VOLUME 22 | NUMBER 01 | JULY 2022
https://ejournal.worldconference.id/index.php/neutron
E-ISSN: 2685-3272 | P-ISSN 1412-0860

Modeling of Flow Rate at Sukkur Barrage using Artificial Neural Networks (ANNs)

Rizwan Jokhio, M. Munir Babar, Pir M. Ajmal
US-Pakistan Center for Advance Studies in Water, Mehran UET, Jamshoro, Sindh, Pakistan
ahmedrizwanjokhio@gmail.com, mmunirbabar.uspcasw@faculty.muet.edu.pk, pirmohammadajmalqureshi@gmail.com.

Abstract

Modeling of flow discharge plays a significant role in effective planning, sustainable usage, development, and management of water resources in short (hourly) and long-term (monthly) temporal categories. Since the inception of managing water resources, various techniques such as conceptual, metric, and physical models have been introduced all of these require a large amount of data, labor, and expense to be incorporated to obtain reliable results, due to which Artificial Intelligence methods were introduced that require less amount of data, time, expense and as well as experience to model flow discharge. In this research study, an attempt was made by employing two different artificial neural network techniques feedforward neural networks (FFNN), and time-lagged neural networks (TLNN) to model and predict the river flow discharge at daily and monthly timescale. 2010 and 503 no. of observations were used for model calibration and validation in daily time scale while 557 and 139 observations were used in monthly timescale. The result of the study revealed that the FFNN modeling approach has captured the daily and monthly stream flow variability very well than the TLNN model with R2 of 0.91 on the daily and 0.71 on the monthly time scale while R2 for the TLNN model was 0.79, and 0.34 for daily and monthly timescale. This indicates that the FFNN technique requires less no. of observations and is more reliable than TLNN and can be used to model river flow.

Keywords :
Artificial Neural Networks, Feedforward Neural Networks, Machine Learning, River Flow Discharge, Stream Flow Prediction, Sukkur Barrage, Timeseries Analysis, Time-lagged Neural Networks

1. Introduction

Rivers serve as a significant source of fresh water for living beings (Bisoyi et al., 2019). Therefore modeling of river flow is of prime concern to guarantee the sustainable development, effective planning, development, and management of water resources (Parisouj et al., 2020). In an irrigation system, information on the flow rate (Q) is vital for various purposes such as; planning, designing, and operating hydraulic structures, managing water distribution among stakeholders, assessing extreme flood events, adopting mitigation measures, and drought warnings and reservoir operation (Hussain & Khan, 2020; Yaseen et al., 2015).

There are two broad modeling approaches used to model river flow: physically based models and data-driven models (Masselot et al., 2016). Physically-based models require a large number of data, finance, experience, complex mathematical tools, and human effort to build, calibrate, and validate the models(Khazaee Poul et al., 2019), while data-driven models are self-adaptive, more easily developed, and capable of capturing non-linear processes numerically using limited data inputs with no knowledge of the underlying physical processes involved (Noori & Kalin, 2016; Wu & Prasad, 2017; Yaseen et al., 2015). Since 1970, many data-driven approaches including linear regression (LR), multi-linear regression (MLR), moving average (MA), and auto-regressive moving average (ARMA) have been used for the modeling of river flow (Yaseen et al., 2015). These models were less accurate, linear, and unable to capture non-linearity b/w hydrological data (Melesse et al., 2011), so researchers have concentrated on developing innovative prediction models which will need not only to be inexpensive but also to be highly accurate and reliable (Hussain & Khan, 2020).

An artificial neural network is an innovative powerful tool, which is inexpensive, accurate, and reliable in terms of capturing non-linear temporal variation in data (Cheng et al., 2020), it is also capable of handling and analyzing large datasets (Buyukyildiz & Kumcu, 2017), so to extract/ recognize the patterns in the historical data, applies those to predict future scenarios (Khazaee Poul et al., 2019). The architecture of an Artificial Neural Network is composed of an input layer, a hidden layer, and an output layer. These three layers contain neurons that are connected through the connecting link called nodes. Compared to conventional methods, artificial intelligence methods can address the noise dataset in a non-linear dynamic system (Adedigba et al., 2017). Artificial Neural Networks have been utilized in various applications of hydrology engineering. Artificial neural networks have been utilized for solving various hydrological problems that include; groundwater simulation (Taormina et al., 2012), modeling of rainfall-runoff relationship (Patel & Joshi, 2017), prediction of river flow (Hussain et al., 2020), and simulation of water quality (Kadam et al., 2019). Various investigations were carried out from time to time by several researchers to reach the final efficient modeling approach that can not only save time but also provide reliable prediction results, some of those studies are described below;

(Büyükşahin & Ertekin, 2019) utilized a conventional modeling technique (ARIMA), and an artificial neural network modeling technique on a monthly stream flow time series. The result revealed that the ANN model performed better than ARIMA. (Ghorbani et al., 2016) carried out a comparative study, in which flow discharge was modeled by (ANN), (MLR) (RC), and (SVM), and their performances were compared with each other. The result revealed that values predicted by ANN, and SVM were more accurate and reliable than those by RC and MLR.

(Khazaee Poul et al., 2019) conducted a comparative study, in which monthly flow discharge was modeled and predicted by ANN, ANFIS, KNN, and MLR. The result of the study revealed that all three methods act significantly better than the MLR models. The MLR model tried to fit a linear relationship between the inputs and outputs. MLR was not able to estimate the monthly flows precisely enough. (Hussain & Khan, 2020) used three different machine learning modeling approaches; artificial neural network (ANN), support vector regression (SVR), and random forest (RF) to forecast Hunza River flow. The results found that ANN & RF performed slightly better than SVR. The main objective of the research study was to investigate the forecasting capabilities of the ANNs model for long-lead time on two-time scales.

2. Methodology

2.1. Study Area

Sukkur barrage is located at a longitude of 68.8471°, and a latitude of 27.6733° N, and was constructed on the River Indus in 1932 with the design discharge of 1.5 million cusecs for diverting water to canals at its upstream side with the purpose to ensure sustainable water provision to mankind for irrigating their lands, fulfilling anthropogenic and industrial water demands, flood control, and regulating the flow in the river.

The land that is cultivated by off-taking canals from the Sukkur barrage gives sufficient agricultural productivity to fulfill not only the need for rural and urban livelihoods in Sindh province but also to support the economy of our country. (Sindh barrage improvement project, 2018).

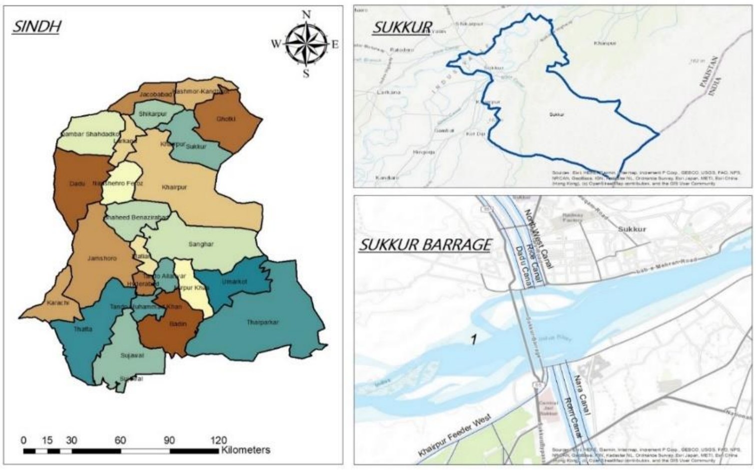

Figure 1. Shows the map of the study area of Sukkur Barrage

2.2. Steps of Model Development

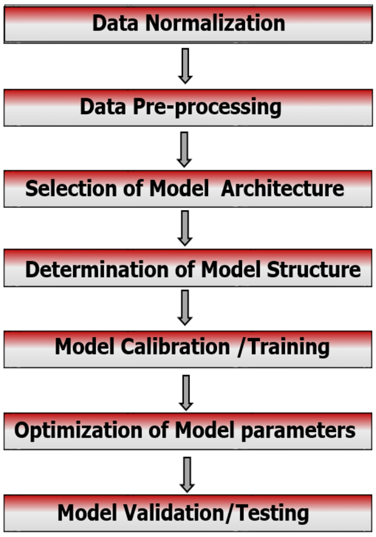

Figure 2. Shows the step-by-step procedure used for the model development.

2.2.1. Data Collection

In this research study, time series of flow discharge from Jan 2013 to December 2019 was acquired from the Department of Sindh Irrigation and Drainage Authority, Sukkur Barrage Division (SIDA).

2.2.2. Data Normalization

The input data were processed and normalized by the below equation to prevent the model from being dominated by variables with large values.

$$QS\ \frac{Qi - Qmin}{Qmax - Qmin} \qquad (1)$$

Where,
Qs = normalized value of data
Qi = observed value of data
Q min = minimum value of data
Q max = maximum value of data

2.2.3. Selection of Model Network

Among various well-known available artificial neural networks; two modeling approaches Multilayer Perceptron (MLP), and Time-Lagged Neural Networks (TLNN) were selected.

2.2.4. Determination of Model Structure

A model structure contains 11 input neurons, 1 hidden layer, and 24 numbers of hidden neurons, log sigmoid as a non-linear activation function was selected and the flow of data was chosen as straight/feedforward

$$af = \sum_{i=1}^{n} XiWi \qquad (2)$$

2.2.5. Data Splitting

The data was divided into two parts (70% and 30%); one for training, and the other part for testing/validating the developed model (Araghinejad, 2014).

2.2.6. Model Calibration

In the training phase initially, random weights were assigned to the data inputs by a non-linear activation function (eq. 3.) to adjust them time by time wherever needed to minimize the error between observed and predicted data.

$$f(af) = \frac{1}{(1+e^{-af})} \qquad (3)$$

2.2.7. Model Validation

After the development, and calibration of the model, validation was performed by using different statistical indices that include; Coefficient of Regression (R2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) to test the reliability and predictive performance of the developed model.

$$R2 = 1 - \frac{\sum_{i=1}^{n}(OBS.i - PRE.i)^2}{\sum_{i=1}^{n}(OBS.i - \overline{OBS})^2} \qquad (4)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}(Qp - Qo)^2 \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Qp - Qo)^2} \quad (6)$$

3. Results and Discussion

In this research study, Two Different types of ANN models; Multilayer Perceptron (MLP), and Time-Lagged Neural Network (TLNN) models have been selected and used for Flow Rate Prediction at daily and monthly timescale.

Table 1. shows the statistical Distribution of the Daily and Monthly datasets

Timescale	Total no. of Observations	Observations used for Training	Observations used for Testing	Standard Deviation (sd)	Max (ft3/sec)	Min (ft3/sec)	Mean (ft3/sec)
Daily	2513	2010	503	83208	709316	8980	75125
Monthly	696	557	139	121606	886524	7798	102460

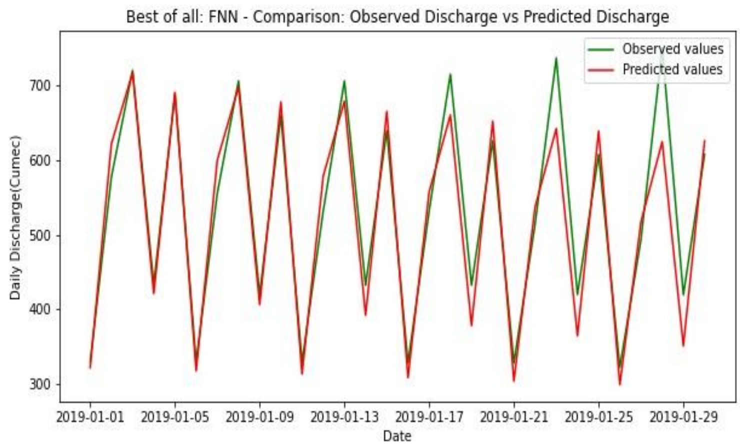

Figure 3. shows the observed and predicted values for January 2019, FNN model captured the daily variability well up to the 17th of January, then under-predicted the peak flows up to the end of the month.

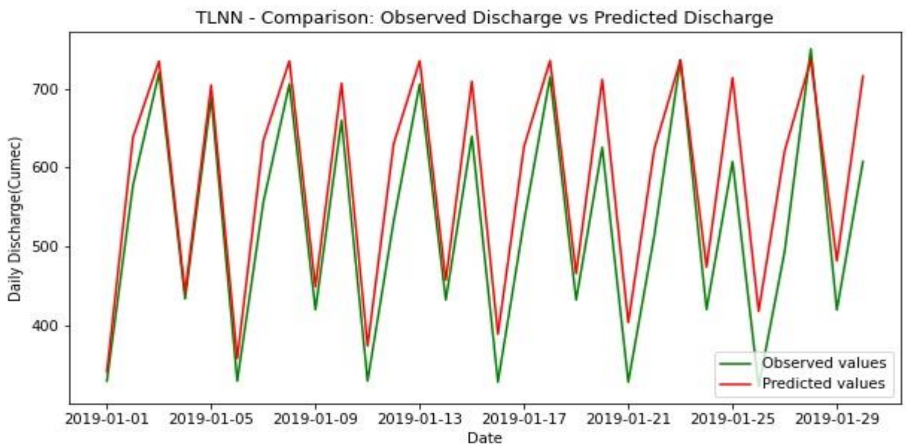

Figure 4. shows the observed, and predicted values by Time-Lagged Neural Network (TLNN), TLNN predicted the base as well as peak well up to the 23 days of the month.

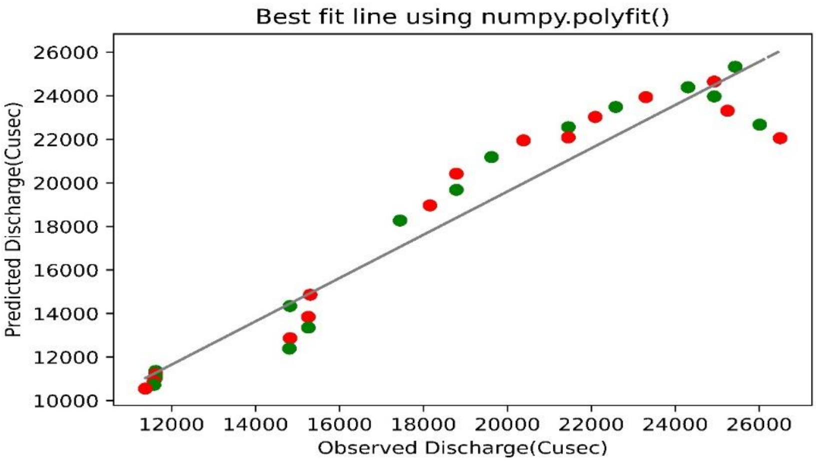

Figure 5. Scatterplot showing the linear relation between observ-ed and predicted discharge

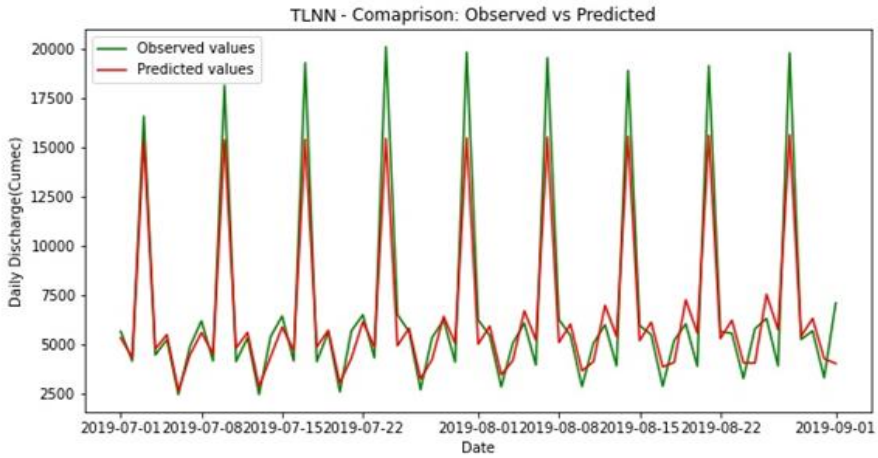

Figure 6. shows the observed and predicted values for the month of July-August 2019, TLNN model captured the base flow well up to July and later failed to capture base flow as well as peak flows

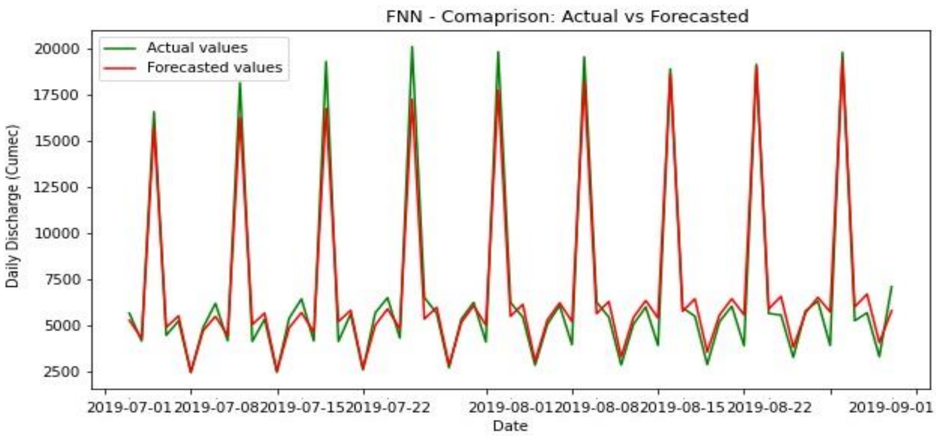

Figure 7. shows the Observed and Predicted values by FNN, which performed very well in capturing base and peak flows but slightly under-predicted the peak flows at the middle of the data points

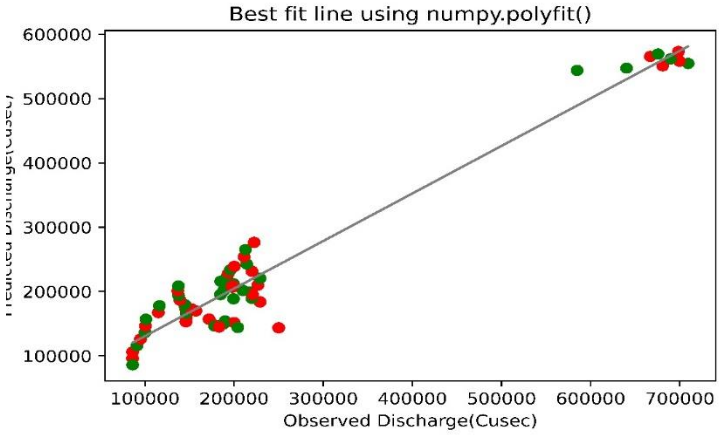

Figure 8. Scatterplot shows the relationship between observed and predicted discharge

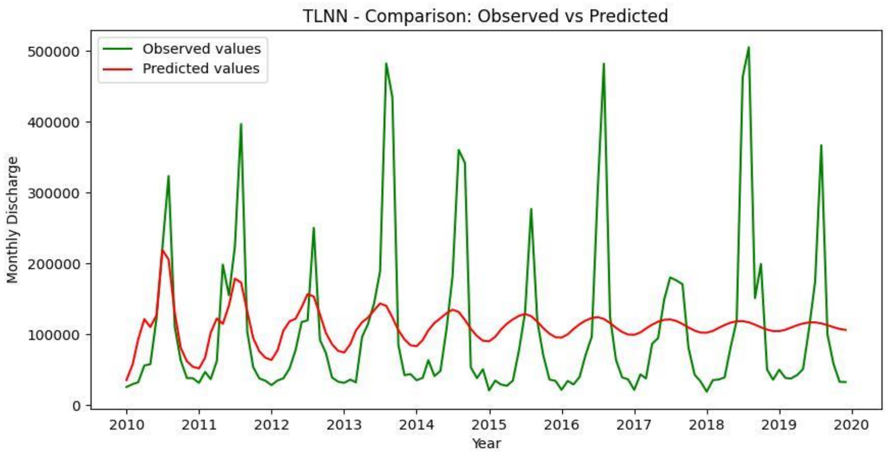

Figure 9. shows the Monthly Observed and Predicted values by TLNN from 2010-2020, which performed worst and failed to capture the base as well as peak flows

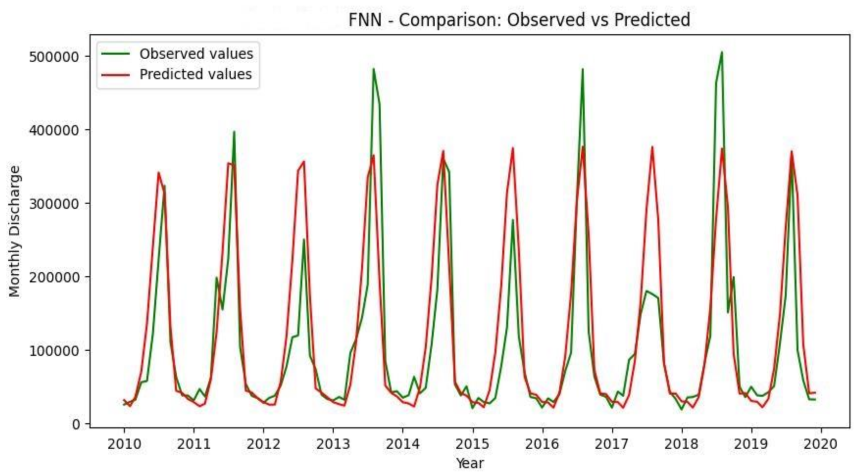

Figure 10. shows the Monthly Observed and Predicted values, FFNN captured the base flow very well at all points, it also predicted peak flows well at some points and slightly under-predicted the peaks in 2013, 2016, and 2018.

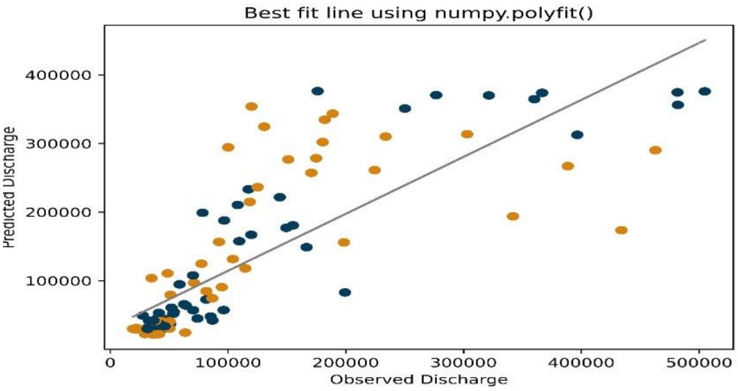

Figure 11. Scatterplot shows the distribution of observed and predicted discharge on a monthly timescale.

Table 2. shows the predictive performance of the developed models based on MAE, RMSE, and R2

Sr, No.	Timescale	Model	MAE (%)	RMSE (%)	R2 (%)
1	Daily (low Flows)	FNN	0.46	0.79	0.91
2	Daily (low Flows)	TLNN	2.01	2.63	0.79
3	Daily (high Flows)	FNN	1.74	2.31	0.96
4	Daily (high Flows)	TLNN	3.21	4.69	0.89
5	Monthly	FNN	17.3	28.29	0.76
6	Monthly	TLNN	200	289	0.33

4. Conclusion

Two ANN modeling approaches have been employed for the prediction of one step ahead at daily and monthly timescales. In the daily timescale among these 2 models, FFNN has performed better for low flows as well as for High flows, while TLNN has captured the daily variability less significantly than FFNN. In the monthly timescale, FFNN has captured base and peak flows very well, but TLNN completely failed to capture the monthly flow discharges, the reason for this is that in this case study, the number of observed monthly datasets used was much smaller than that of daily datasets. TLNN cannot be used where there is less availability of no. of observations for training and testing the model. In addition, it is noted that the number of epochs required for FFNN model convergence was much smaller than the TLNN model. This indicates that the FFNN model is more suitable and can extract nonlinearity characteristics of data more efficiently in comparison to the TLNN model.

References

Adedigba, S. A., Khan, F., & Yang, M. (2017). Dynamic failure analysis of process systems using neural networks. Process Safety and Environmental Protection, 111, 529-543.

Bisoyi, N., Gupta, H., Padhy, N. P., & Chakrapani, G. J. (2019). Prediction of daily sediment discharge using a back propagation neural network training algorithm: A case study of the Narmada River, India. International journal of sediment research, 34(2), 125-135.

Büyükşahin, Ü. Ç., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing, 361, 151-163.

Buyukyildiz, M., & Kumcu, S. Y. (2017). An estimation of the suspended sediment load using adaptive network based fuzzy inference system, support vector machine and artificial neural network models. Water Resources Management, 31(4), 1343-1359.

Cheng, M., Fang, F., Kinouchi, T., Navon, I., & Pain, C. (2020). Long lead-time daily and monthly streamflow forecasting using machine learning methods. Journal of Hydrology, 590, 125376.

Ghorbani, M. A., Khatibi, R., Goel, A., FazeliFard, M. H., & Azani, A. (2016). Modeling river discharge time series using support vector machine and artificial neural networks. Environmental Earth Sciences, 75(8), 1-13.

Hussain, D., Hussain, T., Khan, A. A., Naqvi, S. A. A., & Jamil, A. (2020). A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin. Earth Science Informatics, 13(3), 915-927.

Hussain, D., & Khan, A. A. (2020). Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. Earth Science Informatics, 13(3), 939-949.

Kadam, A., Wagh, V., Muley, A., Umrikar, B., & Sankhua, R. (2019). Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. Modeling Earth Systems and Environment, 5(3), 951-962.

Khazaee Poul, A., Shourian, M., & Ebrahimi, H. (2019). A comparative study of MLR, KNN, ANN and ANFIS models with wavelet transform in monthly stream flow prediction. Water Resources Management, 33(8), 2907-2923.

Masselot, P., Dabo-Niang, S., Chebana, F., & Ouarda, T. B. (2016). Streamflow forecasting using functional regression. Journal of Hydrology, 538, 754-766.

Melesse, A., Ahmad, S., McClain, M., Wang, X., & Lim, Y. (2011). Suspended sediment load prediction of river systems: An artificial neural network approach. Agricultural Water Management, 98(5), 855-866.

Noori, N., & Kalin, L. (2016). Coupling SWAT and ANN models for enhanced daily streamflow prediction. Journal of Hydrology, 533, 141-151.

Parisouj, P., Mohebzadeh, H., & Lee, T. (2020). Employing machine learning algorithms for streamflow prediction: a case study of four river basins with different climatic zones in the United States. Water Resources Management, 34(13), 4113-4131.

Patel, A. B., & Joshi, G. S. (2017). Modeling of rainfall-runoff correlations using artificial neural network-A case study of Dharoi Watershed of a Sabarmati river basin, India. Civil Engineering Journal, 3(2), 78-87.

Taormina, R., Chau, K.-w., & Sethi, R. (2012). Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. Engineering Applications of Artificial Intelligence, 25(8), 1670-1676.

Wu, H., & Prasad, S. (2017). Convolutional recurrent neural networks for hyperspectral data classification. Remote Sensing, 9(3), 298.

Yaseen, Z. M., El-Shafie, A., Jaafar, O., Afan, H. A., & Sayl, K. N. (2015). Artificial intelligence based models for stream-flow forecasting: 2000–2015. Journal of Hydrology, 530, 829-844.