Article

Machine Learning–Based Reconstruction of Missing Streamflow Records in Data-Scarce River Basins

Muhammad Faarid Shah[1]*, Muhammad Farrukh[1], Fahad Shamraiz[1], Touqeer Ali Rind[1], Farjad Aziz[1], Asad Wahab[1] and Qazi Khurshid Ahmed[1]

1 Department of Civil Engineering (DCvE), Ghulam Ishaq Khan (GIK) Institute of Engineering Sciences and Technology, Topi, Khyber Pakhtunkhwa; muhammadfarrukhkhajjak@gmail.com, qureshifahad9999@gmail.com, touqeer.ali@giki.edu.pk, gcv2568@giki.edu.pk, gcv2520@giki.edu.pk, gcv2445@giki.edu.pk

* Correspondence: gcv2560@giki.edu.pk

Abstract

Stable streamflow records are essential in water-resources planning, flood prediction, and effecting climatic change; when plentiful, however, in high-altitude and trans-boundary river basins because of rugged terrain, severe climatic conditions, and insufficient monitoring facilities. The paper fills in missing streamflow records in the Chitral River Basin, a snow- and glacier-fed basin of the Upper Indus Basin, by reconstructing missing discharge at the downstream Arandu station with paradoxical observations of the Chitral station. The statistical and machine-learn models that were used to analyze daily discharge data between 1981-2024 are Linear Regression, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XG Boost). The coefficient of determination (R 2), root mean square error, mean squared error, and mean absolute error were used to assess how well a model works. Although the predictive power of all models was quite robust with R^2 exceeding 0. 95, ANN showed the most balanced and stable performance at both training and testing stages with the lowest prediction errors and best generalization. On these findings, the ANN model was used to recreate missing discharge at the Arandu station to generate a long-run series of streamflow between 1981 and 2024. The restored hydrograph maintains the seasonal snowmelt-based flow regime and fits well with observed records to actually serve as a virtual discharge sensor. The suggested framework is a feasible approach to hydrological reconstruction in mountainous basins with limited data, as well as the enhanced water-resources management in the context of variable climatic conditions.

Keywords: Streamflow reconstruction; Machine learning; Chitral River Basin.

1. Introduction

Access to sound hydrology data, especially streamflow data are one of the persistent issues in most river basins across the globe. The data gaps usually occur because of the extreme climatic conditions, rugged terrain, inadequate monitoring facilities and operational or political limitations, particularly where there are mountains and cross-border areas. The missing streamflow records pose a serious constraint to this water resources planning, flood forecasting and climate change impacts assessment. To overcome these limitations, machine learning (ML) methods have been actively embraced as data-driven approaches towards predicting and reconstructing missing hydrological information, being more flexible and accurate than traditional statistical or empirical procedures.

Recent studies show that ML models have a high potential to be used in imputing streamflow data in a wide variety of hydroclimatic contexts. Modern neural-network-based methods, like the Probabilistic Fusion Imputer with Neural Networks (PROFINN), have been incredibly successful in terms of reconstruction accuracy, averaging RMSE of 0.91 and NashSutcliffe Efficiency (NSE) of 0.93 in the Pamba River basin successfully validating various flow regimes and data-gap situations [1]. The ensemble learning techniques have also demonstrated solid performances. RF with clustering methods has resulted in a mean NSE of more than 0.85 in Indian catchments [2]. On the same note, the predictive capability of Gradient Boosting and Bagging Regressors has been reported to be very high in the estimation of streamflow and the reported accuracy rates of the two methods lie in the range of 0.9737 to 0.9968 [3]. With limited data, simpler ML models like Naive Bayes and k-Nearest Neighbors (KNN) have performed better than traditional methods, and the former has been significantly useful when training data are limited [4].

Although such improvements are made, the performance of ML-based streamflow prediction is made very dependent on the availability of data, climatic variability, and hydrological processes that are unique to the basin. Streamflow gaps can span several months in areas where there is high hydroclimatic variability, which makes models more challenging to train and predict. As a result, ML models should be properly adjusted to local conditions, to guarantee the assurance of successful performance [2]. Recent research also shows that combining ML techniques and physical hydro knowledge enhances the reliability of predictions in ungauged and under-gauged basins, especially where the terrain is complex [5].

The river basins with high altitudes are important in the freshwater cycle of the planet since snow and ice stored in winter are melted during the spring and summer seasons. A classic example is the Hindu Kush-Karakoram-Himalayan (HKH) area that is central and plays a major water tower to South and Central Asia and is a provider of river systems that supply almost 800 million people [6]. The temperature changes are very sensitive to hydrological processes in this region, and directly determine the dynamics of the snow melting and glaciers. Even minor shifts in temperature linked to climate change can have a profound effect of the magnitude of streamflow and the distribution of how it changes throughout the year.

The Upper Indus Basin (UIB) has streamflow regimes that are significantly controlled by seasonal snowmelt and glacier melt. According to historical studies, there are trends towards surface and subsurface hydrology, at a rate of about 0.03-0.05 times per minute per second, which mainly is caused by the intensified melting of glaciers, changes in precipitation levels, and variations in seasonal climate cycles [7]. The trends are extremely dangerous to future water supply and overall, to the possibility of extreme hydrological processes [8].

The UIB and hydrological monitoring of tributaries are still inadequate as topography is rugged, climatic conditions extreme, and several different river systems are transboundary. This has led to the fact that streamflow records are usually incomplete or not available, especially in the upstream and border regions. This has been linked to the well-known ungauged or under gauged catchment problem that poses significant problems to hydrological modeling, flood forecasting and water resources management [8]. Conventional gap-filling methods, including interpolation or transposition of flows within neighboring stations are frequently not suitable to snow- and glacier-fed rivers, due to their poor ability to model complicated flow behavior. This has led to the emergence of regionalization methods that make use of hydrological similarity and upstream-downstream linkages in the transfer of information to poorly measured sites and those better measured, which improves the reliability of data in the mountainous basin [9].

2. Materials and Methods

2.1 Study Area and Hydrological setting

Chitral River Basin is a large mountainous branch of the Kunar-Kabul River system and an important part of the Kabul River Basin (KRB). Although this basin is strategically hydrologically significant to the downstream water supply, it has received relatively little coverage in scientific literature specifically in long-term flow reconstructions and climate sensitivity.

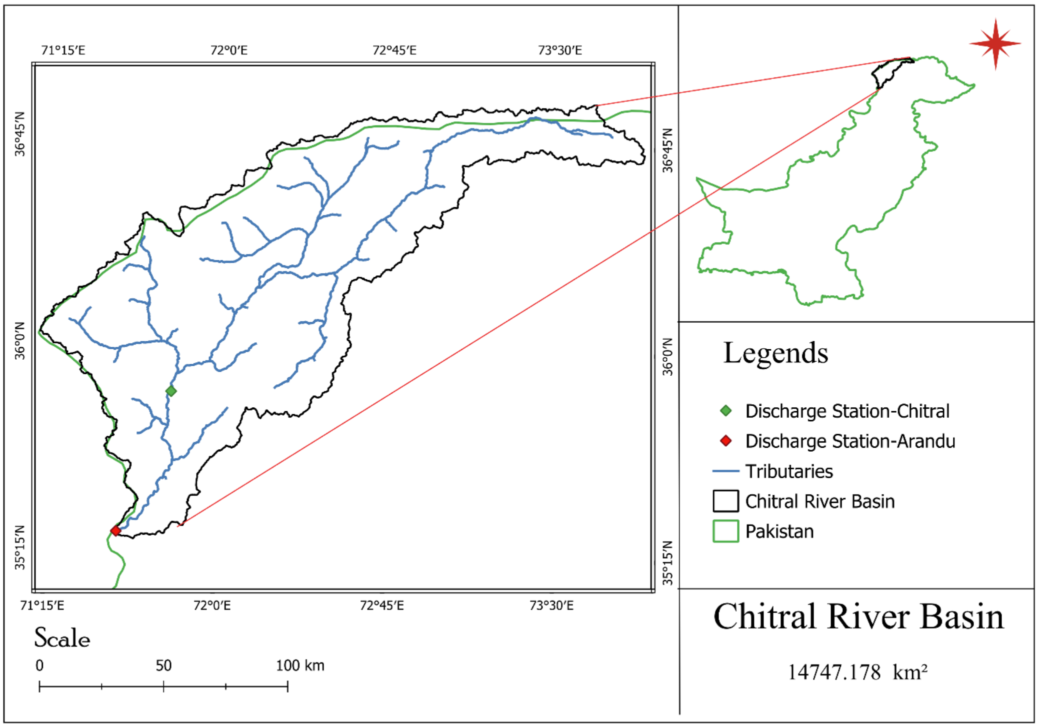

Figure 1: Chitral River Basin

The Chitral River is the one that originates on the mountain range of Hindu Kush and its catchment area is about 14,747km^2 (Figure 1). The basin is largely snow and glacier-driven and the streamflow is mostly controlled by seasonal snowmelt and, to a lesser degree, spring and summer precipitation. The outcome of these cryosphere controls is very strong non-linear hydrological behavior and intra-annual variability.

Monitoring in the basin is very minimal. The downstream Arandu gauging station located at the border area of Pakistan and Afghanistan (35°19′0″ N, 71°34′0″ E), installed in 2008 by WAPDA, is characterized by sporadic records that are short in duration thereby limiting sound interpretation of the long-term flow variability and water availability in the future. Conversely, the upstream Chitral station (35°51'48.0"N, 71°47'15.0"E) installed in 1963 by WAPDA possesses a rather long and continuous discharge history tracing back to the early sixties. This difference forms a significant problem to hydrological evaluation and infrastructural planning in the whole basin.

2.2 Machine Learning Applications

Recent developments in machine learning (ML) and hydro-informatics offer potentially useful tools to overcome such data scarcity. ANNs especially have proven to be highly capable of modeling streamflow in data-constrained mountainous catchments because they can model even complex and non-linear relationships without necessarily being parameterized. Past research has indicated that under those circumstances, the ML-based methods often work better than the conventional autoregressive and conceptual models [10].

Although there is an increase in the use of ML to reconstruct medium-term streamflow even at smaller high-altitude sub-basins within the Hindu Kush-Himalaya (HKH) and UIB, like the Chitral-Arandu system, the long-term reconstruction of streamflow is limited. Majority of the current studies are performed on the large rivers, as small tributaries are underrepresented [7]. Such incomplete long-term discharge information deteriorates water resources planning, design of hydraulic infrastructure and evaluation of climatic-change effects in these sensitive areas [11].

To fill this gap, the current paper uses almost 43 years of hydrological data (1981-2024) to recreate streamflow in the Arandu station using both machine-learned and statistical methods. Linear Regression (LR) and models such as ANN, Support Vector Regression (SVR), random forest (RF), Decision Trees (DT) and Extreme Gradient Boosting (XG Boost) are used.

2.3 Data Collection and Study Period

Two hydrometric stations on the basin were used to acquire data on daily discharge, namely Chitral and Arandu located at (35°19′0″ N, 71°34′0″ E) and (35°51'48.0"N, 71°47'15.0"E). The Chitral station has a continuous discharge record since 1981 to 2024, and the Arandu station has continuous discharge record since 2008 to 2024, although some years are missing, 2013-2016. At the two stations, discharge is reported through cubic meters/s (m3/s).

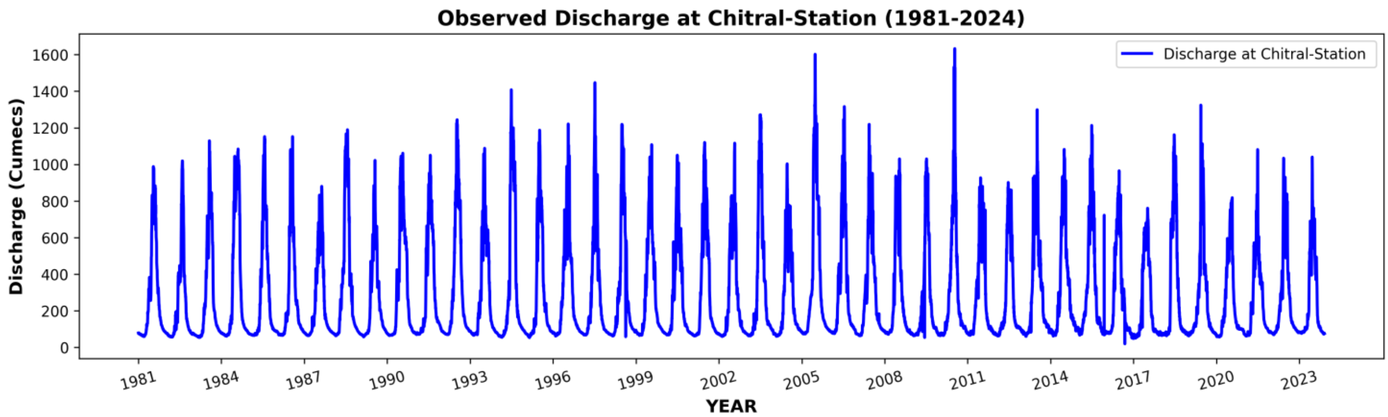

Figure 2: Observed discharge at Chitral Station (1980-2024)

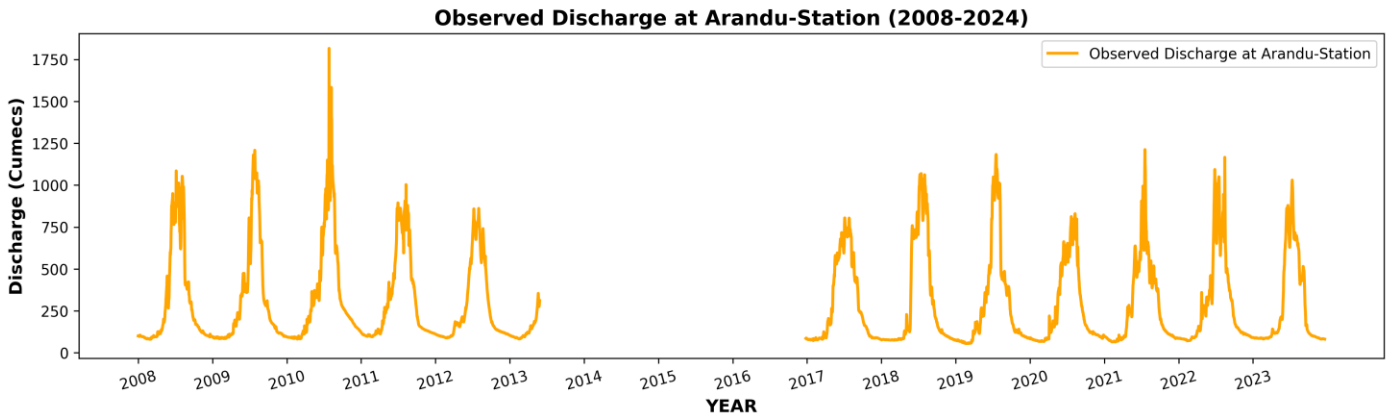

Figure 3: Observed Discharge at Arandu Station (2008-2024)

The Chitral station was chosen as the main predictor since it had a long and consistent history and its hydrological interest to the downstream Arandu station. The two stations share similar seasons flow regimes, due to melting snow and glaciers, which implies a high level of hydrological connectivity between the upstream and the downstream areas as reported in earlier research [2]. This connectivity provides the reason why upstream discharge is an explanatory factor in the ways of reconstituting downstream flows.

2.4 ML Models

.In this study, three machine-learning models of varying complexity were used to forecast the missing streamflow data at the downstream station. Linear Regression (LR), Artificial Neural Networks (ANN) and Extreme Gradient Boosting (XGBoost) were chosen to serve as a representation of the linear, non-linear and the ensemble-based methods of learning, respectively. This is because the application of a variety of models enables them to assess their comparative performance in data-limited and highly varying hydrological conditions.

2.4.1 Linear Regression (LR)

LR was used in order to develop a direct linear correlation between upstream and downstream discharge as a baseline model. Due to its simplicity and interpretability, LR offers a reference framework in which the added value of machine-learning approaches could be evaluated. Nevertheless, it is not applicable in cases when hydrological processes are highly un-linear and this is typical of snow and glacier-fed river systems [12].

2.4.2 Artificial Neural Networks (ANN)

ANNs were used to learn non-linear input/output relationships that do not explicitly parameters like physical ones. The feed-forward type of network was adopted because ANNs are highly applicable to the modeling of complex hydrological systems that are affected by the cryosphere processes. They have been found to be highly predictive in streamflow modeling and especially in low-data situations [12],[13].

2.4.3 Extreme Gradient Boosting (XG Boost)

XG Boost was used as an ensemble method of learning that constructs sequential decision trees with the aim of reducing errors in prediction via gradient boosting. XG Boost is computationally efficient and it is able to deal with non-linear relationships, heterogeneous inputs and missing values. The latter features are what make it applicable to model mountain streamflow that is highly variable and prone to adverse effects of noise and limited observations on the simpler models.

3.5 Model Evaluation

There were four statistical indicators that were used to evaluate the model performance: the coefficient of determination (R^2), root mean square error (RMSE), mean squared error (MSE) and mean absolute error (MAE). These measures are regularly used in hydrological modeling to measure the correspondence between observed and modeled streamflow as well as to gauge the predictive capacity within distinct flow circumstances.

With coefficient of determination (R^2), the proportion of variance in the observed discharge that is explained by the model predictions is measured. It gives a measure of goodness of fit of a model and capability to reproduce temporal variability in streamflow but fails to give an actual measure of the magnitude of prediction error[14].

Root mean square error (RMSE) is a square root of the mean of squared errors of the observed and predicted values. RMSE is also keen on huge deviations, and thus it puts more emphasis on the errors in peak-flow, hence it is especially applicable in examining the performance of the model in high-flow [15].

The mean squared error (MSE) is the average of squared differences in the observed and simulated discharge. It gives a direct account of total prediction error, and is typical as an objective term when calibrating a model, but is highly susceptible to outliers as it squares errors [16].

Mean absolute error (MAE) is used to compute the average value of absolute differences between predicted and observed values. MAE is also less susceptible to extreme values than RMSE and MSE and provides a powerful estimate of typical error of predictions in highly varying hydrological time series [14].

3. Results

3.1. Model Performance in Training

Linear regression (LR), Artificial Neural Networks (ANN) and XG Boost were initially tested on the training set. Table 1 is a summary of training performance measures, whereas Figure 4 shows the correlation between the inflow (Chitral) and the predicted outflow (Arandu) of all three models.

Table 1: Training Data Evaluation Metrices

Model	R^2	RMSE	MSE	MAE
Linear Regression	0.95	63.6	4045.3	36.16
ANN	0.95	63.91	4084.56	34.56
XG Boost	0.96	55.06	3031.93	31.56

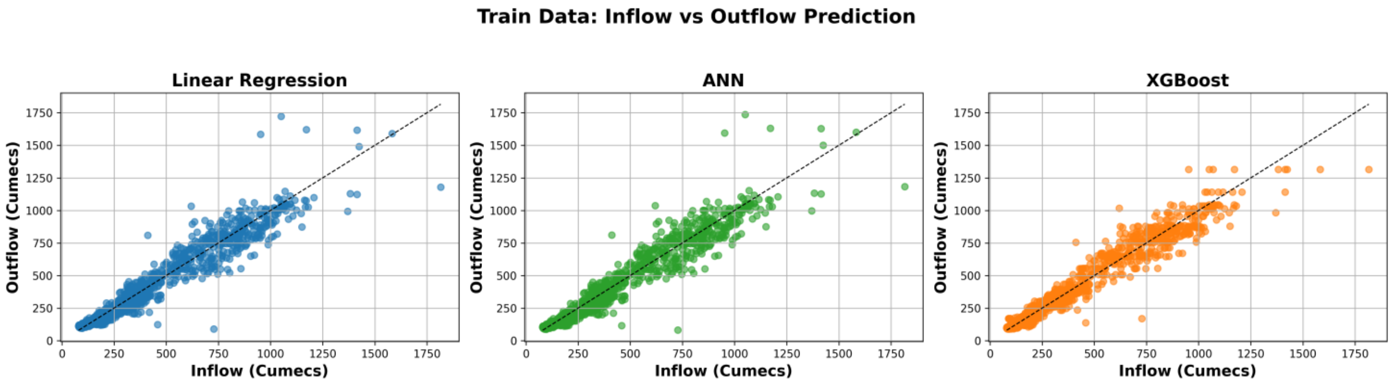

Figure 4: Training Data's Scattered Plots

All models revealed a high level of agreement between observed and simulated discharge during the training with coefficients of determination (R^2) more than 0.95. Linear Regression was able to model downstream discharge variability with an R^2 of 0.95 and thus a large percentage of the variability could be attributed to a linear relationship between upstream inflow. Nevertheless, the values of RMSE (63.6 m3/s) and MAE (36.16 m3/s) are relatively high, indicating lower accuracy in a high-flow regime.

The ANN model had a similar explanatory power (R^2 = 0.95) but the error properties, especially MAE (34.56 m3/s), were better, which represents the average discharge conditions better. The ANN scenario plot shows a higher concentration around the 1:1 line than that of LR particularly at medium-to-high flows.

XG Boost performed better in the training phase as it gave a high R^2 (0.96) and a low RMSE (55.06 m3/s), MSE (3031.93) and MAE (31.56 m3/s). The training scatter plot shows that it has good alignment with observed discharge throughout the entire flow range implying that it has a good fitting capability.

3.2 Model Performance in Testing

The generalization of the model was checked on the independent testing dataset (Table 2, Figure 5). The three models were all very predictive, which proved the strength of the relationship between upstream and downstream discharges.

Table 2: Testing Data Evaluation Metrices

Model	R^2	RMSE	MSE	MAE
Linear Regression	0.97	47.24	2231.4	34.75
ANN	0.97	45.48	2068.74	31.24
XGBoost	0.96	55.13	3039.29	38.7

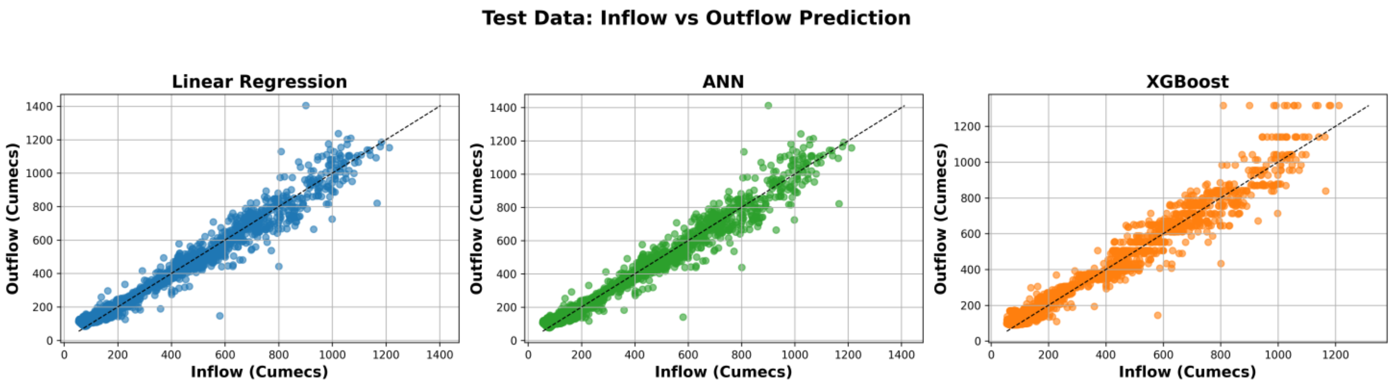

Figure 5: Testing Data's Scattered Plots

Linear Regression resulted in R^2 of 0.97, RMSE and MAE of 47.24 m3/s and 34.75 m3/s respectively. Although the high R^2 is an indicator of good overall agreement, this dispersion was more at higher flows implying that the 1:1 line could not capture non-linear variability effectively.

ANN model proved to be the most balanced in terms of performance in the course of testing as it was the most accurate among all models. The lowest RMSE (45.48 m3/s), MSE (2068.74) and MAE (31.24 m3/s), were observed. Testing scatter plot indicates less bias and narrows the clustering around the 1:1 line than LR and XGBoost especially under peak flow conditions.

To the contrary, XGBoost showed a deterioration in performance when tested. Even though the R^2 was still large (0.96), the measures of errors have worsened significantly (RMSE = 55.13 m3/s; MAE = 38.7 m3/s). The higher the scatter and deviation with increased discharges, the more evidence that there was some overfitting in the training and that it would no longer be able to generalize.

3.3 Choosing the most effective model.

Despite the better results of XG Boost in training (Table 1), its lower accuracy in testing implies that it cannot be used for independent prediction. The assumption that Linear Regression is stable is true but it indicated a much higher error (Table 2)and was not able to represent non-linear hydrological behavior.

The ANN model proved that the most consistent and reliable performance was observed both during the training and testing sessions, and the explanatory power is high, and the total prediction errors are small. ANN was chosen as the best model in the reconstruction of missing streamflow data at the Arandu station due to its superior generalization capability as well as balanced error structure.

3.4 Missing Discharge at Arandu Station Reconstruction (1981-2024).

The trained ANN model was used to recreate missing discharge values in the Arandu station by referring to the long-term upstream station discharge history in the Chitral station. The full reconstructed discharge series of Arandu in 1981-2024 is shown in, which is a combination of measured values and ANN-based predictions.

Figure 6

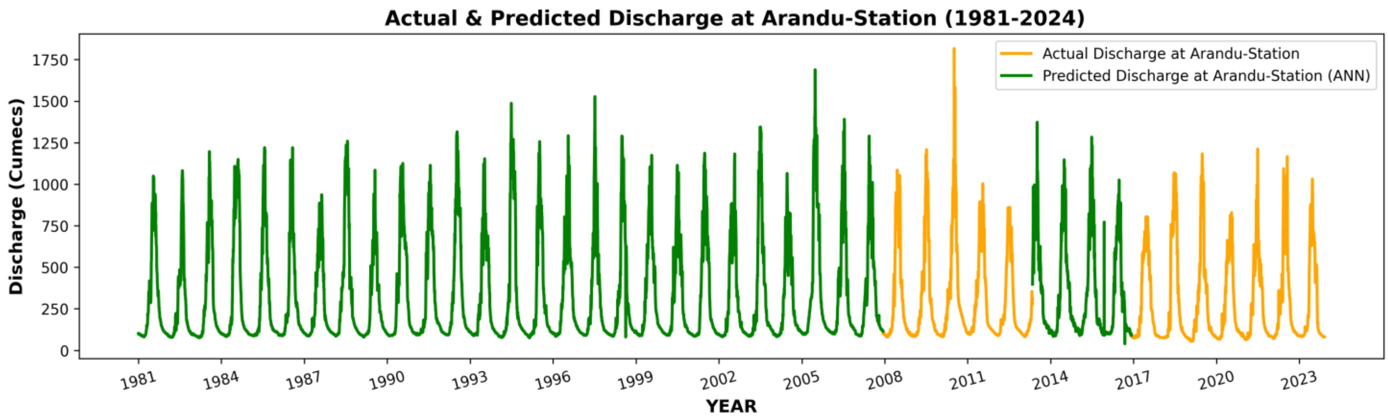

Figure 6: Actual and Predicted discharge at Arandu (1981-2024)

The reconstituted hydrograph maintains the seasonal snowmelt-dominated flow regime within the basin, and annual peaks in the spring and summer and low flows in the winter. The ANN-based estimates blend seamlessly with the observed records of discharges, meaning that there is consistency in terms of time and time-respecting plausibility. This re-created long-term data is, in effect, a virtual discharge sensor, which has been extended by over 20 years of useful Arandu data

4. Discussion

The findings of this work have validated that machine-learning models can be used as a useful tool in recreating missing records of streamflow in snow- and glacier-based mountain basins where traditional hydrological surveillance practices are constrained. The high coefficients of determination throughout all the models confirm that there is a high hydrological relationship between the upstream Chitral and downstream Arandu stations, which supports the validity of upstream-downstream information transfer in the basin. Nevertheless, the distinction between model behavior is observed when metrics of error and generalization are analyzed.

Though relatively simple and easy to interpret, Linear Regression demonstrated greater prediction inaccuracies and dispersion during medium-high flows, which demonstrates its inability to capture the non-linear nature of the processes related to cryosphere melting. XG Boost performed well in terms of fitting in the training phase and poor in accuracy in the testing phase, indicating that it is partially overfitting and not that robust when it is used on independent data. The sensitivity of complex ensemble models to variability of hydroclimatic conditions and small sample sizes in mountainous settings is emphasized by this behavior.

Conversely, the ANN model exhibited a high predictive ability both in training and testing with less magnitudes of error and closer fitting about the 1:1 line especially at the peak flows. This balanced performance depicts a higher generalization and stability, and ANN is therefore more applicable in long term streamflow reconstruction. The reconstructed series at Arandu of reconstructed discharges retains the snow melting regime which is seasonally controlled and has temporal consistency with observed discharges, which confirms its physical plausibility. The usage of reconstructed series as a virtual hydrometric station would offer a viable way to expand streamflow records in under-gauged parts of the Upper Indus Basin and other high-altitude catchments.

5. Conclusions

This paper shows that machine-learning methods can be used to reconstruct lost streamflow data as in the Chitral River. Using the high hydrological interrelations between upstream and downstream stations, a reconstruction of the missing discharge of the Arandu station was achieved over a 43-year span. Though all the analyzed models demonstrated a high level of explanatory power, the Artificial Neural Networks continued to outperform Linear Regression and XG Boost in regard to making predictions and their capacity to generalize. The reconstruction of the ANN based reconstruction provided a continuous and physical realistic discharge series that maintains the snowmelt dominated seasonality of a basin. The long dataset, in effect, serves as a virtual discharge sensor and will play a major role in improving availability of long-term hydrologic data on the basin. The proposed framework has the benefit of providing a transferable and firm solution to close the data divide in other under-gauged mountain systems and transboundary rivers to aid in better hydrological evaluation, infrastructure design, and analysis of impact of climate changes.

6. Patents

No Patents have resulted from the work reported in this manuscript.

Author Contributions:

The conceptualization, methodology development, data collection, formal analysis, model testing, and data validation, initial drafting and final pape preparation were carried out by Muhammad Faarid Shah, Muhammad Farrukh and Fahad Shamraiz. Touqeer Ali Rind played a key role in the initial drafting, methodology development, and critical analysis of the results. Farjad Aziz contributed to data validation and model implementation. Asad Wahab and Qazi Khurshid Ahmed supported the manuscript preparation, including reviewing and editing.

All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This study did not involve humans or animals and therefore did not require ethical approval.

Informed Consent Statement: Not applicable. This study did not involve human participants.

Data Availability Statement: The discharge data used in this study were obtained from hydrometric stations operated by the Water and Power Development Authority (WAPDA), Pakistan. Restrictions apply to the availability of these data, which were used under permission for the current study and are not publicly available. Processed data supporting the findings of this study may be made available from the corresponding author upon reasonable request and subject to data-sharing approvals.

Acknowledgments: The authors acknowledge the Water and Power Development Authority (WAPDA), Pakistan, for providing hydrological discharge data used in this study. The authors also acknowledge the use of generative artificial intelligence tools during the preparation of this manuscript for language refinement and structural editing. The authors have reviewed and edited all generated content and take full responsibility for the accuracy and integrity of the final manuscript.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
LR	Linear Regression
XG Boost	Extreme Gradient Boosting
UIB	Upper Indus Basin
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error

References

[1] G. R. Arathy Nair, S. Adarsh, A. El-Shafie, and A. N. Ahmed, "Enhancing hydrological data completeness: A performance evaluation of various machine learning techniques using probabilistic fusion imputer with neural networks for streamflow data reconstruction," *Journal of Hydrology*, vol. 639, p. 131583, Aug. 2024, doi: 10.1016/j.jhydrol.2024.131583.

[2] H. Solanki and V. Mishra, "Filling Streamflow Data Gaps in Indian Catchments Using Machine Learning." Accessed: Jan. 27, 2026. [Online]. Available: https://meetingorganizer.copernicus.org/EGU25/EGU25-15030.html

[3] H. Dastour and Q. K. Hassan, "A Machine-Learning Framework for Modeling and Predicting Monthly Streamflow Time Series," *Hydrology*, vol. 10, no. 4, p. 95, Apr. 2023, doi: 10.3390/hydrology10040095.

[4] J. L. Ng *et al.*, "Machine Learning-Based Approach for Filling Gaps in Streamflow Data," *Semarak Int. J. Mach. Learn.*, vol. 5, no. 1, pp. 46–63, May 2025, doi: 10.37934/sijml.5.1.4663a.

[5] K. Ramesh *et al.*, "Machine Learning to augment global flood modelling in ungauged basins." Accessed: Jan. 27, 2026. [Online]. Available: https://meetingorganizer.copernicus.org/EGU25/EGU25-7069.html

[6] W. W. Immerzeel *et al.*, "Importance and vulnerability of the world's water towers," *Nature*, vol. 577, no. 7790, pp. 364–369, Jan. 2020, doi: 10.1038/s41586-019-1822-y.

[7] A. F. Lutz, W. W. Immerzeel, A. B. Shrestha, and M. F. P. Bierkens, "Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation," *Nature Clim Change*, vol. 4, no. 7, pp. 587–592, Jul. 2014, doi: 10.1038/nclimate2237.

[8] M. Hrachowitz *et al.*, "A decade of Predictions in Ungauged Basins (PUB)—a review," *Hydrological Sciences Journal*, vol. 58, no. 6, pp. 1198–1255, Aug. 2013, doi: 10.1080/02626667.2013.803183.

[9] L. Oudin, V. Andréassian, C. Perrin, C. Michel, and N. Le Moine, "Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments," *Water Resources Research*, vol. 44, no. 3, p. 2007WR006240, Mar. 2008, doi: 10.1029/2007WR006240.

[10] ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, "Artificial Neural Networks in Hydrology. I: Preliminary Concepts," *J. Hydrol. Eng.*, vol. 5, no. 2, pp. 115–123, Apr. 2000, doi: 10.1061/(ASCE)1084-0699(2000)5:2(115).

[11] F. Akhtar, C. Borgemeister, B. Tischbein, and U. K. Awan, "Metrics Assessment and Streamflow Modeling under Changing Climate in a Data-Scarce Heterogeneous Region: A Case Study of the Kabul River Basin," *Water*, vol. 14, no. 11, p. 1697, May 2022, doi: 10.3390/w14111697.

[12] B. He and K. Takase, "Application of the Artificial Neural Network Method to Estimate the Missing Hydrologic Data," *JOURNAL OF JAPAN SOCIETY OF HYDROLOGY & WATER RESOURCES*, vol. 19, no. 4, pp. 249–257, 2006, doi: 10.3178/jjshwr.19.249.

[13] G. R. D. Souza, I. P. Bello, F. V. Corrêa, and L. F. C. D. Oliveira, “Artificial Neural Networks for Filling Missing Streamflow Data in Rio do Carmo Basin, Minas Gerais, Brazil,” *Braz. arch. biol. technol.*, vol. 63, p. e20180522, 2020, doi: 10.1590/1678-4324-2020180522.

[14] D. R. Legates and G. J. McCabe, “Evaluating the use of ‘goodness-of-fit’ Measures in hydrologic and hydroclimatic model validation,” *Water Resources Research*, vol. 35, no. 1, pp. 233–241, Jan. 1999, doi: 10.1029/1998WR900018.

[15] C. J. Willmott and K. Matsuura, “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance,” *Climate Research*, vol. 30, pp. 79–82, Dec. 2005, doi: 10.3354/cr030079.

[16] D. S. K. Karunasingha, “Root mean square error or mean absolute error? Use their ratio as well,” *Information Sciences*, vol. 585, pp. 609–629, Mar. 2022, doi: 10.1016/j.ins.2021.11.036.