ISSN No:-2456-2165

Development of Marking System based Water Quality Index for Chambal River and their Prediction Model Using Artificial Neural Network

Prerna Sengar[1], Jay Singh Rajput[2],
[1]Student of M. Tech., [2]Research Scholar
(Environmental Engineering),
Madhav Institute of Technology & Science,
Gwalior, Madhya Pradesh, India

Prof. A. K. Saxena[3]
[3]Associate Professor,
Madhav Institute of Technology & Science,
Gwalior, Madhya Pradesh,
India

Abstract:- In India, river qualities are getting degraded diurnally. Actually, it is due to dumping of waste, discharging of untreated water and industrial waste etc. in to the river. Consequently, the river gets contaminated due to such anthropogenic activities. In this regard, Water quality index (WQI) is widely utilized to monitor the quality of river. WQI is a single unique number which represent the quality status of river water. In this study, a marking system based WQI has been developed for Chambal River. In which, the permissible limits (mentioned in IS-Code) of each water quality parameters have been utilized for developing sub-Indices. Subsequently, an average operator has been applied to agglomerate the sub-indices in to a single number termed as WQI. To achieve this, water quality parameter data has been imported from Central pollution control board database that composed the concentration of Bio-Chemical Oxygen demand (B.O.D), Dissolved Oxygen (D.O.), Conductivity, pH, Nitrate, Total Coliform, and Faecal Coliform of 10 different location of Chambal River. Additionally, a prediction model also has been developed by using Artificial Neural Network with artificial dataset. An artificial dataset was generated by utilizing the actual dataset with random sampling. In this direction, Levenberg Marquardt (LM), Baysian Regularization (BR), and Scaled Conjugate Gradient (SCG) algorithms were trained and tested with different settings of hyper-parameter. As a result, Baysian Regularization algorithm shows good result (RMSE = 0.00, R^2 = 0.99) followed by Levenberg-Marquardt algorithm (RMSE = 1.89, R^2 = 0.99), Scaled Conjugate Gradient (RMSE = 1.94, R^2 = 0.97).

Keywords:- *Water quality index (WQI), Artificial Neural Network (ANN), Chambal River.*

I. INTRODUCTION

Not just for a state or a nation, but for all of mankind, water is the most vital natural resource. The careful utilization of this resource is crucial for nation's success. As a result, water, which flows in rivers and streams, may be said to be a nation's basic wealth. This proves the significance of rivers, and no more explanation is necessary to emphasize their significance. Because water knows no political borders, river basins have been recognised as a domain for planning and management all across the globe. India's rivers, which are revered by its people, are one of the country's most distinguishing characteristics. Indian rivers, which cover a large geographical area of 329 million hectares, have been a major contributor to India's rural development. Its various rivers are of tremendous significance to India in terms of cultural, economic, geographical, and religious growth. In India, rivers are revered as Gods and Goddesses, and Hindus even worship them. They provide visitors a fascinating glimpse into India's historical, cultural, and traditional features. The riverine system is a unique sort of ecology among the numerous types of inland fresh water basins. Each watershed's quality and features are determined by the size of the drainage basin, the volume of water flowing through the system, the percentage of natural vs settled regions, and man's direct influence.

The country's water crisis has begun to influence people's lives as well as the environment surrounding them. The following are some of the key concerns that need immediate attention:

Drinking water is scarce in many regions of the nation during the summer months due to excessive groundwater extraction to fulfill agricultural, industrial, and household needs.

Approximately 10% of the rural and urban populations do not have regular access to clean drinking water, and many more are at risk. To satisfy their daily requirements, the majority of them rely on contaminated water sources. Furthermore, water shortages in towns and villages have resulted in significant amounts of water being gathered and delivered by tankers and pipelines across long distances. In this regards, Fluoride, arsenic, and selenium, among other chemical toxins, represent a severe health risk to the people. Roughly 70 million people in 20 states are thought to be at danger from excessive fluoride, while around 10 million people are thought to be at risk due to excessive arsenic in ground water. Aside from that, the rise in chloride, TDS, nitrate, and iron concentrations in groundwater is a major source of worry for a long-term drinking water program. All of these issues must be addressed comprehensively. The concentration of dissolved constituents/ionic concentrations is steadily rising as a result of excessive groundwater extraction.

The purposes of the investigation of WQI are (i) To provide an overview of the water quality of the basin, (ii) To determine the spatial distribution so that the trend of the

ISSN No:-2456-2165

water quality can be assessed for future development plans, (iii) To map surface and groundwater quality changes in the study area using GIS and Geo-statistical techniques, (iv) To identify potential equivalences between Artificial Neural Networks (ANN) and statistical regression model to find the best modelling approach for the study area, and (v) To give recommendations based on the study of the Chambal River Basin to water quality management authorities on how the results can be integrated for sustainable catchment management strategies of the basin.

The present study considers the following objectives: (i) To identify optimized locations for different usages in the present study area. (ii) To study the water quality trends in the basin. (iii) To identify the interrelationship among the bio-physico and chemical parameters of the basin water quality using a statistical approach for both surface and ground water.(iv) To find the Water Quality Index (WQI) of the basin.

II. STUDY AREA

The Chambal River originates from the summit of Janapav Hill in the Vindhyan range in Mhow, Indore, M.P., India and merge with Yamuna River in Etawaha district of U.P. India see Figure 1. It has been protected as the National Chambal Sanctuary for the conservation of Gharial. The area is semiarid and the temperature in the region varies from 2^0 to 48^0C during winter and summer, respectively. The Chambal River has an average width of 400 m, and a depth of 1–26 m. The riverbanks are rocky in the upstream areas leading to muddy and sandy banks further downstream and characterized by forest, shrub, and grasses. Unlike many rivers of the greater Ganges River system, the Chambal River is relatively unpolluted. In this study, various location of Chambal River has been considered (see Table 1).

S.No.	LOCATION	STATE
1	Chamal Entering Rajasthan At Gandhi Saga	Rajasthan
2	Chambal At Gandhi, Near Dam Rampura	Madhya Pradesh
3	Chambal At Jawnpawa, From Origin Point Indore	Madhya Pradesh
4	Chambal At Kota D/S (2km) From City	Rajasthan
5	Chambal At Kota U/S Intake Pt.Near Barrage	Rajasthan
6	Chambal At Nagda D/S	Madhya Pradesh
7	Chambal At Nagda U/S Water Intake Point	Madhya Pradesh
8	Chambal At RameshwarghatNr.Sawaimadhopur	Rajasthan
9	Chambal At Tal Village Near Bridge Ujjain	Madhya Pradesh
10	R. Chambal At Dholpur	Madhya Pradesh

Table 1: Various location of Chambal River

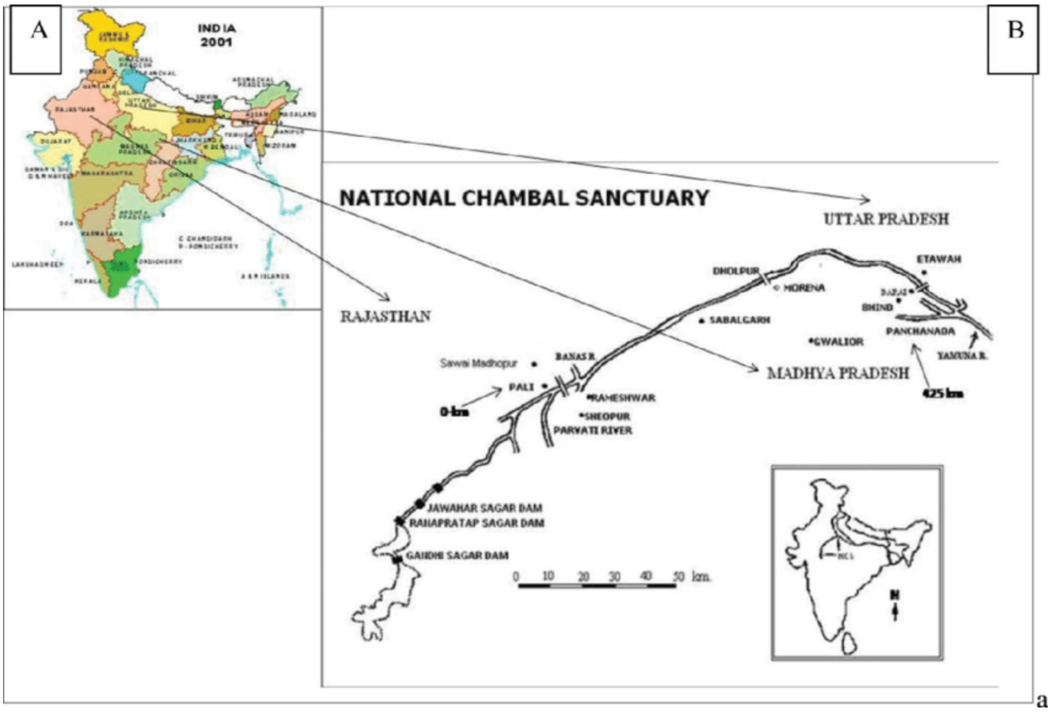


Fig. 1: Map of Chambal River and their tributaries.

III. RESEARCH METHODOLOGY

A. Data Acquisition, preparation, and Normalization

The water quality data of Chambal River are directly acquired from Central Pollution Control Board (CPCB) database. Further, dataset is arranged and prepare in MS Excel. In general, each water quality parameters have their own dimensions to represent water quality. This condition may rises problem in the prediction of dataset in prediction model. Therefore, data normalization is necessary to perform using Equation (1) to normalize the values of the measured parameters that will be used in the computations in order to increase prediction accuracy. In this way, the entirevariable comes under a same dimensionless category.

$$Normalized\ Data\ (x') = \frac{x - min(x)}{max(x) - min(x)} \quad (1)$$

B. Development of Water Quality Index

The purpose of the water quality index is to notify relevant water body authorities about the quality of the water bodies so that they may make informed decisions. Because water quality must be tied to a specific purpose such as aquaculture, irrigation, and drinking, combining the numerous factors into a single number that depicts whether the water is excellent, terrible, or poor is a big difficulty that the water quality index attempts to answer. The following are the actions that must be taken in order to construct a water quality index and also see Figure 2:

- Choosing water quality parameters.
- Evaluation of each quality parameter's and its upper and lower concentration limits.
- Creating of Sub-indices.
- The value of the water quality index is worked by combining sub-indices.
- Ranking of water quality index value is ranked from bad to excellent.

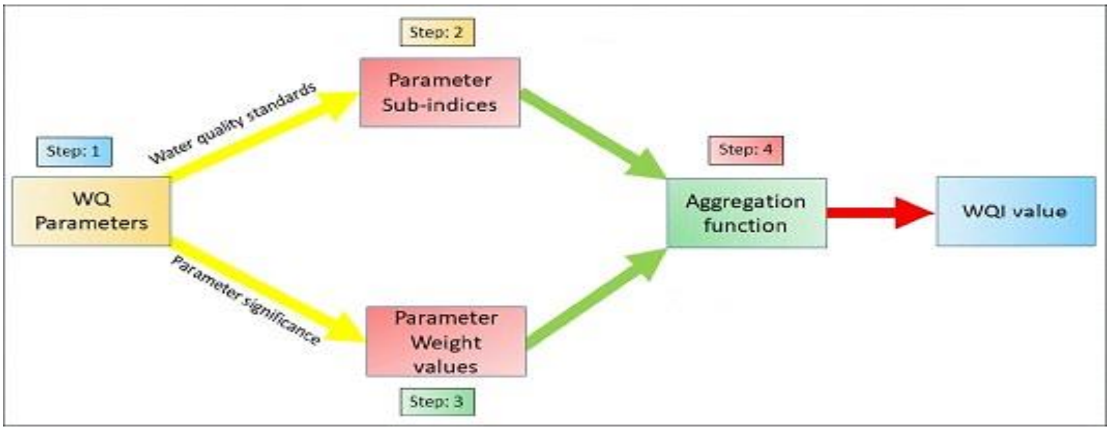


Fig. 2: General steps involved in development of water quality index

In this manner, selected water quality parameters are pH, D.O., B.O.D, Conductivity, Nitrate, Faceal Coliform and Total Coliform. In which, pH signify about the nature of the water in terms of acidity and alkalinity, D.O. accounted for presence of excess amount of available oxygen in the river, B.O.D stands for representing availability of organic contaminate or need of oxygen, Conductivity shows presence of ions, Nitrate accounted for availability of nitrogen ions, and Faceal& Total Coliform represent the amount of pathogens and microorganism available in the river. Therefore, all these parameters have shown chemical and biological characteristic of river water and are sufficient to represent the quality of the water. Subsequently, the standard limits of each water quality parameters have been assessed to define their upper and lower limits as shown in Table 2.

S. No.	Parameters	Upper limits	Lower Limits
1.	D. O.	12 mg/l	5 mg/l
2.	pH	7	<6.5, >8.5
3.	Nitrate	0	45 mg/l
4.	B.O.D	0	3 mg/l
5.	Faecal Coliform	0	2500
6.	Total Coliform	0	5000
7.	Conductivity	0	300 mg/l

Table 2: Upper and lower limit of each water quality parameters

Furthermore, a water quality index based on a marking system is created, which is analogous to a student's test report card. Existing literature has been extensively examined in this respect, with the goal of extracting the most essential river water quality measure. To accomplish it, sub-indices have been determined with respect to the permissible limits of water quality parameters. In this process, a marking scale from 0 to 100 is formed with least count of 10 in which 30 accounted for the permissible criteria, water quality parameter has been assigned to fail category if their sub-indices found below 30. Therefore, the lower limit of each water quality parameters have been set to 30 sub-indices and their upper limits are assigned to 100 sub-indices and further values are assigned with equal division between upper and lower limits in the scale. These scales are shown in the Table 3.

Sub-Indices	100	90	80	70	60	50	40	30	20	10	0
pH Value - 1	7	7.21	7.42	7.63	7.84	8.05	8.26	8.47	8.68	8.89	9.1
pH Value - 2	7	7	6.9	6.8	6.7	6.7	6.6	6.5	1.4	0.7	0
D.O.	12	11	10	9	8	7	6	5	4	3	2
Nitrate Conc.	0	6.43	12.86	19.29	25.72	32.2	38.58	45	51.44	57.87	64.3
B.O.D	0	0.43	0.86	1.29	1.72	2.15	2.58	3	3.44	3.87	4.3
Faecal Coliform	0	357	714	1071	1428	1785	2142	2500	2856	3213	3570
Total Coliform	0	714.3	1429	2142.9	2857.2	3571.5	4285.8	5000	5714.4	6428.7	7143
Conductivity	0	42.86	85.72	128.58	171.44	214.3	257.16	300	342.88	385.74	428.6

Table 3: Sub-Indices scale of each water quality parameter

Finally, an average of all sub-indices were determined to agglomerate the sub-indices of each water quality parameter and which further account for water quality index. Also, marking system for water quality index has been created to represent the final quality of the river shown in Table 4.

Status	Score Secured
Excellent	81-100
Good	61-80
Moderate	41-60
Poor	21-40
Very poor	0-20

Table 4: Marking System of water quality index.

C. Development of prediction Model

a) Artificial Neural Network:

ANN is a network-based system that can handle huge dataset in parallel processing manner. All activities related to data processing, data recognition, data prediction, and other difficult activities can be possible with ANN. In general, architecture of ANN is composed of three layers such as input layers, hidden layers, and output layers as shown in Fig 3. In which, input layer receives raw data in the form of specified variables and sends it to be processed further. Thereafter, hidden layer comes after the input layer and processes the data for result validation. At last, the output layer is used to display the current result. Further, ANN has wide application in Decision Making, Risk Assessment,Construction costs productivity, Prediction and forecasting, Strategy and business study etc.

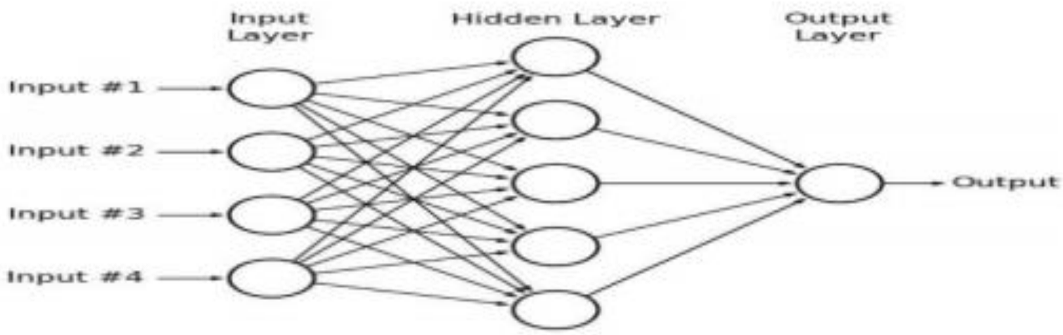


Fig. 3: Architecture of ANN

In this study, a multilayer processing structure (Feed-forward neural network) is utilized. In which, each layer is meant to work independently and to display the outcome before moving on to the next. Many hidden layers work between the input and output layers to conduct complicated tasks. Further, a Back-propagation algorithm approach is utilized to optimize the outcomes. The proposed network's weights are modified using back-propagation methods in ANN. These algorithms repeat cycles in order to get the desired result. Along with it, three training algorithms were utilized for training of ANN model such as Levenberg-Marquardt – it have high efficiency since it regenerates the weight values after each cycle, Baysian Regulization - This approach takes longer to run, but it produces good results for challenging and huge datasets, and Scaled Conjugate Gradient - The benefit of employing this method is that it uses less memory and produces quicker results. Additionally, MATLAB version 2020 user interface of ANN has been used for all computation and development of prediction model.

D. Evaluation of prediction Model

a) Root mean square error (RMSE):
This error function is used to assess the model's performance by comparing the actual and predicted data from the training and testing datasets as shown in equation (2). One will get a trustworthy result if the error value is near to zero.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Predicted\ Outcome_i - Actual\ Outcome_i)^2}{N}} \quad (2)$$

b) Correlation:
This statistic is used to assess how well the actual and anticipated datasets match using equation (3). The correlation value of real and anticipated data may also be used to confirm the model's performance. The correlation value ranges from 0 to 1, with 1 denoting a strong link.

$$R = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (3)$$

IV. RESULTS AND DISCUSSION

The actual datasets are directly imported from CPCB database. Further, more datasets are required to develop a prediction model using ANN. For this reason, an artificial datasets are generated through random sampling by using descriptive statistic of actual dataset. In this way, 1500 data points have been generated for each water quality parameter that further utilized as an input parameter in the prediction model during training of each ANN algorithms. The descriptive statistic of actual and artificial datasets is shown in Table 5. Further, WQI has been calculated using marking system for actual and artificial dataset as shown in figure 4.

Descriptive Statistic	pH		D.O.		Conductivity		BOD		Nitrate		Fecal Coliform		Total Coliform	
	Actual	Artificial	Actual	Artificial	Actual	Artificial	Actual	Artificial	Actual	Artificial	Actual	Artificial	Actual	Artificial
Mean	6.3	6.3	7.8	7.8	1260.2	1284.5	5.1	5.6	1.4	1.4	2966.8	4276.0	3871.0	4701.9
Standard Error	0.2	0.0	0.1	0.0	362.7	53.7	1.5	0.2	0.2	0.0	2706.5	479.7	2610.8	425.5
Standard Deviation	1.2	1.2	0.5	0.6	2115.1	2078.0	8.6	9.3	1.2	1.1	15781.2	18580.3	15223.6	16481.3
Sample Variance	1.4	1.4	0.3	0.3	4473809.0	4317980.1	74.8	86.2	1.4	1.3	249046201.1	345226165.9	231758918.8	271632612.2
Kurtosis	0.8	0.6	12.3	8.8	8.5	6.5	6.3	4.0	8.2	7.1	33.5	18.3	21.6	14.2
Skewness	-0.7	-0.7	-2.6	-2.5	3.0	2.7	2.7	2.3	2.5	2.4	5.8	4.5	4.6	3.9
Range	5.3	5.3	3.2	3.2	9488.5	9488.5	35.3	35.3	5.9	5.9	92002.0	92000.5	80800.0	80796.5
Count	34.0	1500.0	34.0	1500.0	34.0	1500.0	34.0	1500.0	34.0	1500.0	34.0	1500.0	34.0	1500.0

Table 5: Descriptive Statistic of Actual and Artificial Dataset of Water Quality Parameters

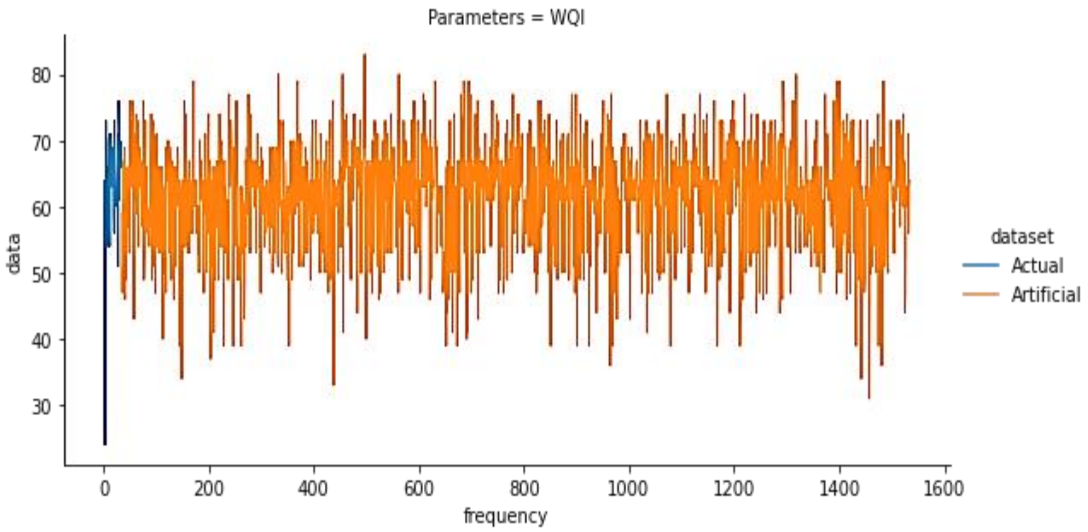


Fig. 4: Data distribution of calculated water quality index of actual and artificial datasets

A. Artificial Neural Network based Prediction Model:

There are only three ANN based algorithm was developed for prediction of water quality index. In this process, some common hyper-parameter of algorithms were selected such as data splitting (Training, Validation & Testing), and Number of hidden neuron. Further, the optimized hyper-parameter value was obtained with hit & trial approach, in which 140 timeseach algorithm was run to obtain RMSE value and Coefficient of correlation between actual outcome and predicted outcome. Subsequently, the best combination was selected based on their Root mean square error (RMSE) and coefficient of correlation (R^2).

B. Training of the Model with Artificial Dataset

The algorithm was run several times to obtained optimized hyper parameter. In this way, no. of hidden neuron was selected in each training set based on their minimum RMSE value. The obtained optimize number of hidden neuronin different algorithm during training of the prediction model with artificial datasets are shown in Table 6.

Training Set	**Levenberg-Marquardt**			**BaysianRegulization**			**Scaled Conjugate Gradient**		
	Hidden Neuron	**RMSE**	**R^2**	**Hidden Neuron**	**RMSE**	**R^2**	**Hidden Neuron**	**RMSE**	**R^2**
50	16	0.02	1.00	13	0.01	1.00	12	0.08	0.87
55	17	0.03	0.98	13	0.01	1.00	19	0.10	0.80
60	11	0.02	0.99	19	0.01	1.00	18	0.31	0.82
65	11	0.02	0.99	20	0.01	1.00	18	0.10	0.77
70	13	0.02	0.99	18	0.01	0.97	13	0.10	0.76
75	14	0.02	0.99	19	0.01	1.00	17	0.10	0.79
80	13	0.02	0.99	16	0.01	1.00	13	0.02	0.69

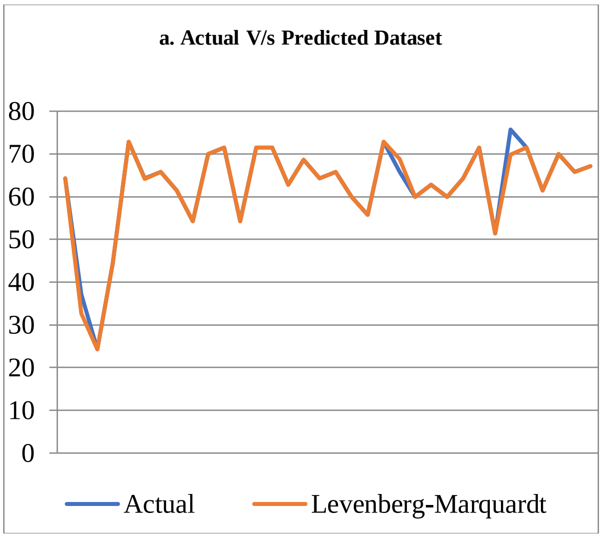
Table 6: Optimized no. of hidden neuron in different ANN algorithm

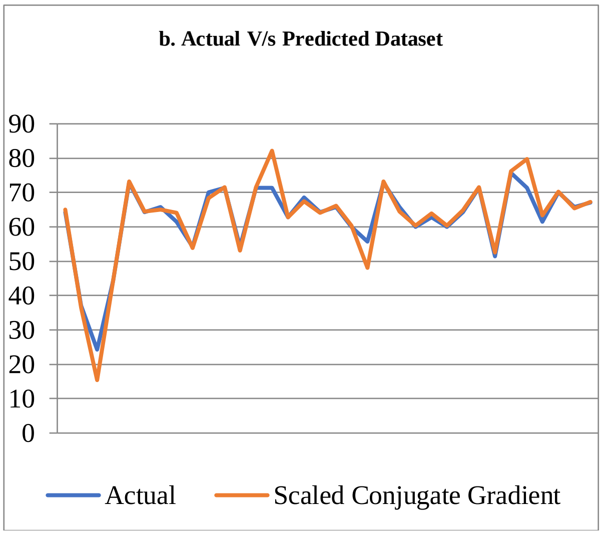
C. Testing of the Model with Actual Dataset

Subsequently, the actual dataset was tested in the different algorithm of ANN. This time the hyper-parameters related to each algorithm were taken from the Table 6. In this manner, each algorithm was run with actual dataset as an input parameter and obtained their Root mean square error (RMSE) and coefficient of correlation (R^2) as shown in Table 7. It is observed that the Baysian Regularization algorithm shows RMSE = 0.00 and R^2 = 0.99 and represent as a best model for the prediction of water quality index. However, Levenberg-Marquardt algorithm shows RMSE = 1.89 and R^2 = 0.99, and Scaled Conjugate Gradient shows RMSE = 1.94 and R^2 = 0.97 promising result. Therefore, any of the algorithm can be utilized for prediction of water quality index.

Algorithm	Data Split	No. Hidden Neuron	RMSE	Correlation
Levenberg-Marquardt	75	14	1.89	0.99
Baysian Regularization	70	18	0.00	0.99
Scaled Conjugate Gradient	50	12	1.94	0.97

Table 7: Best ANN based predicting model of water quality index





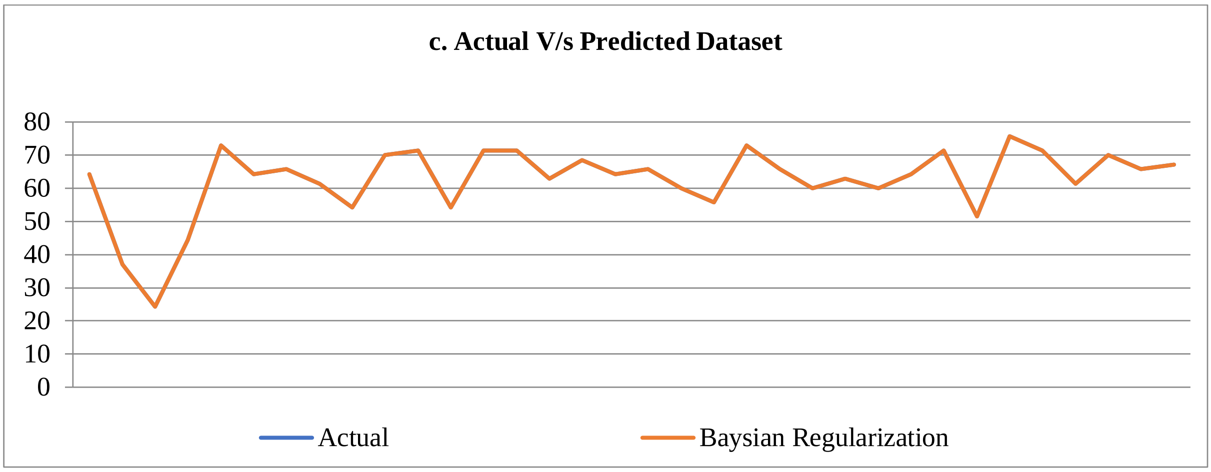


Fig. 5: Actual V/s Predicted Outcomes a. Levenberg-Marquardt algorithm, b. Scaled Conjugate Gradient algorithm, and c. Baysian Regularization algorithm.

V. CONCLUSION

The present study is dedicated to develop a new methodology for the computation of water quality index of Chambal River and develop its prediction model using artificial neural networks based back-propagation feed forward algorithms. To accomplish this task, water quality data (included pH, D.O., Conductivity, B.O.D, Nitrate, Faecal Coliform, Total Coliform) is directly imported from CPCB official website during 2012 to 2018 at the different stream of Chambal River from 10 location. Further, a unique marking system approach is used develop water quality index. In this process, a scale was formed based on permissible limits of each water quality parameter that further used to transform water quality parameter data into sub-indices. Thereafter, an average operator was applied on the sub-indices of all water quality parameters to aggregate it into a single value. Further, a prediction model was developed using back-propagation feed forward algorithms of artificial neural networks such as Levenberg-Marquardt, Baysian-Regularization, and Scaled Conjugate Gradient. In this context, hyperparameters (No. hidden Neuron, sets of training and testing data) of the algorithm were optimized on the basis minimum root mean square error criteria in hit and trial manner. As a result, Baysian Regularization algorithm shows good result (RMSE = 0.00, Correlation = 0.99) followed by Levenberg-Marquardt algorithm (RMSE = 1.89, Correlation = 0.99), Scaled Conjugate Gradient (RMSE = 1.94, Correlation = 0.97).

REFERENCES

[1.] chauhan and S. singh, "Evaluation of Ganga water for drinking purpose by water quality index at Rishikesh, Uttarakhand, India" Rep opin, vol.2, no. 9, pp. 53-61, 2010.

[2.] Ambica, "Groundwater quality characteristics study by using water quality index in Tambaram area, channai, Tamil Nadu" Middle East J Sci Res, Vol.20, no.11, pp. 1396-1401,2014.

[3.] R. Saxena and M. Sharma, "Qualitative and quantitative evaluation of water sources of few areas in and around Gwalior, Madhya Pradesh, India" Journal of Environmental Research and Development, vol.8, no.3, pp. 459, 2014.

[4.] S. S. Parihar, A. Kumar, R. N. Gupta, m. Pathak, A. Shrivastava & A. C. Pandey, "Physico-chemical and microbiological analysis of Underground water in and around Gwalior city, MP, India" Research journal of recent sciences, vol.1, no.6, pp. 62-65,2012.

[5.] N. kalara, R. Kumar, S. S. Yadav and R. T. Singh, "Water qualiyy index assessment of groundwater in koilwar block of Bhojpur (Bihar)" Journal of chemical and pharmaceutical Research, vol.4, no.3, 2012.

[6.] R. Shah & A. Tiwari, "Study of physicochemical parameters of Underground water at Takenpur, Gwalior, MP".

[7.] S. Das, P. K. Roy and A. Mazumdar, "Development of water quality index for groundwater in Kolkata city, West Bengal, India" ARPN Journal of Engineering and Applied Sciences, vol.8, no.12, 2013.

[8.] K. Rai, B. Paul, N. Kishor "A study on the sewage Disposal on Water Quality of Harnu River in Ranchi City Jharkhand, India" International Journal of plants Animals and Environmental science, vol.2, pp. 102-106,2012.

[9.] K. sharma, N. Parashar and R. sharma, "Monitoring of water quality of Yamuna River at Mathura, U.P – Physicochemical cgaracteristics" International journal of Research in environmental science and technology, vol.3, no.4, pp. 156-159, 2013.

[10.] J. N. Halder and M.N. Islam, "water pollution and its impact on the Human Health" Journal of research in environmental science and technology, vol.1, no.4, pp. 1-5, 2008.

[11.] Singh and U. C. Singh, "Assessment of Groundwater quality of parts of Gwalior (India) for agricultural purposes" Indian Journal of science and technology, vol.1, no.4, pp. 156-159-2013.

[12.] S. Paramar and V. Singh, "Water quality parameters of river Yamuna in Delhi after 20 years of the Yamuna action plan" International Journal of science and nature, vol.6, no.4, pp. 662-669,2015.

[13.] S. Tyagi, B. Shuman, P. Singh and R. Dobhal, "Water quality assessment of water quality index"

International Journal of water Resources, vol.1, no.3, pp. 34-38, 2013.
[14.] M. K. Chaturvediand J. K. Bassin, "Assessing the Water quality Index of Water treartment plant and bore wells, in Delhi, India," Environ. Monil. Assess, Vol. 163. pp. 449-453, 2010.
[15.] D. Kumar and B. Alappai, "NSF-Water quality index: Does it represent the experts opinion" pract, period. Hazard, Toxic Radioact, Waste manage, vol.13, no.1. pp. 75-79. 2009.
[16.] S. Krishna kumar, A. logeshkumaran, N. S. Magesh, P. S. Godson, N. Chandrasekhar, " Hydro-geochemistry and application of water quality index for groundwater Quality assessment, Anna Nagar, part of Chennai city, Tamil nadu, India" 2014.
[17.] Li Pei-Yue, QianHUi and Wu Jian-Hua, " Groundwater quality assessment based on improved water quality index in pengyan country, Ningxia, Northwest China" E- Journal of chemistry, 2010.
[18.] APHA 2005- Standard methods for the examination of water and waste water, 21st Ed. Washington D.C.
[19.] BIS 1991- Drinking water specification IS No. 10500, Bureau of Indian Standards.
[20.] WHO (1993) Guidelines for drinking water quality Recommendations, 2nd edition, world health organization, Geneva.
[21.] Chemistry for Environmental Engineering, 4th Edition clair N. Sawyer, Perrg Z Me. Carty, Gener F. Parkin.
[22.] Guide manual: Waste Water and Drinking water, Central pollution control Board, New Delhi, httpc//www.cpcb.nic.in/upload/latest/latest_67_guidentmanualw&wwananalysis.pdf (Accessed January 2020).