavelet decomposition was used to bring out the time series patterns within data by decomposing into sub-bands and artificial neural network models were trained for prediction of future sub-bands. Sub bands are combined back into a single band, which is the required prediction data. The results were improved by using various wavelets and by fine-tuning neural network models.*

***Key Words*: Water quality parameters, Machine Learning, Kallai River, ANN model, Forecasting**

1. INTRODUCTION

Water is the most important in shaping the land and regulating the climate. It is one of the most important compounds that profoundly influence life. Without water life on Earth will be discriminated. The world is facing a dilemma of natural resource scarcity, especially that of water in view of population growth and economic development. Most of the freshwater bodies all over the world are getting polluted, thus decreasing the potability of water. Depending on water life exists, water exits in nature in many forms like the ocean, river, lake, clouds, rain, snow and fog etc. Around 1.4 billion km volume on earth is covered by water which is about 70% of the total area of the earth. 2.5% is the amount of fresh water which is going to be used by all living species. The rivers, lakes and atmosphere underground hold 1% or less of all the pure water.

Water is used for drinking purpose as it flows from river to lake and then canals and ponds. Water is available in river without any outer protection therefore it is exposed to various forms of polluting agents and in that way or the other it gets polluted.

The problem of prediction of water quality parameters plays a significant role in river basin planning and water pollution control. In last decades, there has been an altering necessity for water quality monitoring of many rivers by uniform measurements of different water quality variables. The probability of a pollutant being discharged in to rivers by industrial and municipal waste disposal brings constant attention to those amusing and using water from rivers. For the prediction of water quality parameters in a river under assumptions of interest, different deterministic models have been tried in the past. A computational method animated by the studies of the brain and nervous system of biological organisms is called an Artificial Neural Network (ANN). It performs highly idealized mathematical models for our present understanding of such complex systems. One of the best characteristics of the neural networks is their ability to learn. The process of learning for ANN is called training the neural network. The training of ANN regulates itself to develop an internal set of features that utilizes to classify information or data. [2] In contrast with the ordinary methods, ANNs afford incomplete data, sacrificial results, and have less vulnerability to outliers.

The artificial neural networks are developed by using the neural network tool box in MATLAB. ANN has sustained a volatile growth in the application in almost all areas of research. The ANN method has various advantages over semi-empirical or traditional phenomenological models because they involve known set of input data without any assumptions. The ANN acquires a mapping for the input and output variables, which can afterwards be used to predict the required output as a function of desirable inputs. Any smooth, measurable function between input and output vectors can be approximate with a multi-layer neural network by choosing a suitable set of connecting weights and transfer functions. ANN models have been exceedingly employed for water quality problems. The major aim of the present work is to build an artificial neural network (ANN) model in MATLAB for Kallai river water quality parameters and explain its application in the complex data of water quality as how the model has the ability of the improvement for the interpretation of the results [3].

1.1 Objectives

The aim of the study is to investigate the possibility of predicting future river water quality parameters from the previously collected data, with the help of an artificial neural network. Unlike, modelling the water quality parameters, which depends on water quality parameters to predict another set of water quality parameters at the same time frame. This study focuses on predicting the water quality parameters at a future time frame. The objectives of the study are as follows:

- To collect water quality parameter data's at the Kallai Bridge and Mampuzha Bridge from various institutional authorities, who collects and manages these pieces of information.
- Understanding various water quality parameters and choosing the most suitable data for this purpose.
- Understanding various software tools for detailed analysis.
- Preprocessing the stored data to make it suitable for processing.
- Develop Artificial Neural Network models for prediction.

Finding suitable techniques to enhance the prediction results.

2. METHODOLOGY

2.1 Study Area

The study area selected here for this project was the Kallai river, shown in Figure 1. Kallai River which has a length of 40 km is located in Kozhikode district of Kerala state. It is one of the smallest rivers of Kerala. This river originates from Cherikkulathur village, Kozhikode District, Kerala of Western Ghats. This river joins with the Chaliyar River and finally enters Arabian Sea at Kozhikode.

Long ago, the logs and timber from the Nilambur forest are tied together and rolled into Kallai. This method of transporting logs was easier. But now even the movement of small boats through this river is difficult due to the presence of silt. Earlier Kallai was considered as the Lifeline of Kozhikode but now it is dying due to pollution.

The water quality parameters of the Kallai river are monitored by the Kerala State Pollution Control Board and the Kerala Water Authority. The monthly water quality data of the Kallai river are collected from the Kerala state pollution control board and the Kerala Water Authority. The water quality data of the two stations Mampuzha Bridge and Kallai Bridge of the year 2011 to 2019 are collected.

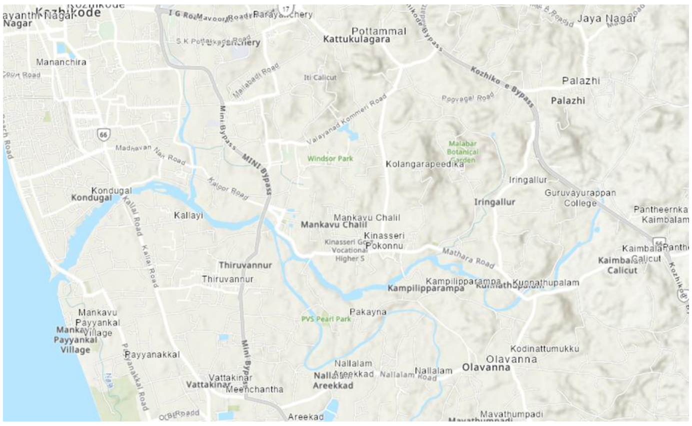

Fig -1: Map showing the course of Kallai river

2.2 Sampling Points

Two sampling points were selected along the course of the river. The sampling points selected are at Kallai Bridge (fig1) and Mampuzha bridge (fig 2). This sampling points almost 10km distance apart from each.

Sampling point 1

- Kallai bridge
- Geographic Coordinates- 11°14'15.6"N 75°47'14.7"E

Fig -2: map showing sampling point 1

Sampling point 2

- Mampuzha bridge
- Geographic Coordinates - 11°14'11.6"N 75°50'38.7"E

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072

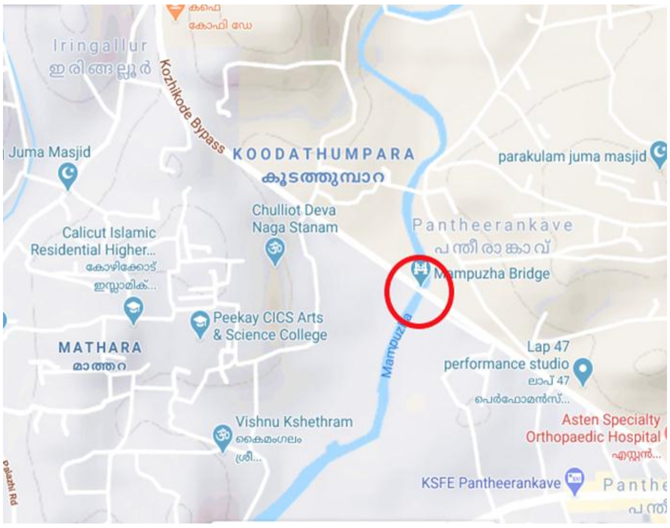

Fig -3: map showing sampling point 2

2.3 Data Collection

Data was collected from three sources. Most of the data where yearly or partial. Only Kerala Water Authority was able to provide the data suitable for the purpose.

Data was collected from ENVIS Kerala, Kerala pollution control board and kerala water authority. ENVIS has water quality data from the year 2002 to 2016. Out of which it has only data of Kallai River at Kallai Bridge and the data is available from 2008 to 2016. Only yearly data available, with the minimum, maximum and mean values.

The Kerala State Pollution Control Board collects river data as part of their studies. They have published the yearly and seasonal data of Kallai River at Kallai Bridge. Seasonal data contains data from the year 2015 to 2018 for Pre-Monsoon (Jan-May), Monsoon (June-Sep) and Post Monsoon (Oct-Dec). Data contains 5 water quality parameters such as pH, Conductivity, D.O., B.O.D., T.C. Yearly data contains data from the year 2009 to 2018. Data contains 8 water quality parameters such as pH, Conductivity, D.O., B.O.D., Nitrate, Ammonia, T.C, F.C. The Data from Kerala Pollution Control Board Office has monthly data for 36 months from the year 2010 to 2019. Data had 26 water quality parameters. Lot of in between months are missing in the obtained data. For water quality parameter modelling this would have been enough. But, data forecasting requires data at regular intervals.

The data collected from Kerala Water Authority is from the year of 2011, 2012 and 2013. Data had water quality parameters at Kallai Bridge and Mampuzha Bridge. Data had 8 water quality parameters such as pH, Turbidity, Chloride, Iron, Nitrate, C.O.D., D.O., and Total Coliform. Data was monthly and continuous without any missing months (36 entries or each location). Which was most suitable for the purpose.

2.4 Data Preprocessing

Prediction of the future data based on past and present data can be most commonly done by the analysis of trends. To achieve the objective of the study, the following integrated approach has been adopted in this work. Block diagram of the approach is shown in Figure 4. Raw Data is the chosen dataset of water quality parameters collected for a single sampling point. Raw data must be a time series of data. A time series is a series of data points indexed listed in time order. What needed here is a time series sequence taken at successive equally spaced points in time.

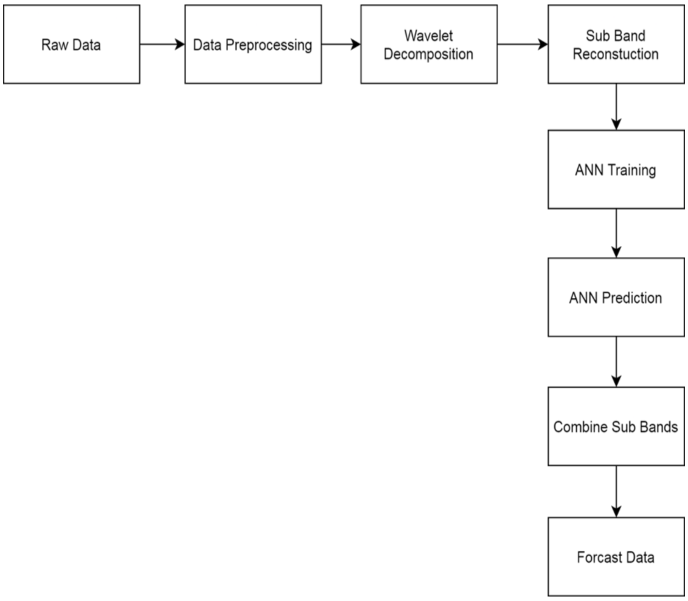

Fig -4: Block Diagram

In this case, a monthly sequence. Such a dataset has been collected from the Kerala Water Authority for two sampling points in Kallai River. One sampling point is the Kallai Bridge and the other is the Mampuzha Bridge. Collected data may not be suitable for any processing as it is. The collected data may contain non-numeric data which is not suitable for processing like "Nil", "BDL", "trace" etc. Which has to replace with corresponding numeric equivalents. Sometimes some of the values may be missing also. These can be filled by linear interpolation of neighbouring non-missing values. Then unnecessary data needs to be removed if there is any. These processes that make the raw data suitable for computation is called Data Preprocessing. In any Machine Learning process, Data Preprocessing is that step in which the data gets transformed, or Encoded, to bring it to such a state that now the machine can easily parse it. In other words, the features of the data can now be easily interpreted by the algorithm.

2.5 Wavelet Decomposition

Wavelet transforms are a mathematical means for performing signal analysis when signal frequency varies over time. For certain classes of signals and images, wavelet analysis provides more precise information about signal data than other signal analysis techniques. After data preprocessing, each water quality parameter in data is decomposed using a Multilevel Wavelet Transform. In this case, a 3 level Daubechies Wavelet Transform is suggested. Figure 5 shows 3-level Wavelet Decomposition.

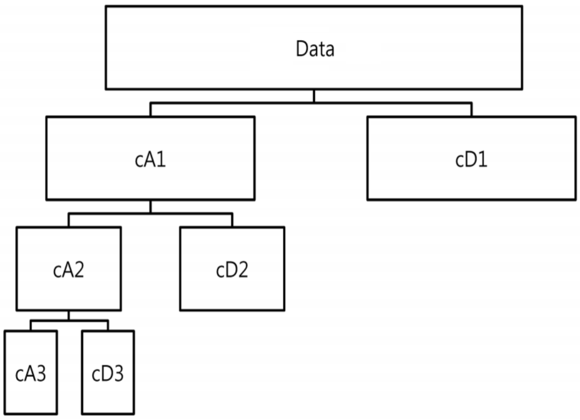

Fig -5: 3-level Wavelet Decomposition

Depending on the data, a number of levels and used wavelet transform may vary. Figure 5 shows the decomposition of a 3 level Wavelet Transform. Data mentioned in the figure represents the data of a single water quality parameter for all time. At the first level of decomposition, data is decomposed into approximation coefficients(*cA1*) and detail coefficients(*cD1*). Approximation coefficients are generated by passing the data through a lowpass filter and down sampling. Whereas detail coefficients are generated by passing the data through a high pass filter and downsampling. At the second level, previously generated approximation coefficients(*cA1*) is further decomposed into its approximation coefficients(*cA2*) and detail coefficients(*cD2*). At each level further in, approximation coefficients are again and again divided into its approximation coefficients and detail coefficients. The final set of coefficients are all the detail coefficients and the approximation coefficients from the final decomposition. That is, for a 3 level decomposition, there will be 4 sets of coefficients. These coefficients are individually reconstructed, generating sub-bands of the original data. This process is repeated for all the water quality parameters.

2.5 Wavelet Reconstruction

The discrete wavelet transforms provide a perfect reconstruction of the signal upon inversion. This means that you can take the discrete wavelet transform of a signal and then use the coefficients to synthesize an exact reproduction of the signal to within numerical precision. Decomposed components can be assembled back into the original signal without loss of information. This process is called reconstruction, or synthesis. The mathematical manipulation that effects synthesis is called the *inverse discrete wavelet transform* (IDWT). Where wavelet analysis involves filtering and downsampling, the wavelet reconstruction process consists of upsampling and filtering. Upsampling is the process of lengthening a signal component by inserting zeros between samples. Fig 6 shows an example of one of the reconstructed sub bands

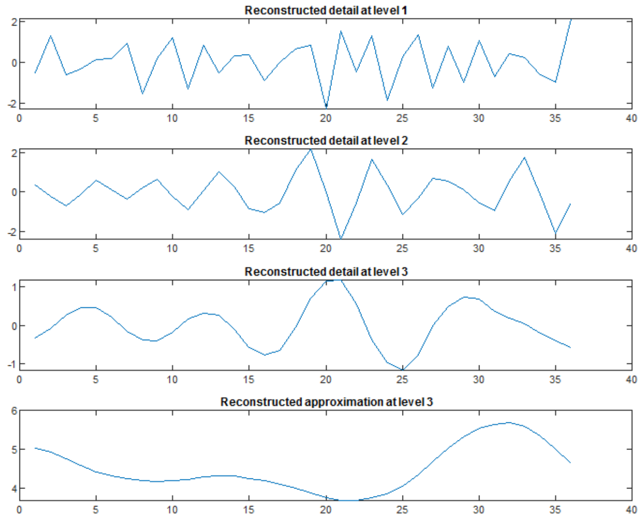

Fig -6: Reconstructed sub bands from detail and approximation coefficients of Dissolved Oxygen

2.6 Artificial Neural Network

ANN is a computational model based on the structure and functions of biological neural networks. Artificial Neural Networks are massively parallel distributed data processing systems consisting of a large number of highly interconnected artificial neurons with performance characteristics resembling biological neurons of the human brain. The processes in developing the neural network model can be mainly divided into 2 phases, the training phase and forecasting phase. For this purpose a training algorithm is made use of. The type of algorithm used in the study is feed forward back propagation algorithm.[5] Small Artificial Neural Networks are developed for prediction of each sub band. Figure 6 shows the preferred network with 3 layers. The first layer being the input layer with 3 neurons, the second layer being the hidden layer with 3 neurons and a third layer being the output layer with 1 neuron. Trained using the Levenberg-Marquardt Algorithm, with input for training being $x(t)$, $x(t-1)$ and $x(t-2)$, where x is each sub-band. Target for training will be $x(t+1)$ for input $[x(t),x(t-1),x(t-2)]$. During prediction, trained neural network produces $x(t+1)$ for each sub-band on providing $[x(t),x(t-1),x(t-2)]$ as inputs to the neural network, shown in figure 7. Predicted sub-bands for each water quality parameter are combined to form a single band. Combining of sub-bands can be done by just summing the sub-bands, forming the final prediction data or forecast data for each water quality parameters.

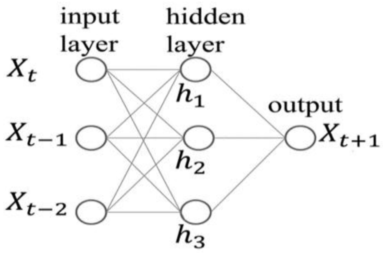

Fig -7: Multilayered Perceptron Back-propagation Neural Network

3. RESULT AND DISCUSSION

ANN models developed where able to predict the water quality parameters fairly well using small dataset, even though there were many unknown factors affecting the water quality parameters of a flowing river ANN models can work even with not so correlated predictors. All the values can be considered on the same foot while constructing single – hidden layer ANN models. To justify the best predictor combination of ANN model, the data available with Kerala Water authority (from 2011 to 2013) were used as training and validation dataset. In present work, the feed-forward neural network algorithm is used by the neural network tool in MATLAB software. The data in a neural network are categorized into three sets: training, validation and test. The result shows that the proposed artificial neural network-based model has a great potential to stimulate and predict the parameter which affects the water quality. Predicted results are shown in figure

Dissolved oxygen model

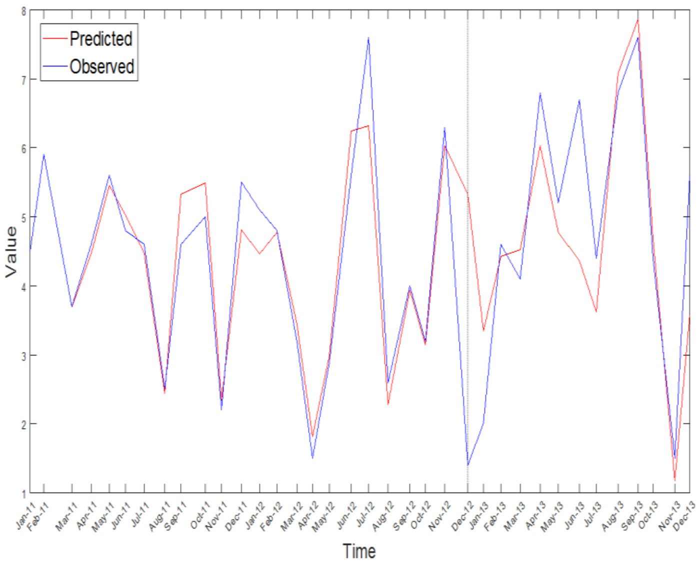

Fig -8 Dissolved Oxygen parameter predictions of Location 1 by ANN

The ANN-based intelligent predictive model was developed to simulate and predict the Dissolved Oxygen at Kallai River (in 2 sampling points). Most aquatic organisms needs oxygen to survive and grow. Figure 8 and figure 9 shows the Dissolved Oxygen parameter predictions by ANN

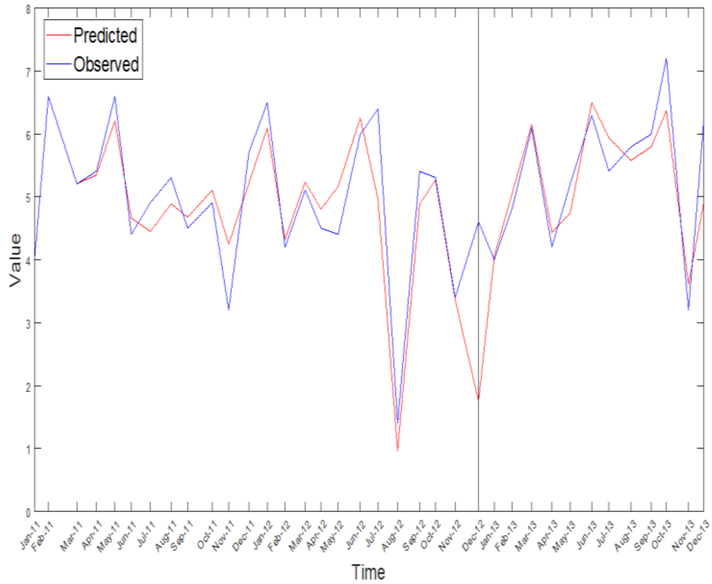

Fig -9 Dissolved Oxygen parameter predictions of Location 2 by ANN

COD Model

The ANN-based intelligent predictive model was developed to simulate and predict the COD at Kallai River (in 2 sampling points). The chemical oxygen demand (COD) is a parameter that measures all organics: the biodegradable and the non-biodegradable substances. ANN model predicts COD at a reasonable level. Figure 10 and figure 11 shows the COD parameter predictions by ANN

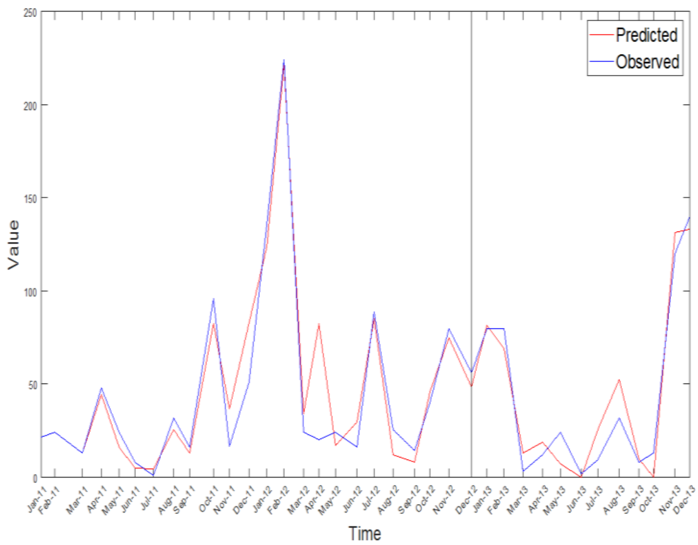

Fig -10 COD parameter predictions of Location 1 by ANN

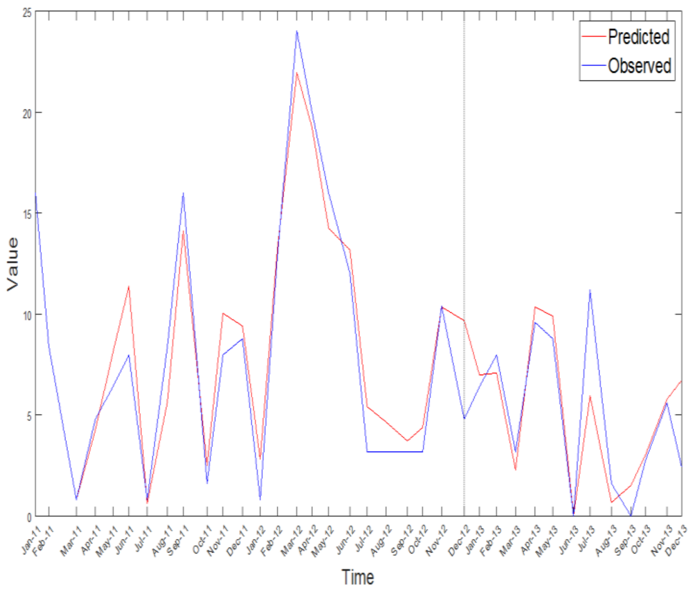

Fig -11 COD parameter predictions of Location 2 by ANN

Nitrate Model

The ANN-based intelligent predictive model was developed to simulate and predict the Nitrate at Kallai River (in 2 sampling points). The high concentration of nitrate, in surface water, can stimulate the rapid growth of the algae which degrades the water quality. Nitrates can enter the groundwater from chemical fertilizers used in the agricultural areas. Figure 12 and figure 13 shows the Nitrate parameter predictions by ANN

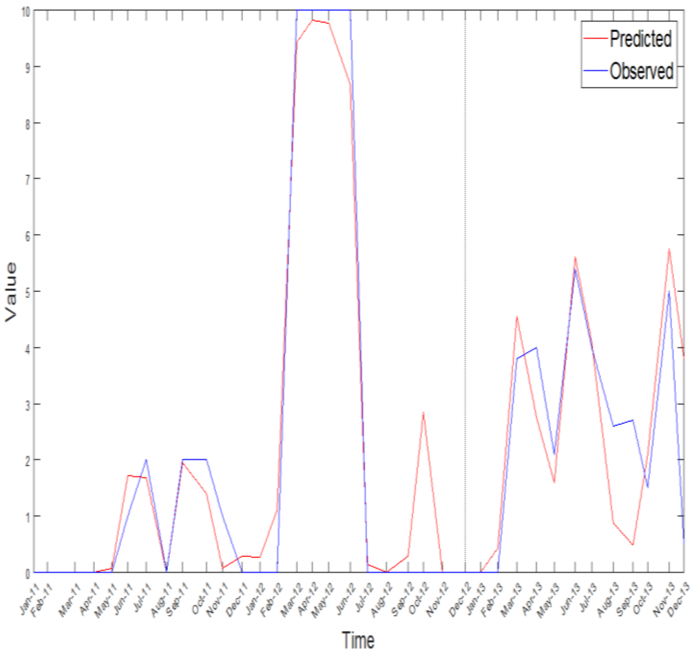

Fig -12 Nitrate parameter predictions of Location 1 by ANN

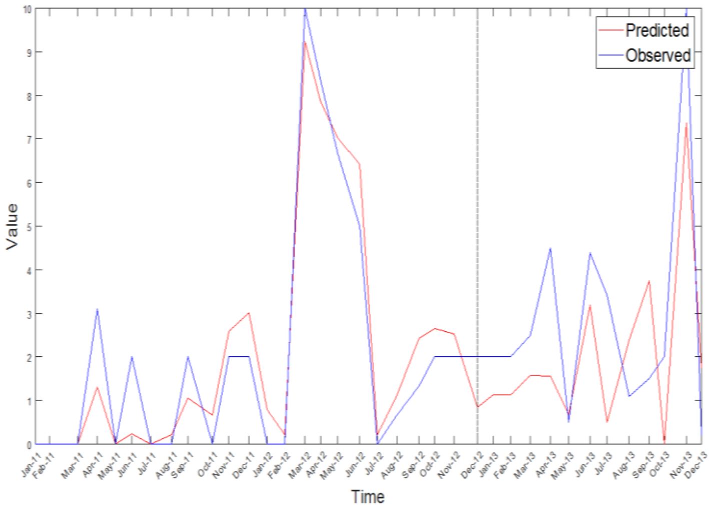

Fig -13 Nitrate parameter predictions of Location 2 by ANN

Turbidity Model

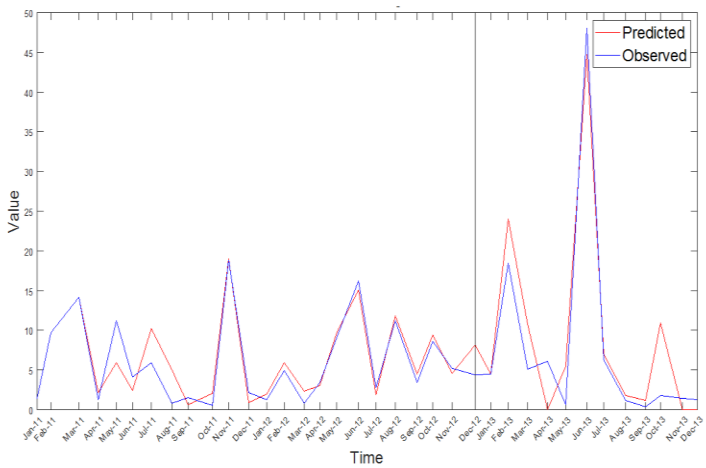

Fig -14 Turbidity parameter predictions of Location 1 by ANN

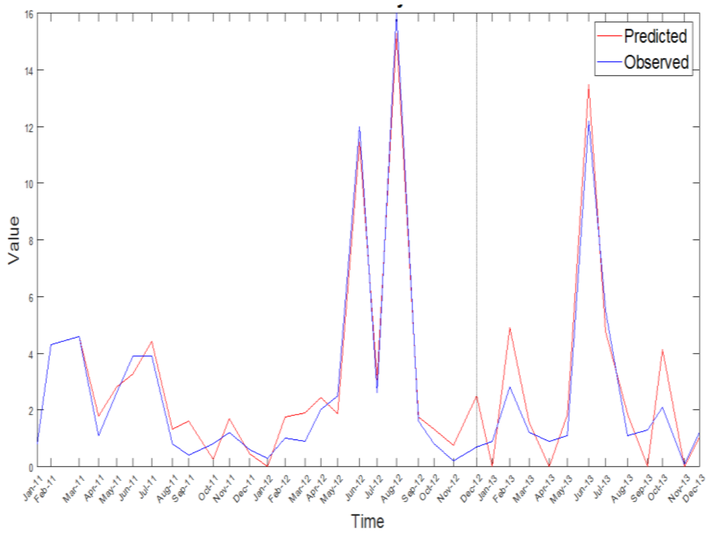

Fig -15 Turbidity parameter predictions of Location 2 by ANN

The ANN based intelligent predictive model was developed to simulate and predict the Turbidity at Kallai river. It is a measure of the amount of suspended particles in the water. Figure 14 and figure 15 shows the turbidity parameter predictions by ANN

pH Model

The ANN based intelligent predictive model was developed to simulate and predict the pH at kallai river. Generally, the pH of water has small variations due to buffering action with carbon dioxide. Regarding Kallai River the pH value lies in the permissible limit. Figure 16 and figure 17 shows the pH parameter predictions by ANN

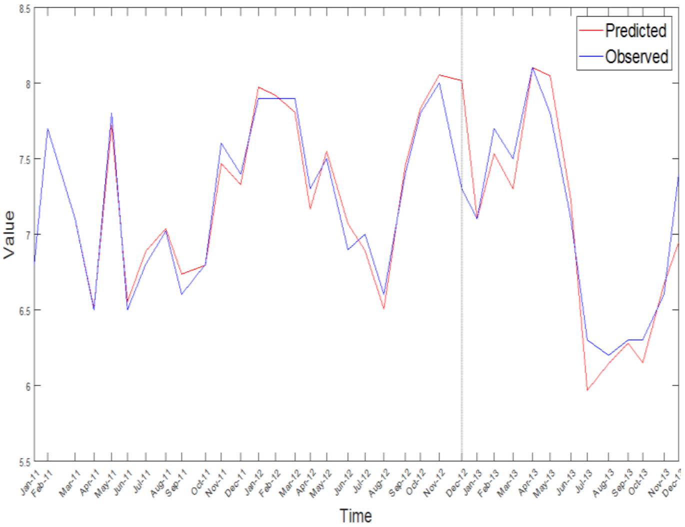

Fig -16 pH parameter predictions of Location 1 by ANN

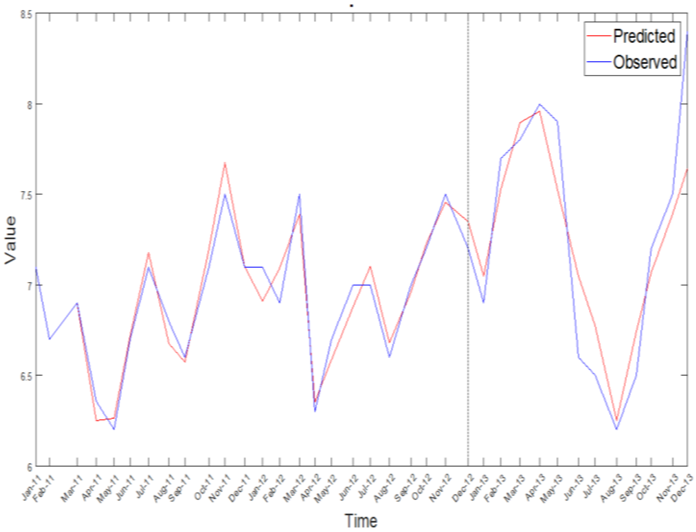

Fig -17 pH parameter predictions of Location 2 by ANN

ANN captures the embedded spatial and unsteady behavior in investigated problems using its architecture and non-linearity nature compared with other classical modelling techniques. Neural networks are being used in a wide variety of applications as an important decision making tool. This work suggests the use of ANN-based water quality parameter prediction model for Kallai River. Five main parameters have been studied viz pH, Turbidity, Nitrate, COD, DO, In fact, the proposed ANN model requires no prior knowledge of natural physical processes of these water quality parameters. Despite the highly stochastic nature of proposed water quality parameters, the proposed models are capable of mimicking the water quality parameters accurately with small prediction error. The ANN-based model exhibit robustness and reliable performance in predicting viz pH, Turbidity, Chloride, Iron, Nitrate, COD, DO, TC. It is possible to predict such a fair results using small networks because of the ability of the decomposition process, to decompose the time series data into sub band components, in which some of them were showing cyclic nature. Dealing with such cyclic components resulted in the fairly well results. Another optimization that enhanced the results were using the most appropriate wavelet filter for each parameters, which added to the performance. The root mean square error and R^2 value for the prediction result is shown in Table 1 and Table 2 respectively. The results can be further enhanced by adding more hidden layers to the artificial neural network, which also requires a larger dataset.

Table -1: Table shows the Root Mean Square Error for the prediction.

Dataset	Sampling point 1		Sampling point 2	
	Training	Testing	Training	Testing
pH	0.3319	0.4063	0.2757	0.8524
Turbidity	0.379	0.8155	0.1732	0.2962
Chloride	0.1267	0.3248	0.2596	0.4718
Iron	0.3956	0.2776	0.1667	0.4861
Nitrate	0.1993	0.3604	0.2918	0.6564
COD	0.3234	0.2243	0.2823	0.3317
DO	0.5679	0.661	0.6209	0.4335
Total Coliform	0.2251	0.0815	0.2431	0.2508

Table -2: Table shows the R^2 value for prediction.

Dataset	Sampling point 1		Sampling point 2	
	Training	Testing	Training	Testing
pH	0.8844	0.9016	0.918	0.7968
Turbidity	0.8486	0.8899	0.9677	0.8749
Chloride	0.9825	0.8707	0.927	0.9537
Iron	0.8444	0.8634	0.9698	0.7242
Nitrate	0.9572	0.4201	0.9094	0.4642
COD	0.8871	0.9401	0.9186	0.6715
DO	0.6487	0.6638	0.5905	0.7702
Total Coliform	0.9448	0.7646	0.938	0.8851

4. CONCLUSION

In this study, Artificial Neural Network and Wavelet were applied to two sampling points in Kallai River, to predict the future values of river water quality parameters producing a baseline for early identification of unprecedented changes in river water quality. This project shows the parameters can be predicted to a certain degree. Which is awesome to be able to do this with so little data. If the authorities are able to implement this for all the rivers. Then, it opens a new possibility for auto-detection of river water quality irregularities that may have happened due to a new source of pollutant or pollution sources, making it possible to provide early warnings. Which is not possible though looking at the collected data alone, due to seasonal and climatic changes that is being shifted due to human interference with nature. Also, the proposed method doesn't require any additional expenses or arrangements for implementations and the only raw material needed is the data, which is already collected by authorities like state water authorities and state pollution control boards. The proposed method is not just for the river but can be used anywhere where real-world data is collected. Which altogether makes it an optimal solution for automation towards preservation and protection of natural resources.

REFERENCES

[1] Polluting the river, [online], Available: https://ypte.org.uk/factsheets/river-pollution/polluting-the-rivers.

[2] Ammar Salman Dawood, Haleem K. Hussain, Aymanalak Hassan (2016), Modelling of river water quality parameters using Artificial Neural Network- A Case Study, International Journal of Advances in Mechanical and Civil Engineering, Volume 03, Issue 05.

[3] River water quality, [online], Available: https://www.lenntech.com/rivers-pollution-quality.html

[4] Gajendran C, Water quality assessment and prediction modelling of Nambiyar river basin, [online], Available: http://hdl.handle.net/10603/9477.

[5] Roshni R, and Elizabeth C Kuruvila. "BOD MODELLING USING ARTIFICIAL NEURAL NETWORK." International Journal of Advance Research And Innovative Ideas In Education 2.4 (2017): 32-41.

[6] Emamgholizadeh, Samad & Kashi, Hamed & Maroufpoor, Eisa & Zalaghi, (2013). Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models. International Journal of Environmental Science and Technology. 11. 645-656. 10.1007/s13762-013-0378-x.

[7] Archana Sarkar, Prashant Pandey. "River Water Quality Modelling using Artificial Neural Network Technique". International Conference on Water Resources, Coastal and Ocean Engineering (ICWRCOE 2015). Volume 4, 2015, Pages 1070-1077.

[8] Nidhisha C, Elizabeth C Kuruvila. "Modelling BOD and COD Using Artificial Neural Network with Factor Analysis". International Research Journal of Engineering and Technology (IRJET 2019). Volume: 06 Issue: 04 | Apr 2019.

[9] Joy Parmar, Mosin I Hasan, "Forecasting of River Water Quality Parameters". International Journal Of Scientific Research in Engineering – IJSRE-2015. Volume 01, Issue 05 May 2015.

[10] Sujana Prajithkumar, Dr Shweta Verma, S. J. Mane, "Prediction of Water Quality Index of Pavna River using ANN Mode", International Research Journal of Engineering and Technology (IRJET 2019). Volume 03, Issue 12 (December 2014).

[11] Akkaraboyina, Mahesh & Raju, B.S.N. "Assessment of water quality index of River Godavari at Rajahmundry". Universal Journal of Environmental Research and Technology. vol 2. Pages 161-167.

[12] Sarda, Dr Purushottam & Sadgir, Parag. "Computation of Water Quality Parameters and Prediction Tool ANN for Modeling of Water Quality of Reservoir". International Journal of Innovative Research in Science, Engineering and Technology. 4. 8906-8911. 10.15680/IJIRSET.2015.0409086.

[13] Nayla Hassan Omer, "Water Quality Parameters", Submitted: March 2nd 2019 Reviewed: September 10th 2019 Published: October 16th 2019. Doi:10.5772/intechopen. 89657.

[14] Bennett, Kristin P.; Campbell, Colin (2000). "Support Vector Machines: Hype or Hallelujah" (PDF). SIGKDD Explorations. 2 (2): 1–13. doi:10.1145/380995.380999.

[15] Quiet die the Kerala's rivers, [Online], Available: http://timesofindia.indiatimes.com/articleshow/62878503.cms.

[16] Waterpollution, [Online], Available: http://www.kerenvis.nic.in/Database/Waterpollution_834.aspx

[17] Salim Aijaz Bhat, Gowhar Meraj, Sayar Yaseen (2014) "Statistical Assessment of Water Quality Parameters for Pollution Source Identification in Sukhnag Stream: An Inflow Stream of Lake Wular (Ramsar Site), Kashmir Himalaya". Journal of Ecosystems. Hindawi Publishing Corporation. https://doi.org/10.1155/2014/898054.