International Journal of Multidisciplinary Research and Growth Evaluation.

Flow prediction in Kabul River: An artificial intelligence based technique

Salah Ud Din
University of Engineering and Technology, Peshawar, Pakistan

* Corresponding Author: **Salah Ud Din**

Article Info

ISSN (online): 2582-7138
Impact Factor: 5.307 (SJIF)
Volume: 05
Issue: 02
March-April 2024
Received: 15-02-2024;
Accepted: 19-03-2024
Page No: 854-857

Abstract
This study employs random forest regressor to forecast future flow in Kabul River. Utilizing CMIP6 projected climate data for the SSP585 scenario from the IPSL-CM6A-LR climate model. The random forest regressor demonstrate efficacy in predicting flow, achieving an R^2 of 0.77. The study highlights the importance of modern artificial intelligence-based techniques for precise flow and flood predictions and suggests an increase in flash flood events in Kabul River in response to a warming climate.

DOI: https://doi.org/10.54660/.IJMRGE.2024.5.2.854-857

Keywords: Random Forest, CMIP6, Kabul River, Floods, Climate Change, Artificial Intelligence

1. Introduction

In 2010 Kabul river has experienced devastating flash flood that caused severe human and financial losses [1]. Kabul river being a key tributary of the Indus river with its basin located at the verge of monsoon region is exposed to flash floods, especially during monsoon periods [2].

The recent advances in climate changes revealed that the global as well as the regional temperatures at Kabul basin would increase by around 2 to 5 degrees Celsius by 2100 [3, 4]. With this increase in global and regional temperatures water cycle including flow patterns in rivers and streams would significantly change [2]. Such changes would be nonuniform, with increase and decrease of hydrological events in various regions are expected [3, 4]. However, studies on Kabul river have shown that flow pattern would significantly change to yield catastrophic floods more frequently due to glacial and snow melt [2, 5].

Conventionally, hydrological predictions are made using statistical methods that are oriented on process (water cycle). Given its estimative nature and recent advances in predictions using artificial intelligence these hydrological predictions can be easily made using the artificial intelligence techniques [6].

In this study an artificial intelligence based algorithm named random forest is employed to predict future flow of Kabul river.

2. Literature Review

With warming climate global as well as regional water cycle is impacted, affecting uses dependent on the hydrological regimes mainly, hydropower energy, groundwater, agricultural production and ultimately the availability of water for domestic purposes [7, 8, 9], bringing frequent floods [5] that are highly dependent on basin climate parameters and can lead to damages of bridges and other hydraulic structures [10].

Studies have shown that in the Indus River due to climate warming, the summer flows have already increased since 1986 [11]. Kabul river as a key tributary to Indus River is also impacted due to climate change, with its flow patterns changed to yield more floods events [2].

The hydrological science and most of the problems that are worked out in it are based on statistical estimations, for which modern artificial intelligence based regression algorithm provide as ease [6].

The use of such techniques has significantly increased in recent advances, in which the random forest based model have shown better performance along with Support Vector Machine (SVM) and Multilayer Perceptron Neural Network (MLP) [12]. Similarly, Bayesian Linear Regression (BLR), Boosted Decision Tree Regression (BDTR), and Neural Network Regression (NNR) have also been used to predict rainfall events [13, 14].

Random forest is a technique based on multiple decision trees that are trained to give predictions. Each tree can give a prediction and average of all giving final prediction result of random forest regressor [15].

3. Methodology

3.1. Study Area

Kabul river basin is the border region of Pakistan and Afghanistan, with catchment 92,650 sq. km. Kabul River is mainly fed by snow and glacial melt with monsoon precipitation yielding floods during monsoon period [2].

Fig 1: Kabul River Basin

3.2. Material and Methods

Training data comprise the daily flow data of the Kabul river from Jan 2002 to Nov 2022. The flow data contained missing values which are marked NAN.

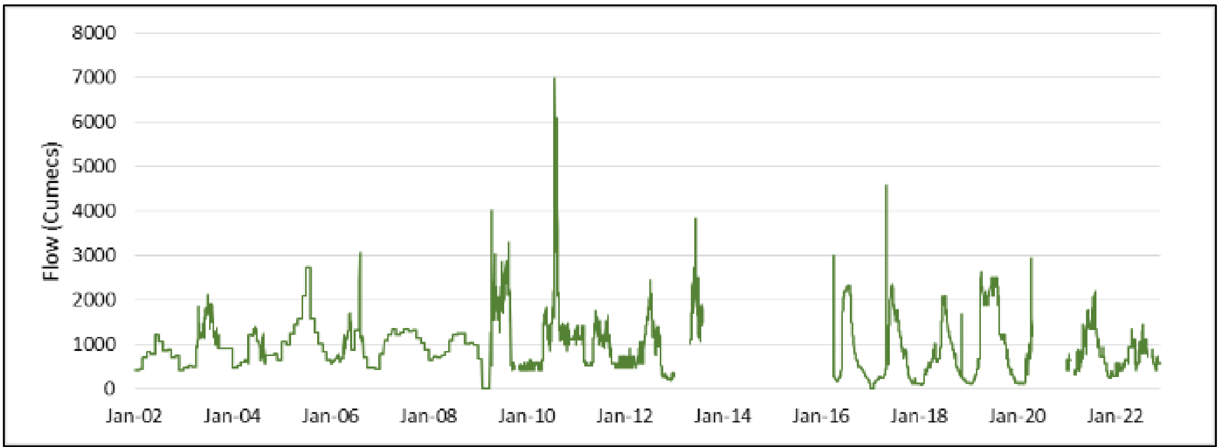

Fig 2: Daily flow data of Kabul River from 01.01.2002 to 31.11.2022, collected at Nowshera gauge station.

Climate data from the climate model IPSL-CM6A-LR (France) is another part of the training dataset. Precipitation and temperature data from IPSL-CM6A-LR for SSP585 scenario of CMIP6, is used as climate parameter as both play a key role in determining the flow pattern in the rivers [16]. The precipitation and temperature data from IPSL-CM6A-LR is already bias-corrected downscaled [17].

Daily precipitation (Ppt) along with daily max temperature (Tmax) and daily minimum temperature (Tmin) are the three parameters that formed the independent variable set of the training dataset.

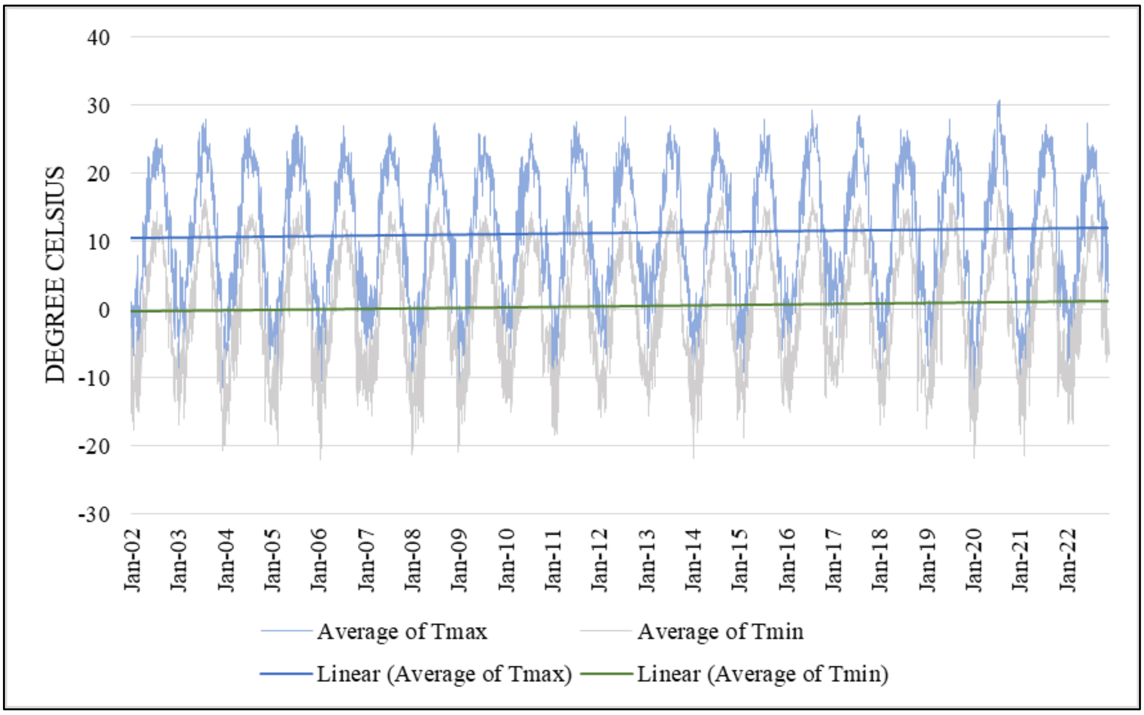

Fig 3: Variation of basin average daily max temperature and daily min temperature from 01.01.2002 to 31.11.2022 for Kabul River

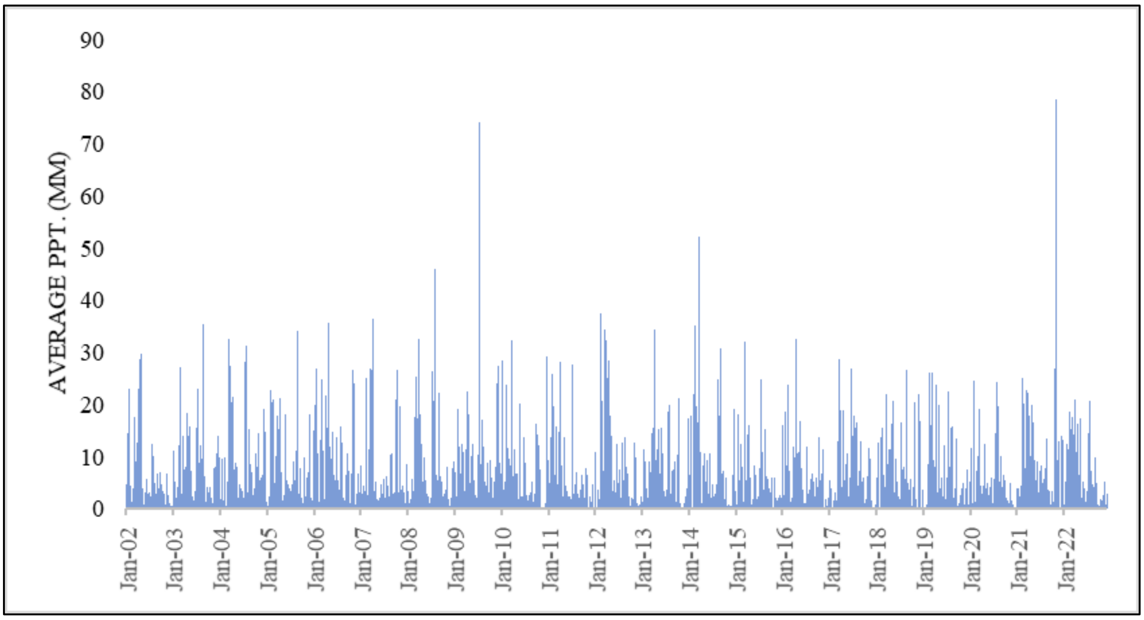

Fig 4: Variation of basin average daily precipitation from 01.01.2002 to 31.11.2022 for Kabul River

Consequently, two datasets are prepared the first one is the training dataset containing independent variable set (Ppt, Tmax and Tmin) and dependent variable as daily flow of Kabul River from Jan 2002 to Nov 2022, the second dataset is the prediction dataset containing values of Ppt, Tmax and Tmin for future period; from 2031 to 2050. The training dataset is further split in 80:20 splits as calibration and validation samples.

The random forest regressor form sklearn documentation is then used to build the model that is trained using the calibration sample. Various values of number of trees and tree depths are randomly checked for calibration and finalization of the architecture of the random forest regressor for yielding the best result.

The accuracy of random forest regressor for both calibration and validation sets are checked used R^2. Finally, using the prediction dataset flow values for the future period is predicted.

4. Results & Discussions

The random forest regressor give better performance for flow predication using 100 numbers of trees with max depth value of each tree as 20. R^2 calibration value of 0.94 and R^2 validation value of 0.77 indicated that the random forest regressor satisfactorily detect pattern in calibration data and its performance in validation data is reasonably good.

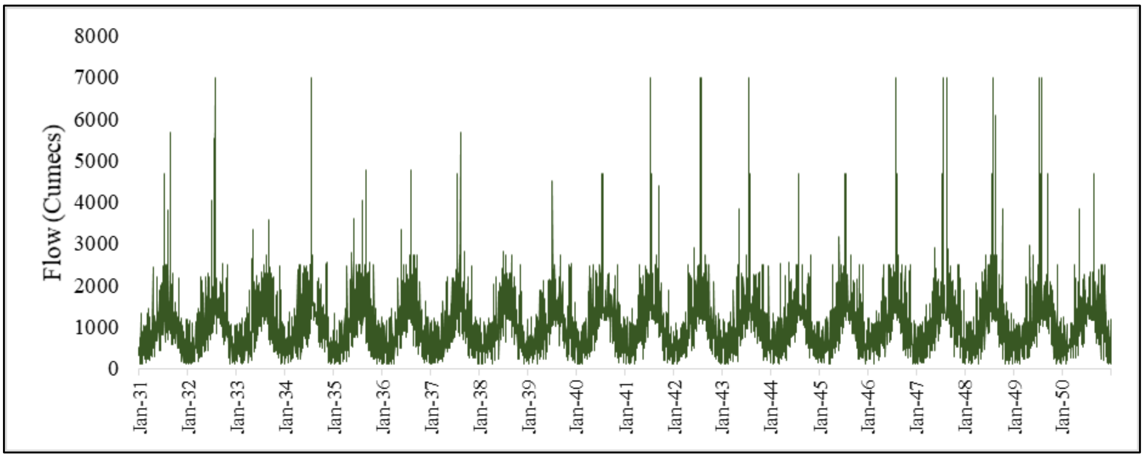

Fig 5: Result of flow prediction for Kabul River from random forest regression model

The temperature data IPSL-CM6A-LR revealed (Figure 2) that basin average daily temperature is following an increasing trend with daily maximum increased by around 3 % from 2002 to 2022. The results of predicted flow (Figure 5) from the random forest regression model for future indicated that in future the flow pattern of Kabul River would significantly change and flood would occur more frequently.

5. Conclusions

The performance of the random forest regressor is found reasonably good. The flood pattern for future period is effectively detected, depicting that flood pattern in Kabul River would significantly vary in future and more occurrence of flood events are expected. However, the performance of random forest regressor for climate models other than IPSL-CM6A-LR and for other CMIP6 scenario i:e SSP245 etc, need to be assessed.

6. References

1. Sayama T, Ozawa G, Kawakami T, Nabesaka S, Fukami K. Rainfall-runoff inundation analysis of the 2010 Pakistan flood in the Kabul River basin. Hydrol Sci J. 2012 Feb; 57(2):298-312. doi: 10.1080/02626667.2011.644245.
2. Iqbal MS, Dahri ZH, Querner EP, Khan A, Hofstra N. Impact of climate change on flood frequency and intensity in the Kabul river basin. Geosci. 2018 Apr; 8(4). doi: 10.3390/geosciences8040114.
3. Jarraud M, Steiner A. Summary for policymakers. 2012. doi: 10.1017/CBO9781139177245.003.
4. Achuta Rao K, Adhikary B, Allan RP, Armour K, Bala G, Barimalala R, Berger S. Technical Summary IPCC Sixth Assessment Report by Working Group 1. The Intergovernmental Panel on Climate Change. 2019 [Online]. Available: https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/
5. Hirabayashi Y, Alifu H, Yamazaki D, Imada Y, Shiogama H, Kimura Y. Anthropogenic climate change has changed frequency of past flood during 2010-2013. Prog Earth Planet Sci. 2021 Dec; 8(1). doi: 10.1186/s40645-021-00431-w.
6. Lange H, Sippel S. Machine Learning Applications in Hydrology. 2020. doi: 10.1007/978-3-030-26086-6_10.
7. Penmetsa V, Holbert KE. Climate Change Effects on Solar, Wind and Hydro Power Generation. IEEE. 2019.
8. Touseef M, *et al*. Assessment of the future climate change projections on Streamflow Hydrology and water availability over upper Xijiang River Basin, China. Appl Sci. 2020 Jun; 10(11). doi: 10.3390/app10113671.
9. Ravazzani G, Barbero S, Salandin A, Senatore A, Mancini M. An integrated Hydrological Model for Assessing Climate Change Impacts on Water Resources of the Upper Po River Basin [Online]. Available: http://dx.doi.org/10.1007%25
10. Allahbakhshian-Farsani P, Vafakhah M, Khosravi-Farsani H, Hertig E. Regional Flood Frequency Analysis Through Some Machine Learning Models in Semi-arid Regions. Water Resour Manag. 2020 Jul; 34(9):2887–2909. doi: 10.1007/s11269-020-02589-2.
11. Arfan M, Lund J, Hassan D, Saleem M, Ahmad A. Assessment of spatial and temporal flow variability of the Indus River. Resources. 2019 Jun; 8(2). doi: 10.3390/resources8020103.
12. Hagen JS, Leblois E, Lawrence D, Solomatine D, Sorteberg A. Identifying major drivers of daily streamflow from large-scale atmospheric circulation with machine learning. J Hydrol. 2021 May; 596. doi: 10.1016/j.jhydrol.2021.126086.
13. Ridwan WM, Sapitang M, Aziz A, Kushiar KF, Ahmed AN, El-Shafie A. Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia. Ain Shams Eng J. 2021 Jun; 12(2):1651–1663. doi: 10.1016/j.asej.2020.09.011.
14. Lee S, Kim JC, Jung HS, Lee MJ, Lee S. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Nat Hazards Risk. 2017 Dec; 8(2):1185–1203. doi: 10.1080/19475705.2017.1308971.
15. BREIMAN L. Random Forests. Mach Learn. 2001; 45:5–32. [Online]. Available: https://link.springer.com/content/pdf/10.1023/A:1010933404324.pdf
16. Rajbhandari R, Shrestha AB, Kulkarni A, Patwardhan SK, Bajracharya SR. Projected changes in climate over the Indus river basin using a high resolution regional climate model (PRECIS). Clim Dyn. 2015; 44(1–2):339–357. doi: 10.1007/s00382-014-2183-8.
17. Khan A. Climate Change Assessment Report For Pakistan. 2021 Apr.