e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

River Water Level Prediction Modelling using Artificial Neural Network and Multiple Linear Regression

Pankaj Goswami[1] and Haina Brahma[2]
[1]Assistant Professor, Civil Engineering Department, Assam Engineering College, INDIA
[2]Student, Civil Engineering Department, Assam Engineering College, INDIA

[2]Corresponding Author: hainabrahma1004@gmail.com

ABSTRACT

Nowadays, Prediction modelling has become one of the most popular research areas among researchers/scientists around the world. In this study, the size of the training data is about 60%, validation data and testing set is about 20% of the total available data. In this paper, we have developed and tested feed-forward neural network architectures optimized with Levenberg-Marquardt back-propagation with transig activation function in hidden and output layers in predicting monthly river water elevation. Also, in this approach, the multiple linear regression equation to estimate monthly river water level was generated by using precipitation, discharge and return period as predictor variables. In this project, the results show the coefficient of determination (R^2) between the predicted and actual output using both Artificial Neural Network and Multiple Linear Regression model for the monthly peak, monthly average and monthly minimum of Brahmaputra, Pagladia and Puthimari River.

Keywords- Feed-Forward Neural Network, Levenberg-Marquardt Back Propagation, Prediction Modelling, Transig Activation Function, Multiple Linear Regression, Coefficient of Determination

I. INTRODUCTION

Assam, a state in North-eastern region of India, is full of natural resources and agricultural state and with its vast network of rivers. Assam State is comprised of two valleys namely the Brahmaputra and Barak Valley and it is situated in between 90°-96° North Latitude and 24°-28° East Longitude. It is prone to natural disasters like flood and erosion which has a negative impact on overall development of the state.

Predicting flood disasters are good potential research areas for its impact to the publics and the economics of the affected country [1]. Based on the application of a parallel artificial neural network (ANN) model the approach uses state variables, input and output data, and previous model errors at specific time steps in order to predict the errors of a physically based model [2]. In recent years research on modern coastal water level modelling and prediction techniques has been growing concerns [3].

Flood prediction modelling has become one of the most famous research areas among the researchers or scientists all around the globe. The objective of this project is to study the river water level prediction modelling of Brahmaputra River, Pagladia River and Puthimari River using Neural Network Toolbox of Matlab Software and also using Multiple Linear Regression method.

II. STUDY AREAS

For this project work, the study was done for the Brahmaputra River and its two tributaries namely, Pagladia River and Puthimari River. Figure 1 shows map showing the position of Pagladia River, Puthimari River and Brahmaputra River.

Figure 1: Map showing the position of Pagladia, Puthimari and Brahmaputra River

The data used for this study have been collected the discharge data and water level data from the Water Resource Department under Lower Assam Investigation Division and the rainfall data from World Weather Online. The data used in this project are rainfall, discharge and water level for the year 2008 to 2017. Then all the data are arranged into monthly peak, monthly average and monthly minimum order for each river.

III. METHODOLOGY

This work is licensed under Creative Commons Attribution 4.0 International License.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

A. Artificial Neural Network

An Artificial Neural Network (ANN) is a highly inter-connected network of many simple processing units, called neurons. Neurons in an ANN are arranged into groups, called layers. Neuron in each layer operates in logical parallelism. ANN is also called as black-box model, used for modelling complex hydrological processes like rainfall-runoff modelling, water quality modelling, groundwater modelling, and precipitation prediction.

ANN consists of layers of neurons. The model is characterized by a network of three layers of simple processing units, which are put together to each other. The first layer is called an input layer, which receives input information. The third layer is called an output layer, which generates output information. Between output and input layers, there are hidden layers. There can be one or more hidden layers and information is transmitted through the connections between nodes in different layers. The available data were split into three data sets: training set, validation set and testing set to assess the forecasting performance of model. Figure 2 shows Artificial Neural Network Architecture.

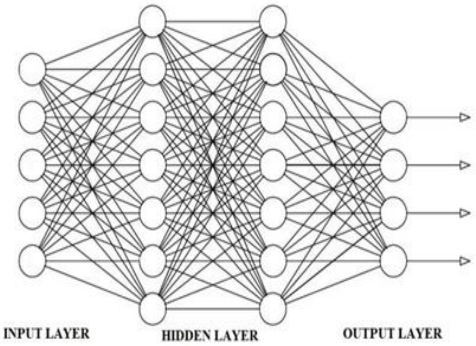

Figure 2: Artificial Neural Networks Architecture

B. Multiple Linear Regression

In a model of simple linear regression, a single response measurement Y is related to a single predictor X for each observation. The critical assumption of the model is that the conditional mean function is linear. Equation 1 shows a simple linear regression equation.

$$\boldsymbol{E\left(\frac{y}{x}\right) = a + bx \ldots\ldots\ldots (1)}$$

In a multiple linear regression model, the numbers of predictor variables are more than one. Equation 2 shows a multiple linear regression equation. This leads to the following "multiple regression" mean function:

$$\boldsymbol{E\left(\frac{y}{x}\right) = a + b_1x_1 + b_2x_2 + \ldots\ldots\ b_nx_n \ldots\ldots (2)}$$

Where a is called the intercept and the b_n are called slopes or coefficients.

The multiple linear regression equation to estimate monthly river water level was generated by using precipitation, discharge and return period as predictor variables.

IV. RESULTS AND DISCUSSION

A. Artificial Neural Network

The data in neural networks are categorised into three sets; training, testing and validation. The size of the training data is 60%, validation data and testing set is 20 % of the total available data. In this paper, we have developed and tested feed-forward neural network architectures optimized with Levenberg-Marquardt back-propagation with transig activation function in hidden and output layers in predicting monthly river water level. The networks had been trained and tested with 10 years of data using Matlab software. Prediction accuracy has been measured by means of mean square error (mse) and correlation coefficient (r). The tests are done for monthly peak, monthly minimum and monthly average data each for the Brahmaputra River, Pagladia River and Puthimari River. In this approach, 10 numbers of hidden layers are used for the test.

Figure 3, 4 and 5 show the regression output for the monthly peak, monthly average and monthly minimum data of Brahmaputra River respectively.

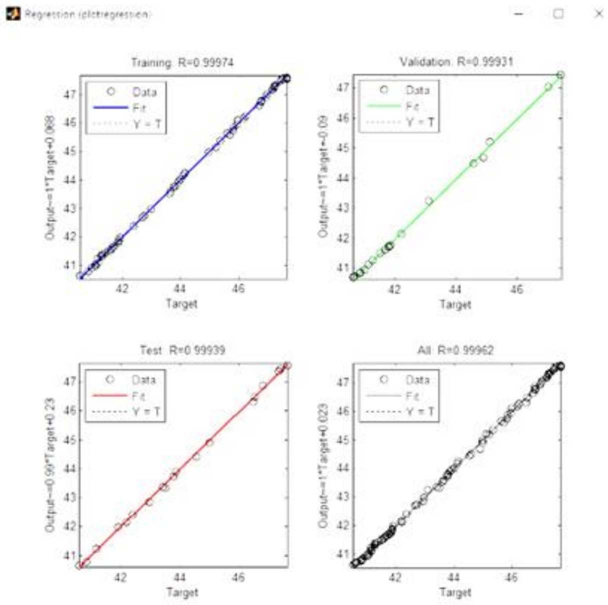

Figure 3: Regression output by ANN for monthly peak of Brahmaputra River

This work is licensed under Creative Commons Attribution 4.0 International License.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

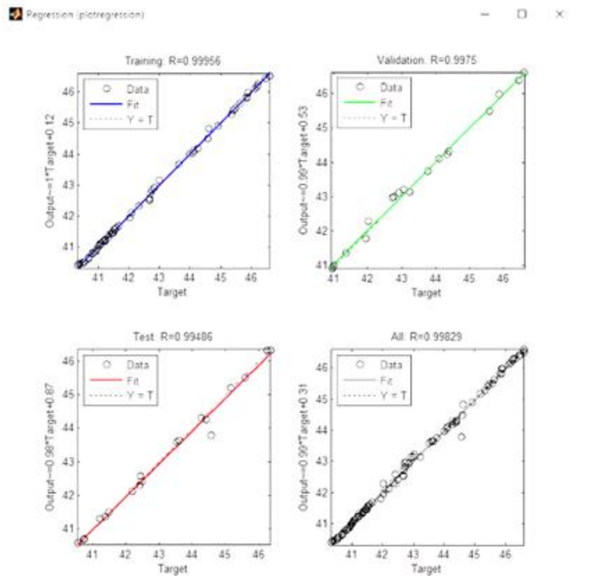

Figure 4: Regression output by ANN for monthly average of Brahmaputra River

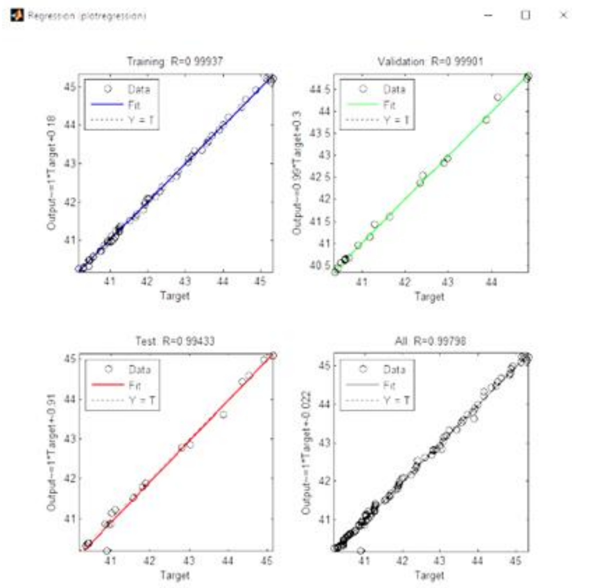

Figure 5: Regression output by ANN for monthly minimum of Brahmaputra River

The regression output gives the correlation coefficients (R) for the monthly peak, monthly average and monthly minimum data are 0.99962, 0.99829 and 0.99798 for the Brahmaputra River as shown in figures 3, 4 and 5 respectively.

Figure 6, 7 and 8 show the regression output for the monthly peak, monthly average and monthly minimum data of Pagladia River respectively.

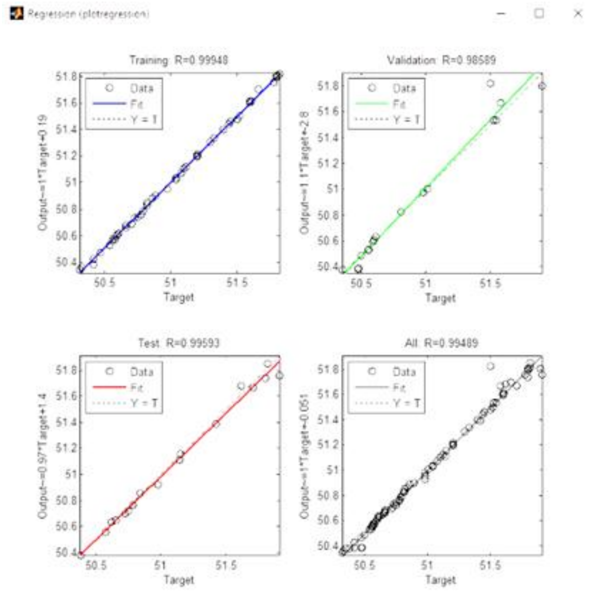

Figure 6: Regression output by ANN for monthly peak of Pagladia River

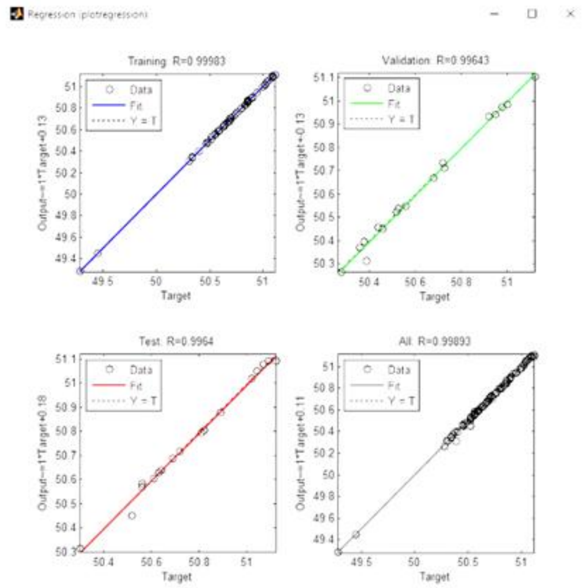

Figure 7: Regression output by ANN for monthly average of Pagladia River

This work is licensed under Creative Commons Attribution 4.0 International License.

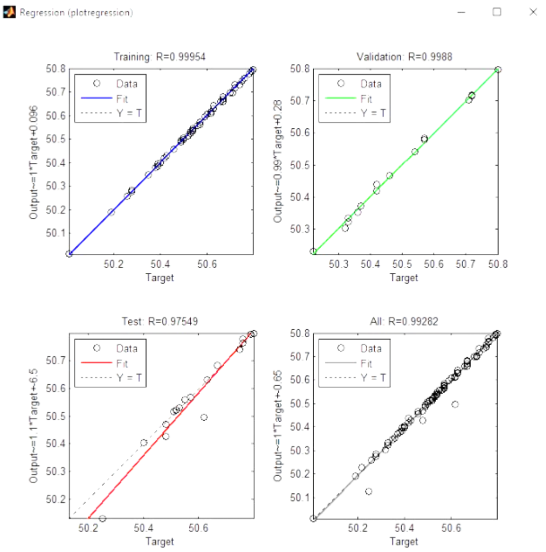

Figure 8: Regression output by ANN for monthly minimum of Pagladia River

The regression output gives the correlation coefficients (R) for the monthly peak, monthly average and monthly minimum data are 0.99489, 0.99893 and 0.99282 for the Pagladia River as shown in figure 6, 7 and 8.

Figure 9, 10 and 11 show the regression output for the monthly peak, monthly average and monthly minimum data of Puthimari River respectively.

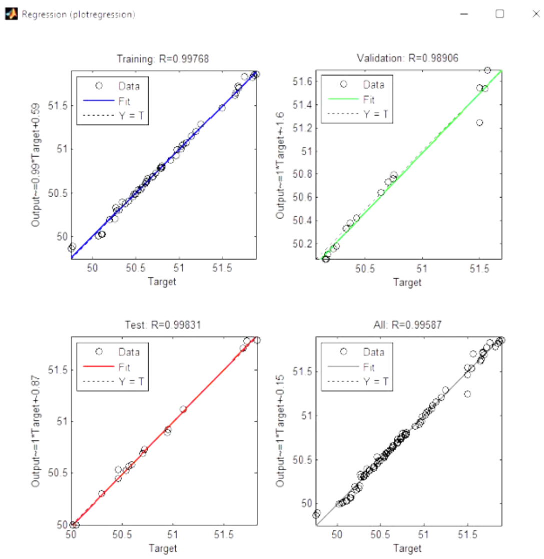

Figure 9: Regression output by ANN for monthly peak of Puthimari River

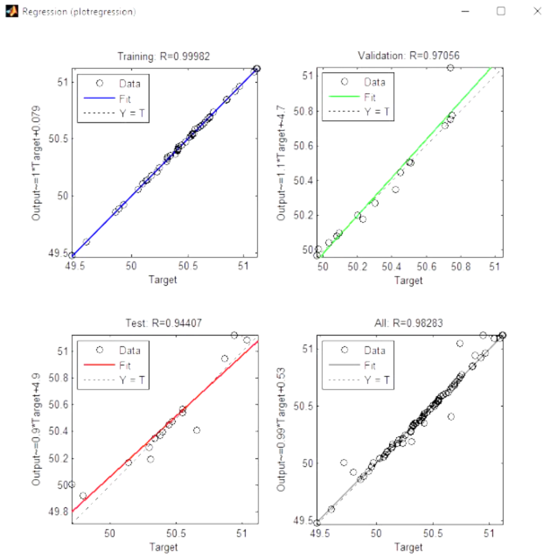

Figure 10: Regression output by ANN for monthly average of Puthimari River

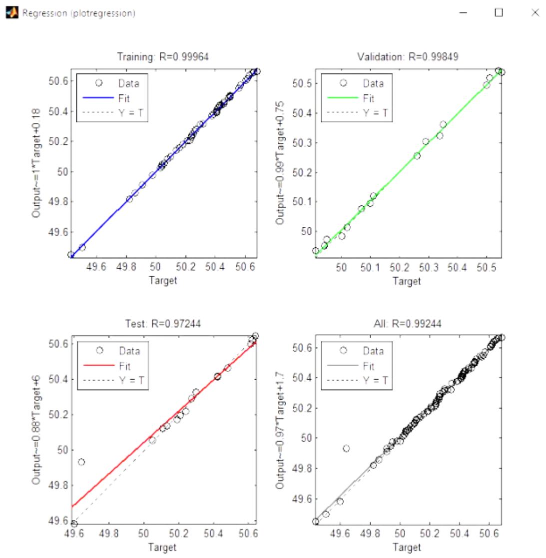

Figure 11: Regression output by ANN for monthly minimum of Puthimari River

The regression output gives the correlation coefficients (R) for the monthly peak, monthly average and monthly minimum data are 0.99587, 0.98283 and 0.99244 for the Puthimari River as shown in figure 9, 10 and 11.

The predicted outputs from the ANN simulations are saved and the predicted outputs are compared with the actual or observed outputs. Table 1 shows the coefficient of determination (R^2) between the predicted and observed or actual water level of Brahmaputra River, Pagladia River and Puthimari River.

 This work is licensed under Creative Commons Attribution 4.0 International License.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

Table 1- The coefficient of determination (R^2) between predicted and observed or actual water level

Rivers	Peak	Average	Minimum
Brahmaputra River	0.9673	0.958	0.9517
Pagladia River	0.888	0.9582	0.953
Puthimari River	0.8213	0.9437	0.9667

The predicted and actual output plots are shown in figure 12, 13 and 14 for the monthly peak, monthly average and monthly minimum of Brahmaputra River

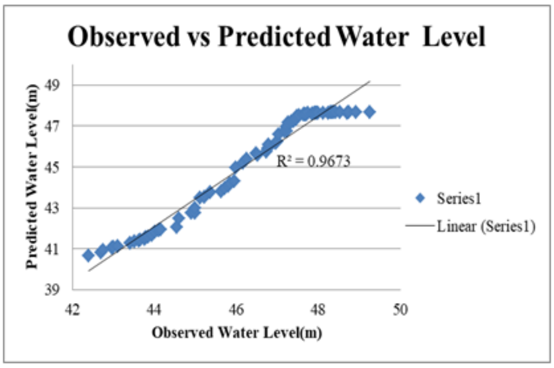

Figure 12: Observed and Predicted water level plot for monthly peak using ANN of Brahmaputra River respectively.

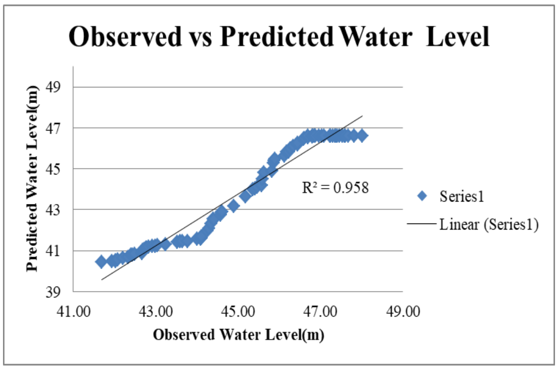

Figure 13: Observed and Predicted water level plot for monthly average using ANN of Brahmaputra River

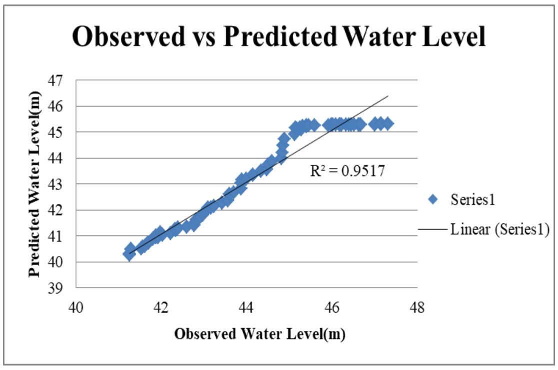

Figure 14: Observed and Predicted water level plot for monthly minimum using ANN of Brahmaputra River

The predicted and actual output plots are shown in figure 15, 16 and 17 for the monthly peak, monthly average and monthly minimum of Pagladia River respectively.

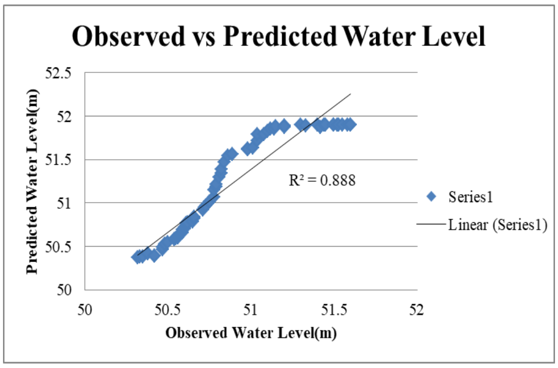

Figure 15: Observed and Predicted water level plot for monthly peak using ANN of Pagladia River

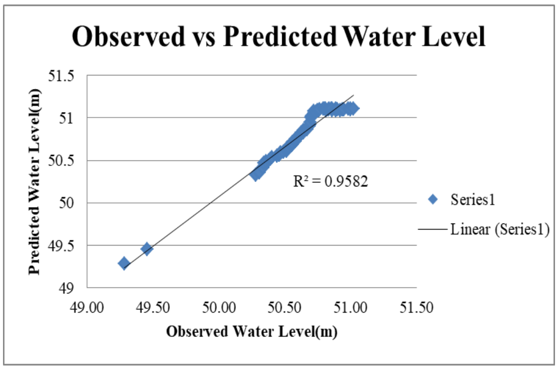

Figure 16: Observed and Predicted water level plot for monthly average using ANN of Pagladia River

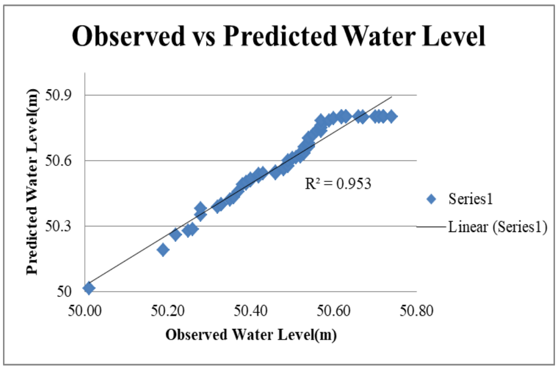

Figure 17: Observed and Predicted water level plot for monthly minimum using ANN of Pagladia River

The predicted and actual output plots are shown in figure 18, 19 and 20 for the monthly peak, monthly average and monthly minimum of Puthimari River respectively.

This work is licensed under Creative Commons Attribution 4.0 International License.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

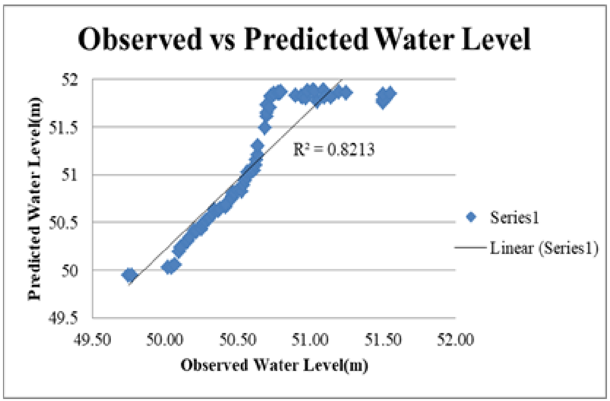

Figure 18: Observed and Predicted water level plot for monthly peak using ANN of Puthimari River

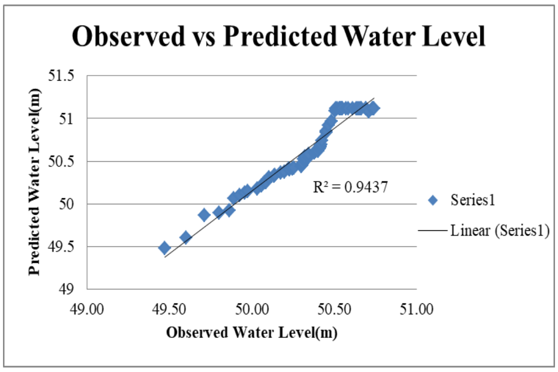

Figure 19: Observed and Predicted water level plot for monthly average using ANN of Puthimari River

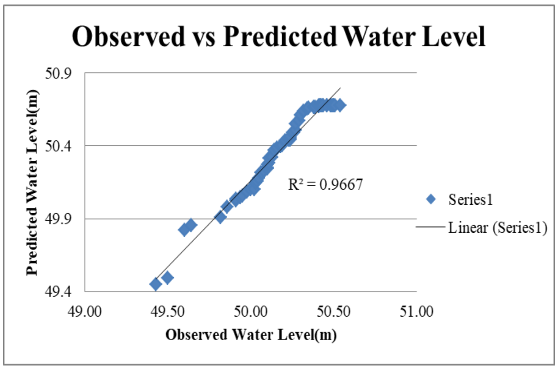

Figure 20: Observed and Predicted water level plot for monthly minimum using ANN of Puthimari River

Multiple Linear Regression

The multiple linear regression equations to estimate monthly peak, average and minimum water level at Pandu Gauge station were generated by regression model as equation 1, 2 and 3 respectively.

$$\textbf{Water Level} = 39.91 + 6.0954e^{-2} * t_r + 5.8198e^{-4} * pptn + 6.7571e^{-5} * Q \;\ldots\ldots(1)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Brahmaputra River from equation 1 is 0.981. The predicted and actual output plot is shown in figure 21.

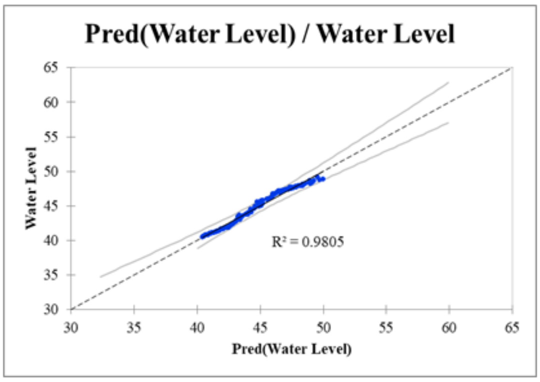

Figure 21: Observed and predicted plot for monthly peak using Multiple Linear Regression of Brahmaputra River

$$\textbf{Water Level} = 39.41 + 3.4903e^{-2} * t_r + 1.0309e^{-2} * Pptn + 1.5213e^{-4} Q \;\ldots\ldots(2)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Brahmaputra River from equation 2 is 0.9972. The predicted and actual output plot is shown in figure 22.

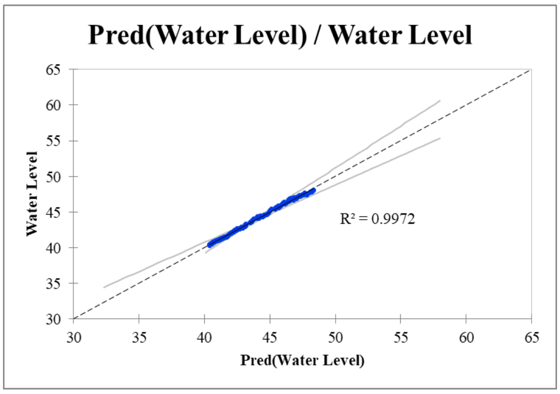

Figure 22 Observed and predicted plot for monthly average using Multiple Linear Regression of Brahmaputra River

$$\textbf{Water Level} = 39.11 + 3.3717e^{-2} * t_r - 2.9187e^{-3} * Pptn + 1.4944e^{-4} * Q \;\ldots\ldots(3)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Brahmaputra River from equation 3 is 0.9914. The predicted and actual output plot is shown in figure 23.

This work is licensed under Creative Commons Attribution 4.0 International License.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

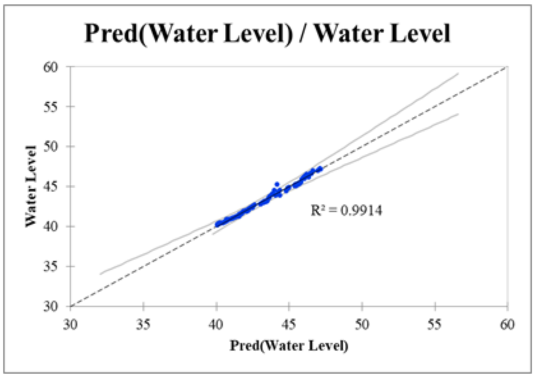

Figure 23 Observed and predicted plot for monthly minimum using Multiple Linear Regression of Brahmaputra River

The multiple linear regression equations to estimate monthly peak, average and minimum water level at Pagladia N.T. Road X-ing Gauge station were generated by regression model as equation 4, 5 and 6 respectively.

$$\textit{Water Level} = 50.27 - 1.3119e^{-4} * \textit{Pptn} + 4.1444e^{-3} * Q + 1.3685e^{-2} * t_r \ldots\ldots(4)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Pagladia River from equation 4 is 0.9796. The predicted and actual output plot is shown in figure 24.

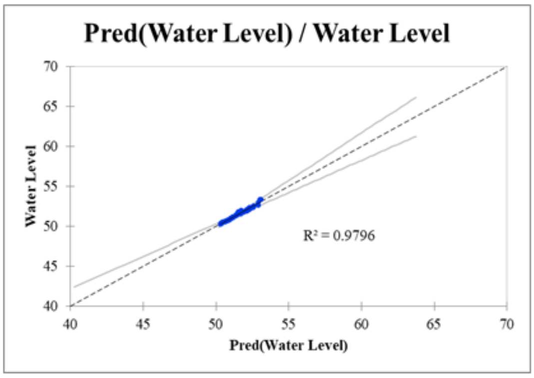

Figure 24 Observed and predicted plot for monthly peak using Multiple Linear Regression of Pagladia River

$$\textit{Water Level} = 50.17 - 7.3769e^{-4} * \textit{Pptn} - 2.0189e^{-5} * Q + 1.2679e^{-2} * t_r \ldots\ldots(5)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Pagladia River from equation 5 is 0.8859. The predicted and actual output plot is shown in figure 25.

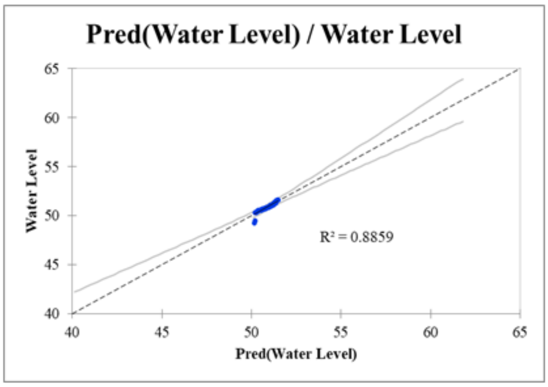

Figure 25 Observed and predicted plot for monthly average using Multiple Linear Regression of Pagladia River

$$\textit{Water Level} = 50.24 - 1.4164e^{-3} * \textit{Pptn} + 3.1660e^{-3} * Q + 6.6322e^{-3} * t_r \ldots\ldots(6)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Pagladia River from equation 6 is 0.9366. The predicted and actual output plot is shown in figure 26.

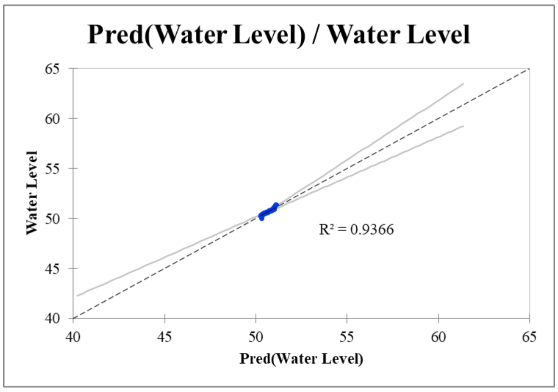

Figure 26 Observed and predicted plot for monthly minimum using Multiple Linear Regression of Pagladia River

The multiple linear regression equations to estimate monthly peak, average and minimum water level at Puthimari N.H. X-ing Gauge station were generated by regression model as equation 7, 8 and 9 respectively.

$$\textit{Water Level} = 49.66 + 7.2620e^{-4} * \textit{Pptn} + 2.3204e^{-3} * Q + 2.4573e^{-2} * t_r \ldots\ldots(7)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Puthimari River from equation 7 is 0.8997. The predicted and actual output plot is shown in figure 27.

This work is licensed under Creative Commons Attribution 4.0 International License.

e-ISSN: 2250-0758 | p-ISSN: 2394-6962
Volume- 9, Issue- 6 (December 2019)
https://doi.org/10.31033/ijemr.9.6.4

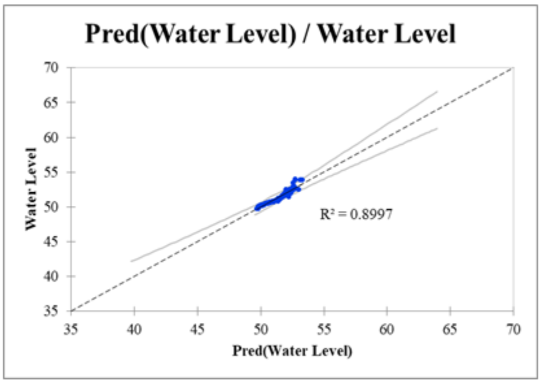

Figure 27 Observed and predicted plot for monthly peak using Multiple Linear Regression of Puthimari River

$$\boldsymbol{Water\ Level} = 49.78 - 2.6946e^{-5} * \boldsymbol{Pptn} + 3.2283e^{-3} * \boldsymbol{Q} + 1.4509e^{-2} * \boldsymbol{t_r} \ldots\ldots(8)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Puthimari River from equation 8 is 0.9559. The predicted and actual output plot is shown in figure 28.

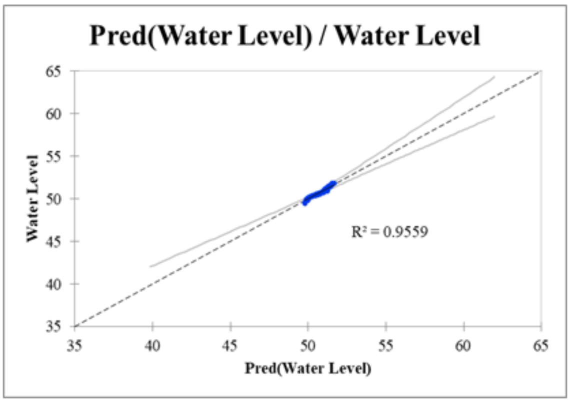

Figure 28 Observed and predicted plot for monthly average using Multiple Linear Regression of Puthimari River

$$\boldsymbol{Water\ Level} = 49.74 - 6.9974e^{-3} * \boldsymbol{Pptn} + 3.9380e^{-3} * \boldsymbol{Q} + 1.1555e^{-2} * \boldsymbol{t_r} \ldots\ldots(9)$$

Where, t_r= Recurence Interval, pptn= Precipitation, Q= Discharge.

The coefficient of determination (R^2) between the predicted and actual output of Puthimari River from equation 9 is 0.9388. The predicted and actual output plot is shown in figure 29.

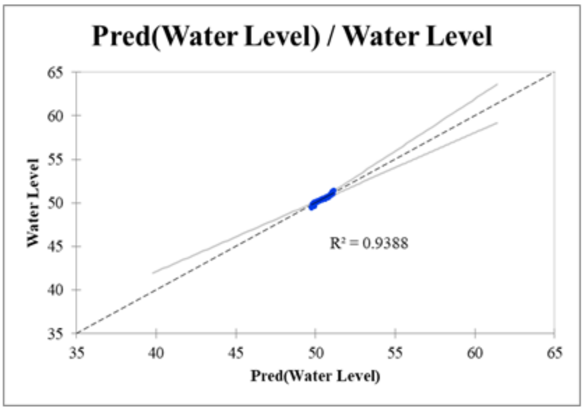

Figure 29 Observed and predicted plot for monthly minimum using Multiple Linear Regression of Puthimari River

V. CONCLUSIONS

River water level prediction modelling of Brahmaputra River and its tributaries Pagladia and Puthimari River are carried out using Artificial Neural Network (ANN) and Multiple Linear Regression. Table 2 shows the all R^2 value for both Artificial Neural Network and Multiple Linear Regression.

Table 2- All R^2 value for both Artificial Neural Network and Multiple Linear Regression

MODEL	RIVERS	MONTHLY PEAK	MONTHLY AVERAGE	MONTHLY MINIMUM
ARTIFICIAL NEURAL NETWORK	BRAHMAPUTRA	0.9673	0.958	0.9517
	PAGLADIA	0.888	0.9582	0.953
	PUTHIMARI	0.8213	0.9437	0.9667
MULTIPLE LINEAR REGRESSION	BRAHMAPUTRA	0.981	0.9972	0.9914
	PAGLADIA	0.9796	0.8859	0.9366
	PUTHIMARI	0.8997	0.9559	0.9388

The coefficient of determination (R^2) for River Brahmaputra, Pagladia and Puthimari have shown satisfactory results as R^2 value very close to 1 for both ANN and Multiple Linear Regression model.

ACKNOWLEDGEMENT

At the very outset, I am highly contented to express my sincere and heartfelt gratitude to my respected guide Dr. Pankaj Goswami, Assistant Professor, Department of Civil Engineering, Assam Engineering College for his valuable guidance, constructive suggestions and full co-operation throughout the course of this study and also in preparing and finishing the report by his diligent scrutiny and correction of the manuscript.

I would also like to express my sincere and heartfelt gratitude to Dr. Palash Jyoti Hazarika, Professor and H.O.D, Department of Civil Engineering, Assam Engineering College for his constant encouragement and for providing the necessary facilities to carry out the project work.

This work is licensed under Creative Commons Attribution 4.0 International License.

I also thank all the faculty members and staff of Department of Civil Engineering, Assam Engineering College for providing valuable help and support to carry out the work.

I also express my gratitude to all my classmates and well-wishers for their constant encouragement, valuable advice and inspiration throughout the work.

REFERENCES

[1] Adnan R., Samad A.M, & Ruslan F.A. (2016). A 3-hours river water level flood prediction model using NNARX with improves modelling strategy. *International Conference on System Engineering and Technology (ICSET)*, pp. 23-27.

[2] Abebe A.J. & Price R.K. (2004). Information theory and neural networks for managing uncertainty in flood routing. *Journal of Computing in Civil Engineering ASCE*, *18*(4), 373-380.

[3] Badejo, Temitope O., Uduodo, & Daniel. (2014). Modelling and prediction of water level for a coastal zone using artificial neural networks. *International Journal of Computational Engineering Research*, *4*(6), 26-41.

[4] The MathWorks Inc. (2009). *Neural network toolbox for use with MATLAB*. Available at: https://www.mathworks.com/access/helpdesk/help/toolbox/neuralnetwork/.

[5] Water Resource. (2019). *Government of Assam*. Available at: https://awrmis.assam.gov.in/.

[6] https://www.worldweatheronline.com.

[7] https://en.m.wikipedia.org/wiki/Brahmaputra_River.

[8] https://www.waterresources.assam.gov.in/portlet-innerpage/brahmaputra-river-system.

This work is licensed under Creative Commons Attribution 4.0 International License.