Application of ANN and Snow Melt Runoff Model for Streamflow Prediction Over Scarcely Gauged Catchments

By
Amreen Aslam
(2020-NUST-MS-GIS-328406)

Institute of Geographical Information Systems
School of Civil and Environmental Engineering
National University of Sciences & Technology (NUST)
Islamabad, Pakistan
August, 2024

Application of ANN and Snow Melt Runoff Model for Streamflow Prediction Over Scarcely Gauged Catchments

By

Amreen Aslam

(Registration No: 328406)

A thesis submitted in partial fulfillment of the requirements for the degree of Master in Science in Remote Sensing and GIS

Institute of Geographical Information Systems
School of Civil and Environmental Engineering
National University of Sciences & Technology (NUST)
Islamabad, Pakistan

August, 2024

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Amreen Aslam (Registration No. MSRSGIS 00000328406), of Session 2020 (Institute of Geographical Information systems)** has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulation, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ______________
Name of Supervisor: Dr Muhammad Azmat
Date: 13-9-24

Signature (HOD): ______________
Date: 4-10-2024

Dr. Muhammad Ali Tahi
HoD & Assoc. Prof. SCEE (IGI
NUST, H-12, Islamabad.

Signature (Principal & Dean SCEE): ______________
Date: 07 OCT 2024

PROF DR MUHAMMAD IRFA
Principal & Dean
SCEE, NUST

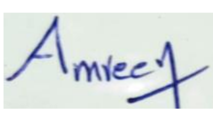
ACADEMIC THESIS. DECLARATION OF AUTHORSHIP

I **Amreen Aslam**, affirm that this thesis and the research contained within are my original work and have been conducted independently by me.

Title: Application of ANN and Snow Melt Runoff Model for Streamflow Prediction Over Scarcely Gauged Catchments

I hereby declare that:

1. This thesis consists of my original work and does not include any material published or authored by others, except where properly cited.
2. Any portions of this thesis that have been submitted for a degree or other qualification at this or any other institution are clearly indicated.
3. I have acknowledged all significant sources of assistance.
4. In cases where this thesis is based on collaborative work, I have clearly distinguished my contributions from those of others.
5. None of this work has been published prior to submission.
6. This work complies with the HEC plagiarism policy and is not plagiarized.



Signed:

Date: 6-09-2024

DEDICATION

To

Allah Almighty

&

My loving parents, my supportive husband, and our precious little ones

I am deeply grateful for the love, care, and encouragement I have received since the beginning of my studies. Thank you to everyone who supported and prayed for me throughout the process of completing this thesis. Your motivation has meant a lot to me.

ACKNOWLEDGEMENTS

All praises to almighty Allah who to whom everything belongs and to whom we all will return. First, I would like to extend my heartfelt gratitude to my supervisor, Dr. Muhammad Azmat for his devoted guidance throughout my research period and thankful to my examination committee members Dr. Salman Atif, and Dr. Latif Anjam. Their support, guidance and suggestions helped me a lot. I am also thankful to the IGIS, NUST and all administration, which supported me to do research in this developing and growing field.

I am also grateful for the direct or indirect, considerate support and help of those whom I didn't mention here. I am thankful for their sincere efforts.

Finally, my gratitude to my family for their trust in me and providing me with moral and financial support to acquire my ambition, I just can't thank them enough.

Amreen Aslam

Table of Contents

List of Figures

List of Tables

Abstract

Currently, climate change, floods, and other extreme weather events are becoming increasingly common, resulting in significant impacts on society. Predicting streamflow is essential for effective hydrology and water resource management. Accurate and timely predictions of streamflow help in water allocation, flood forecasting, and reservoir management. Forecasting streamflow in cold regions presents challenges due to the significant annual and the seasonal changes in the natural processes occurring within catchments. The research aims to contribute to the understanding of stream flow dynamics in scarcely gauged catchments and to offer significant perspectives regarding water resource management in the study area. The purpose of this study is to evaluate the application of ANN and Snowmelt run off model for streamflow prediction over scarcely gauge catchments like Astore river, Hunza and Gilgit river. For this purpose, Climatic variable data, including precipitation, temperature and Observed Streamflow data are obtained from reputable sources like Meteorological-Department and Water & Power Development Authority in Pakistan. Snow covers data from MODIS imageries and DEM from USGS earth explorer. Three different ANNs models are implemented, The Keras Sequential Feedforward Single Layer Perceptron (SLP) Model, The Keras Sequential Feedforward with Backpropagation Model Multi-Layer-Perceptron and Radial-Basis-Function Models are used. In results Multi-Layer Perceptron (MLP) model performed effectively and yielded the best results R^2 98% and NSE 99% during the validation and testing phases as compared to others. The findings of this research indicate the effectiveness of artificial neural network models in capturing the complex relationships between climatic variables, snow cover/glacier dynamics, and stream flow variability,

achieving high predictive accuracy across different temporal scales and spatial locations within the study area as compared to Snowmelt runoff model.

CHAPTER 1

INTROUDCTION

1.1 Introduction

The increasing frequency of climate change-induced anomalies, including floods and erratic precipitation patterns, poses significant challenges to water resource management in mountainous regions such as the Hunza and Astore Valley of Gilgit Baltistan in Pakistan. Streamflow prediction plays a crucial role in mitigating the adverse impacts of these hydrological hazards by facilitating informed decision-making in water allocation, flood forecasting, reservoir management, and ecological preservation. However, accurate streamflow prediction in data-scarce mountainous catchments presents a formidable challenge due to the complex hydrological processes and limited availability of observational data. Addressing these challenges requires the development and application of advanced modeling techniques capable of capturing the intricate interactions between meteorological, glaciological, and hydrological processes. In recent years, ANN (Artificial Neural Network) has become an effective resource for streamflow prediction, leveraging the computational capabilities compelled by the framework and functioning of neural networks that are biological. The versatility of these models lies in their ability to learn complex patterns and relationships from data, handle nonlinearity, and generalize well to unseen conditions. Previous studies, such as those carried out by (Zheng and Mahmood, 2018 and Abid et al.,2014), have established the effectiveness of ANN models in forecasting runoff from various catchment areas, including the Waikato River in New Zealand and water reservoir discharge in Pakistan. By incorporating climatic variables, snow cover/glacier dynamics, and observed streamflow data, ANN models offer a promising approach for streamflow prediction in scarcely gauged mountainous catchments. However, while ANN models provide valuable insights into streamflow dynamics, they are not without limitations. Data scarcity, non-stationarity of hydrological processes, and uncertainties in model calibration pose significant challenges to their application in mountainous regions. Furthermore, the reliance on empirical relationships may limit the interpretability and transferability of ANN models, particularly in ungauged catchments and future climate scenarios. Addressing these challenges requires a holistic approach that integrates data-driven modeling techniques with physical models, like the Snow Melt Runoff (SRM), to enhance reliability and robustness of streamflow predictions. By leveraging a diverse range of input datasets, including climatic variables, snow cover/glacier dynamics, and observed streamflow data, this research aims to enhance the accuracy and consistency of streamflow predictions in data-scarce mountainous catchments. By conducting thorough model calibration, validation, and comparative analysis, this study aims to highlight the strengths and weaknesses of various modeling approaches, offering valuable insights for water resource management and strategies to adapt to climate change in the region.

The Indus Basin, a significant river basin in Asia, supplies over 70% of water to the Pakistan’s dry and low-lying areas (Mukho et al., 2015). Anchored in the Hindukush-Karakoram and Himalaya (HKH) mountain ranges, the Indus River's sustenance chiefly relies on snow and glacial melt water (Adnan et al., 2017). The Indus River's water supply holds multifaceted importance, serving as a critical resource for irrigation, electricity generation, and a significant source of clean water for downstream populations

(Immerzeel et al., 2009). The designation of the Hindukush-Karakoram-Himalaya (HKH) region as the "water tower of Asia" underscores its pivotal role, as warming trends in this area raise serious environmental concerns and garner substantial scientific attention. The far-reaching impacts of global climate change have deleterious effects on snow and glaciers within Pakistan's HKH domain, which contributes more than half to the total run-off of the Indus River basin (Ashraf et al., 2012). The evolving flow patterns, influenced by climate change and other factors, have the potential to intensify tensions between provinces. Downstream areas, particularly the Sindh province, face the dual challenges of reduced water availability during dry seasons and heightened flood risks in wet seasons. Over the past three decades, the HKH region has experienced an increase in temp of 1.5°C, twice the (0.7 ºC) observed in further areas of the country (Rasul, 2012). In an era marked by mounting populations, the repercussions of global warming, and the detrimental impact of pollution on water quality, the significance of Earth's most precious resource, water, is consistently escalating. This intensifying importance stems from the expanding demands across residential, industrial, and agricultural sectors acknowledge these trends (Yeleliere et al., 2018). The principal origin of freshwater rests in precipitation, manifested through rainfall or snowfall. Notably, mountains experience substantial levels of both rain and snow, rendering them the primary well-spring of fresh water.

The pivotal role of snow-covered regions lies in the phenomenon of snow-induced runoff, catalyzed by snowmelt, which predominantly initiates during the spring season coinciding with amplified requirements for water (Krishna et al., 2011). Over one-sixth of the global population relies the water sourced originating by snowmelt for their water supply. The prevailing and foremost challenge confronting the world today is the phenomenon of climate change and its intertwined aspect of global warming. The repercussions of this environmental transformation on water supplies are notably diverse and often unpredictable. As we progress toward the conclusion of the current century, substantial alterations in climate patterns are anticipated due to the persistent escalation of greenhouse gas emissions. The trajectory of climate change is an unequivocal certainty, driven by the escalating global mean temperatures and the profound alterations inflicted upon the atmosphere's chemical composition by human activities. The elevated mountainous terrain has borne a substantial brunt of the repercussions of global climate change in the recent decades. Notably, snow, glaciers, and permafrost exhibit heightened vulnerability to shifts in climatic factors due to their closeness to melting thresholds. Indeed, one of the most conspicuous outcomes of rising temperatures is the transformation in the presence of ice and its cascading effects on the physical dynamics of elevated mountain systems (Haberli, 1990).

Snowmelt and glacier-melt derived water sources are indispensable for over one-sixth of the global populace's water provisioning. The prevailing paramount challenge confronting the world pertains to the events of global climate change. This environmental upheaval's repercussions on water supplies, specifically, exhibit a marked propensity for diversity and unpredictability. The trajectory towards the century's conclusion portends substantial climate alterations, attributed to the relentless surge in greenhouse gas emissions. The inevitable continuity of climate evolution is a foregone conclusion, substantiated by the ascending global mean temperatures, propelled by profound alterations in the atmospheric chemical composition induced by human activities. Over recent decades, the elevated mountainous milieu has borne a pronounced impact of global climate change. Notably, the transformative effects of shifts in air conditions hold resonance for snow, glaciers, and permafrost due to their proximity to critical melting thresholds. In essence, one of the most perceptible outcomes of escalating temperatures could manifest as alterations in ice occurrences and their cascading ramifications for the upward trajectory of temperatures, coupled with shifts in precipitation patterns, amplifies the hydrological cycle, leading to increased stream flow volume and its temporal distribution across the year. This phenomenon, however, gives rise to notable water stress (Houghton et al., 2021). Kundzewicz et al., (2007) depicted the discharge within glacier-fed rivers is anticipated to experience an initial upsurge, followed by a decline over the ensuing decades due to the gradual diminution of ice storage. This underscores the pivotal role of temperature fluctuations in governing the allocation of precipitation. Precipitation, whether in the form of snow, rain, glacier melting, or a combination thereof, significantly contributes to freshwater reserves. The reduction in glacier coverage across mountainous regions has become pronounced due to ongoing global warming trends. A recent investigation by the United States Geological Survey, utilizing historical glacier data and imagery from NASA's TERRA satellite's ASTER instrument, indicates noteworthy reduction in the size of mountain glaciers within areas such as the Andes, the Himalayas, the Alps, and the Pyrenees during the last ten years (Wessels et al., 2002). These findings align with the widespread conclusions drawn from various glacier studies conducted globally, all of which highlight the rapid pace of glacial retreat in recent times. Meier (1997) conducted a comprehensive study that encompassed more than 200 glaciers and found a global retreat in glaciated area ranging from 6,000 and 8,000 km^2 between

1960's and 1990's. Hoelzle (2001) from the organization which monitors the Glaciers of the World (WGMS) noted that observations spanning the past century unmistakably demonstrate a global-scale reduction in mountain glaciers. Their findings indicate that this trend reached its zenith in the early 20th century, followed by a resurgence in glacier growth around 1950. However, during the 1980s, the pace of glacier retreat escalated once again, exceeding the scope of pre-industrial variability. The IPCC's Second Assessment Report (1996), based on comprehensive scientific research, forecasted that approximately 25% of the total mass of mountain glaciers worldwide might vanish by 2050, and as much as 50% by the year 2100 (Rees and Collins, 2004). IPCC's (2007) report further highlighted that in certain moist tropical areas and high latitudes, annual average river discharge and water availability might witness a rise of 10–40% by the mid-century, while in select dry regions and mid-latitudes, they could decline by 10–30%. Presently, over 16% of the global population live in these areas that depend on meltwater from major mountain ranges for sustenance; nevertheless, projections indicate a decline in water reservoirs stocked in glaciers and snow cover throughout the 21st century. Situated between 24° and 38° N and 61° and 78° E, Pakistan stands as one of the many developing nations in the region. The country heavily relies on snow and ice melts in its mountainous areas to fill a huge portion of its freshwater supplies. However, the susceptibility of this water source to climate variations introduces various potential impacts. Pakistan's vulnerability is heightened due to its substantial reliance on uninterrupted river flow for sustained water access and power generation. The Indus River, originating in northern Pakistan and terminating at the Arabian Sea in the south, forms a pivotal component of this system (Hayat et al., 2019). Given Pakistan's diverse climatic conditions, the ramifications of climate change could be particularly pronounced. The complex interaction between melting snow and glacier runoff within the Indus Basin bears great significance for the agricultural-based economy of the country (Archer et al., 2010).

1.2 Objectives

These are the following objectives of the study are:

- Application of ANN models for accuracy assessment to predict streamflow using multiple data limitation scenarios over complex hydrology catchments (i.e. scarcely gauged and high-altitude snow and glacier covered).
- Comparison of ANN models with physical based hydrological models (i.e. Snowmelt Runoff Model) under data limitation scenarios.

CHAPTER 2

LITERATURE REVIEW

The methods for predicting river flow in frigid areas were established in early 1970s, when advancements in calculating technology enabled the calculations of large data sets through digital techniques (WMO, 1975; Singh, 2018). The field of hydrology has witnessed a surge in research focusing on stream flow prediction methods, driven by the growing urgency to address water resource management challenges exacerbated by climate change. Various approaches, including empirical, physical & data models, have been employed to forecast stream flow accurately. Among these, ANN models have garnered significant attention for their capacity to handle complex nonlinear relationships inherent in hydrological data.

The progress in technology today enables researchers to create insightful simulations of the interrelated hydrological processes that play a role in river flow within the water cycle. The progress in technology facilitates effective implementations of flood analysis, assessment and the evaluation of climate change effects on water resources (Cunderlik and Burn, 2003; Velázquez et al., 2013; Kouchak et al., 2015).

The administration of water resources is facing an escalating burden as a consequence of the rising population and the diverse demands for water across various sectors, including agriculture, industry, domestic use, recreation, and environmental needs. Furthermore, climate change serves as a significant challenge for river flow forecasting and hydrological modeling, as rising temperature correlates with alterations in precipitation patterns, snow dynamics, and soil in-filtration processes. These changes impact both the volume and seasonal distribution of the stream-flow in colder regions (Aygun et al., 2020).

A study done by Zaghlou et al., (2022) on simulating river discharge in frigid and unmonitored areas. The Study review of different methods of the modelling river flow in ungauged basins to overcome this restriction, "Predictions in Ungauged Basins" initiative provides various studies focused on enhancing forecasting accuracy. This includes the use of regression, calibration, interpolation techniques. Process models have significantly improved by integrating remote sensors for data to reproduce and analyze complex processes. Moreover, the empirical models that rely on observed data for graphical solutions rather than mathematical formulations are increasingly being enhanced by advancements in machine learning, leading to remarkable forecasting accuracy. Snow cover data from remote-sensing maps (MODIS) can be utilized to calculate snowmelt in hilly regions. The study conducted by Hayat et al., (2019) emphasis is on predicting streamflow in the Astore and Hunza basins located in the Hindukush Karakoram Himalayan region through the application of the Snowmelt Runoff Model (SRM) Climate forecasts derived from both global and regional modeling approaches (Representative Concentration Pathways; RCP 2.6, 4.5, and 8.5) were integrated to SRM for assessing effects on water resources. Tthe successful application of this approach indicates that the SRM could be effectively used to simulate annual stream flow trends in mountainous watersheds characterized by snowmelt dominance. It is very important to recognize that model uncertainty tends to increase significantly with longer forecasting horizons.

Renaud and Robert (2022) focused on stream-flow forecasting, specifically for one to seven days ahead by using an ANN model into northeastern U.S. watersheds. They employed a basic modeling framework, generating training and validation data with deterministic distributed model spanning 16 summers (2000-2015). To evaluate how different input variables affect forecasting accuracy, the study analyzed the ANN model's performance with various combinations of input data. The input datasets included the deep-soil moisture, the surface-soil-moisture and the streamflow from the previous day, and a combination of surface soil moisture and antecedent observed stream flow. The findings indicated that the most effective combination for enhancing forecast accuracy involved using soil moisture content alongside observed stream flow for both watersheds.

Dolling and Eduardo (2002) worked on using artificial neural networks to predict monthly streamflow in mountainous watersheds. The process involves selecting input variables and the variables for training and testing the neural model. Applying artificial neural networks to streamflow prediction entails making decisions on four main aspects: choosing the most relevant variables for explaining runoff, designing optimal network architecture, selecting the best learning strategy, and choosing the system that most accurately reflects the stream flow patterns. It is crucial to create a systematic procedure to develop a network that captures the majority of predictable patterns in the data and can accurately represent different scenarios not included in the training data. The methodology first involves selecting the most relevant variables which are input & output, as well as choosing the learning, testing, and validation datasets for analysis. Proper collection and analysis of relevant information is essential in order to obtain accurate and representative data for the modeling process. This initial step is crucial, as the entire process relies on the quality of the data (Linsley et al., 1988). The second step involves selecting the output variables. Model 1 has only one output variable, while

Model 2 uses seven output variables at seven different months from July to January. The next step is selecting the input variables that affect the streamflow order. Following this, the data sets are divided into three groups wet, normal and dry on the quantity of water in the stream during specific time periods. The model uses 40% of the data for learning and validation, and 20% for testing. The next step involves selecting the model, which includes defining the architecture, training process, validation, and determining the optimal model. The Cascade-Correlation algorithm, introduced by Fahlman and Lebiere (1990), suggested for determining the quantity of neurons in the hidden layer (Zell et al., 1995). The training, validation, and testing sets remain consistent across the five different model types mentioned earlier. Subsequently, the weights needed to achieve the desired output from the ANN models Determinations are made through a process of learning that utilizes the input variables. The optimal set of weights minimizes objective which means square error. Following this, model validation occurs, during which the weights are fine-tuned to reduce the Root- Mean-squar-Error. This methodology effectively produced a model for the stream flows prediction in Argentina, using climatological data from the Pachon meteorological station at an altitude of 1900 meters. The San Juan River basin is situated in a mountainous area of the Andes, where precipitation primarily falls as snow. The network was primed with arbitrary values between ±1, and the convergence of the propagation was analyzed, testing various rates for learning and coefficients. The final values were 0.9 and 0.7. Results were presented in tables and graphs, concluding that the neural network models demonstrated superior performance in predicting monthly spring and summer streamflow compared to alternative methods. This proposed approach offers significant advantages for management of water resources for irrigation and hydroelectric power generation.

Abid et al., (2014) measure the water discharge at Hemren Reservoir as a case study for utilizing artificial neural networks to manage and forecast water reservoir discharge. The prediction of various hydrological phenomena is crucial for effective water resource management. From an engineering perspective, accurately forecasting components of natural reservoirs, such as inflow, is essential for multiple applications. The resulting techniques vary greatly depend on the purposes and the characteristic's and also on data the best methods is to overcomes the uncertainty to predict the Hemren reservoir inflow was used the Artificial neural networks approach has adopted. The datasets used include monthly discharge and precipitation data, specifically focusing on the intensity of rainfall in the intermediate catchment area between the Hemren and Derbendi Khan dams. Various network configurations were tested for employing a comprehensive dataset spanning 24 years (1980-2004), for training and testing purposes. The model featured three input layers, with 40 neurons distributed across two layers which are hidden and the output layer. The correctness of model was assessed by using Mean-Square-Error (MSE) and the correlation coefficient. The network was successfully trained, achieving convergence at an MSE of 0.027 through the early stopping method. The training and testing phases yielded (0.97, 0.77) coefficient correlations indicating a high level of accuracy for the prediction technique.

Zheng and Mahmood (2014) conducted research on implementing an ANN structure to predict overflow in Waikato-Catchment area in the New-Zealand. The effective use of the ANN technique, similar to other modeling methods, depends on the careful selection of input variables. This River is the very longest in New-Zealand's North Island, stretching approximately 425 km and featuring eight dams and nine hydroelectric stations, making it economically significant for the country (Abbas & Saad 2014).

The study utilized daily rainfall and runoff data spanning 10 years, starting from May 24, 2002. The arithmetic mean method was applied to determine average values, with data sourced from the Whangamarino Control Structure, which covers a large portion of the catchment area. The MATLAB neural network tool was employed for the runoff prediction model. A feedforward neural network (FFNN) was used, incorporating the backpropagation algorithm, which is widely utilized in hydrological applications (Adamowski et al. 2012).

In this study, the choice of transfer function for the ANN was critical. The hyperbolic tangent purpose is to select hidden-layer neurons, transforming inputs from negative to positive infinity into outputs ranging from -1 to 1. Linear function was employed, producing outputs within a negative to positive infinity range.

The process of neural network based on initial weights and stopping criteria, this is categorized into three distinct subsets: training, testing and validation. Training subcategory was used to calculate and update the biases and the weights, while validation subcategory helped stop training when the squared error reached its minimum. The testing subset evaluated the network's performance. The backpropagation algorithm adjusted weights and biases to minimize error, with the Levenberg-Marquardt Algorithm (LMA) chosen for training.

A graph was created to illustrate the flow duration curve, categorizing flows from high to low discharge. The research employed the correlation coefficient (R^2), root mean square error (RMSE), and Nash-Sutcliffe Efficiency (NSE) as metrics to evaluate the performance of the model. Two techniques for selecting input vectors were applied: the sequential and pruned time series along with a non-sequential time series approach. Ultimately, this model using non sequential technique and retrieve the highest R^2 while minimizing error.

A pivotal contribution to the validation of snowmelt runoff simulations was achieved through an examination of the Beas River basin, situated at the Pandoh Dam in India (Duan et al. 2019 and Prasad & Roy, 2005). In this endeavor, the basin's topography was meticulously subdivided into 12 elevation zones, each with a maximum elevation of 500 meters. Utilizing input parameters extracted from diverse sources such as satellite data, hydro-logical data, meteorological data and the existing maps, measured and estimated runoffs displayed compelling consistency, affirming the efficacy of the model. Additionally, the application of SRM extended to the Astore River, a component of northern Pakistan's Upper Indus Basin, as investigated by Butt & Bilal (2011). This study further exemplified SRM's versatility and viability in simulating snowmelt runoff dynamics within distinct geographical contexts. To validate the effectiveness of a synergistic application involving the Snowmelt Runoff Model and data obtained through remote sensing for modeling the daily variations in streamflow within catchments characterized by significant snow presence, a comprehensive examination was undertaken in North-East Region of the Iran (Ha et al. 2021; Firouzi et al., 2016).

Within the extensive body of literature, a multitude of studies have undertaken an assessment of the efficacy of the SRM in gauging streamlining simulations across diverse basins worldwide (Azmat et al. 2015). An exemplary instance involves the selection of SRM to model and predict daily discharge for multiple basins within the Spanish Pyrenees (Gómez-Landesa & Rango, 2002). In this endeavor, the snow cover image is derived through a linear pairing of channels 1 and 2 in NOAA. SRM is used to generate real-time forecasts with area snow cover as a key input. The results from these analyses based on SRM significantly contributed to enhancing water resource management practices for hydropower companies operating in the Spanish Pyrenees. Similarly, the applicability of SRM was extended to Nepal’s hilly river basins, where it showcased its potential in water resources planning & management (Nayak et al. 2007; Dhami et al., 2018).

CHAPTER 3

STUDY AREA

This area comprises of three River basins, the Astore, Gilgit and Hunza River, which are part of the mighty (UIB) Upper Indus Basin. The Sub-Indus basins serve as the unique hydrological characteristics and significant relevance to water resource management in the country.

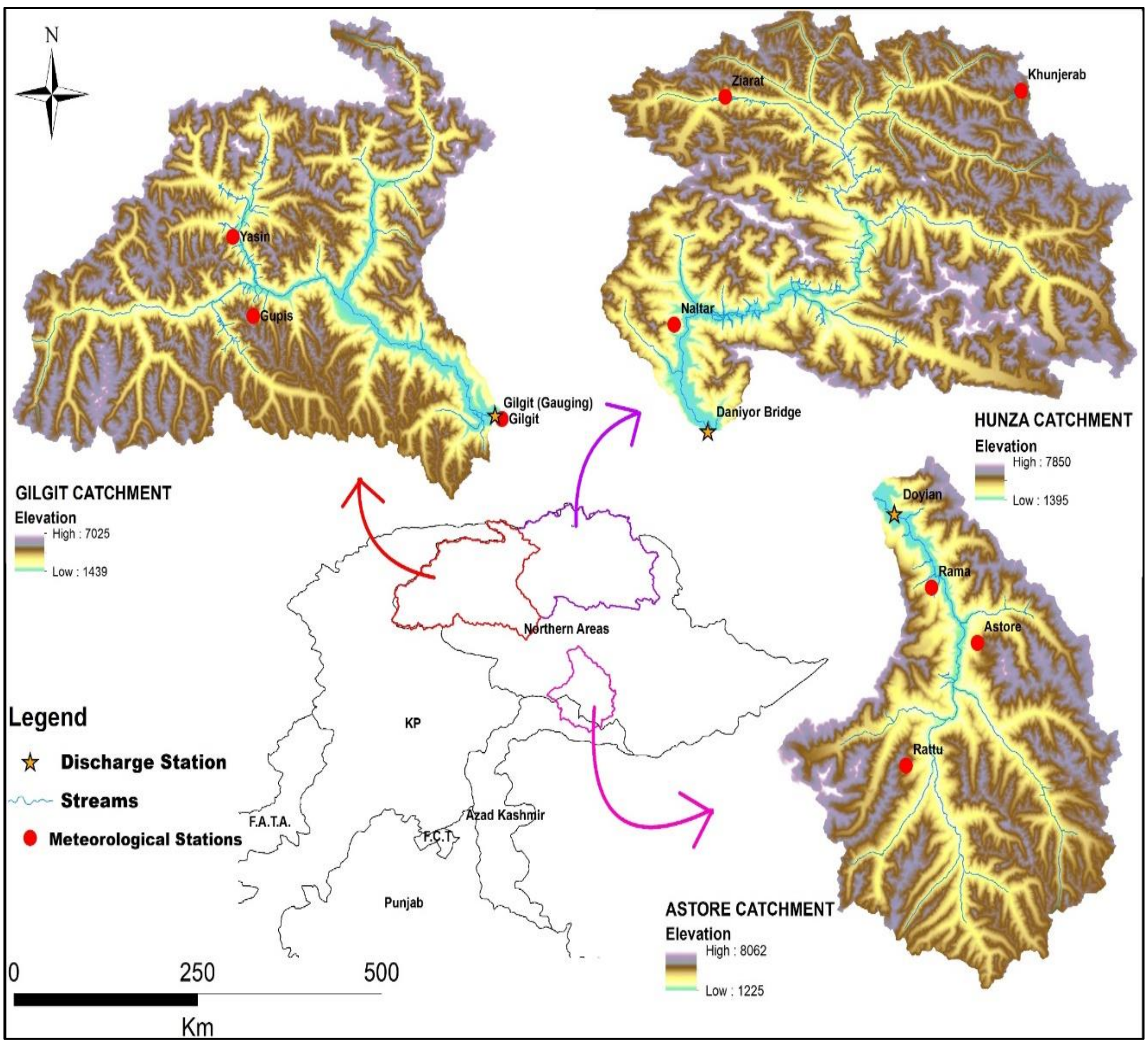


Figure 3.1 Study Area map of Astore, Gilgit and Hunza Rivers

These regions are characterized by their high altitude, rugged terrain, and intricate network of rivers originating from snowmelt and glacier-fed sources. The Hunza River, originating from the Khunjerab Glacier and Gilgit River originating the journey begins at Shandur Lake and continues until it merges with the Indus River in proximity to the towns of Juglot and Bunji and the Astore river originates on the western slopes of Burzil Pass. The Gilgit River basin, which has a large drainage area is situated in Pakistan's Gilgit-Baltistan region, on the eastern side of the Hindukush Range. This valley extends southeast before draining into the enormous Indus River. At the Alam Bridge hydrometric station, the river's discharge is constantly observed. Geographically, the basin is located between 35.80°N and 36.91°N latitudes and between 72.53°E and 74.70°E longitudes. From towering peaks that soar as high as 7,730 meters to low-lying plateaus that are 1,250 meters above sea level, the elevation within the basin varies widely (Ali, 2017). The basin's topography comprises diverse features, with a substantial portion of about 982 km^2 located at elevations exceeding 5,000 meters. A significant proportion of the basin, approximately 8%, is covered by glaciers, making up approximately 4% of the entire Upper Indus Basin's cryosphere extent. This icy domain is characterized by approximately 944 km^2 of clean glacier area and an additional 146 km2 covered by debris. These rivers play pivotal roles in sustaining the local ecosystems and supporting agricultural livelihoods. These rivers traverse through valleys renowned for their scenic beauty and cultural heritage, attracting both tourists and researchers interested in the region's hydrological dynamics.

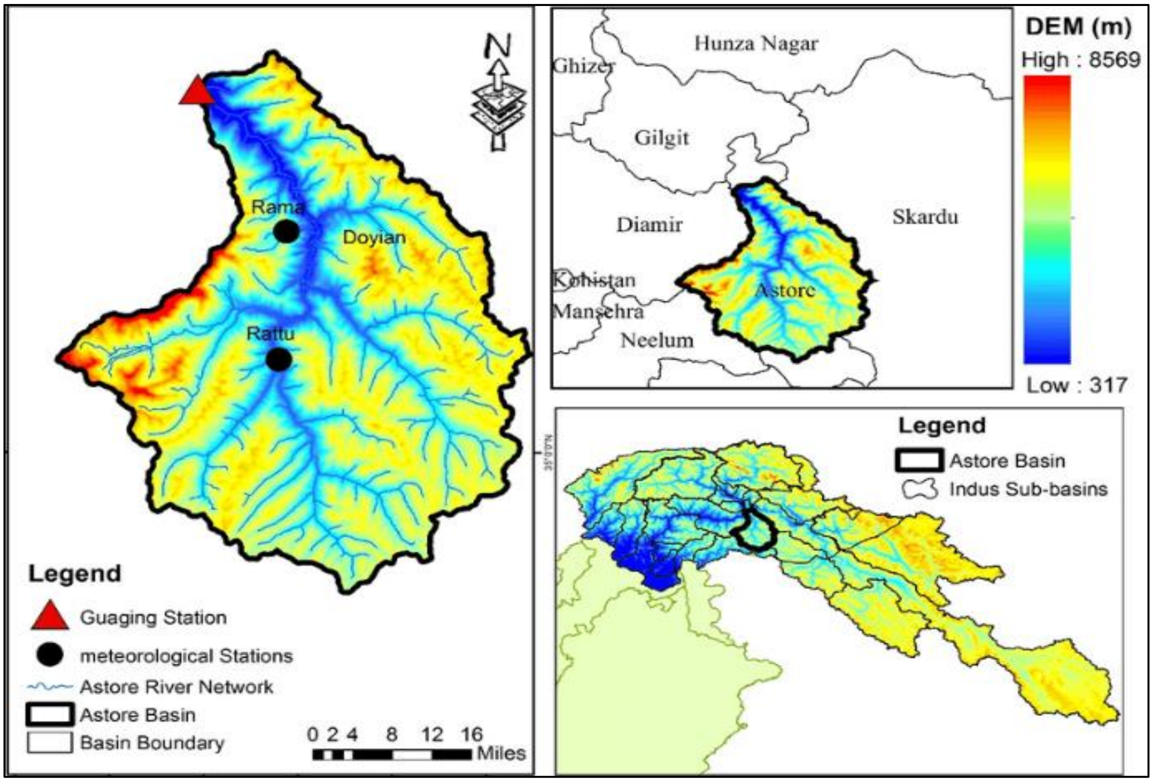


Figure 3.2 Geographical location of Astore River Basin in Upper Indus River Basin.

These are significant water tributaries situated in Upper-Indus-River basin in Northern Pakistan. The length of the Gilgit River is 240 km and Astore river is 220 km and Hunza River is 190 km. The Hunza River catchment, covering a substantial drainage area is situated in high-altitude central Karakoram region of Pakistan. Its geographical boundaries extend from 36.05°N latitude to 37.08°N latitude and 74.04° E longitude to 75.77°E longitude, encompassing a breathtaking landscape. The mean catchment elevation is 4,631 meters, accentuating its high-altitude characteristics. The elevation within the basin exhibits a remarkable range, spanning from 1,432 meters above sea level to towering peaks reaching as high as 7,849 meters. Within the basin, a vast portion of about 4,152 km^2 is covered by glaciers, signifying the significant cryosphere influence in the region. These glaciers serve as vital reservoirs of freshwater, supplying water to the river and contributing to the region's overall hydrology. The Hunza River basin has around 1,384 glaciers. It encompasses a glacier area of 3,673.04 km^2 and a debris covered area of 479.56 km^2 (Ali, 2017).

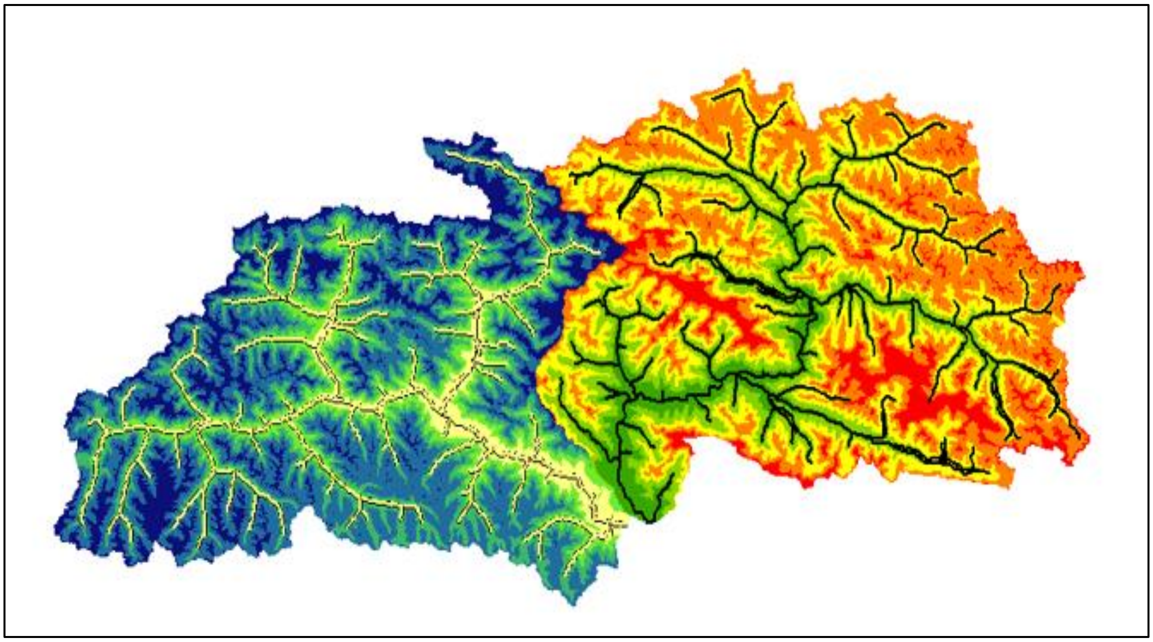


Figure 3.3 Geographical location of Hunza and Gilgit River Basins.

The hydrology of these regions is heavily influenced by seasonal variations in precipitation, snowmelt, and glacier dynamics, making stream flow prediction a challenging yet imperative task. With climate change projections indicating alterations in precipitation patterns and glacier retreat, understanding the hydrological processes governing these regions is crucial for adapting to future water resource scenarios. Additionally, the scarcity of hydrological data in these remote and mountainous regions presents a significant challenge for traditional modeling approaches, necessitating the utilization of innovative techniques, for example ANN models for stream flow forecast (Alqurashi, M. 2021). This seasonal fluctuation in SCA has significant implications for snowmelt runoff, which, in turn, influences the river's flow and water availability downstream. It is contributed by two primary weather systems: westerly disturbances and the summer monsoon. The region receives rainfall and snowfall from westerly disturbances during the winter months. In contrast, the summer monsoon brings moisture-laden winds that lead to heavy rainfall in the basin, replenishing its water resources and supporting the ecological balance. The Gilgit River basin is home to an impressive number of glaciers and glacier lakes, adding to the basin's hydrological complexity. A total of 585 glaciers and 605 glacier lakes are scattered throughout the landscape. Among these glacier lakes, eight are identified as potentially dangerous due to GLOFs, which can present serious risks to downstream communities and infrastructure (Amjad, 2023). Understanding the undercurrents of glacier, lake, and their interactions considering the weather and hydrology is of paramount importance. As climate change continues to exert its influence on this sensitive region, accurate knowledge of the basin's cryosphere processes, and water resources is essential for sustainable water management, disaster risk reduction, and informed decision-making for the well-being of the communities that rely on the Gilgit River basin's resources. The upper Indus Basins hold strategic importance in Pakistan's water resource management landscape, serving as vital sources of freshwater for irrigation, hydropower generation, and domestic consumption. The viability of these water resources is vital for the socio-economic development of local communities and resilience of region to impact of weather change. However, susceptibility of these regions to hydrological risks including flooding and landslides underscores the need for accurate stream flow prediction models to inform early warning systems and disaster preparedness measures.

Despite the ecological and socio-economic significance of these regions, their hydrological dynamics remain relatively understudied, particularly in the context of stream flow prediction using advanced modeling techniques. This research aims to address this knowledge gap by concentrating on the application of ANN models for stream flow prediction, aiming to unravel the complex interactions between climate variability, glacier dynamics, and stream flow regimes in these scarcely gauged catchments. Through an in-depth analysis of the study area's hydrological processes and the development of robust prediction models, this research endeavors to provide valuable insights for sustainable water resource management and strategies for adopting to climate change in the Hunza and Astore Valleys of Pakistan.

3.2 Data Collection and Preprocessing

Collecting high-quality data is essential for the success of any hydrological modeling endeavor, particularly in scarcely gauged catchments such as the Hunza, Gilgit and Astore Valley regions of Pakistan. Daily streamflow data was required to be used as an input in model for calibration and further validation purposes. This data was collected from the Water & Power Development Authority's (WAPDA) project of surface water hydrology. This research draws upon a diverse range of data sources to capture the complex hydrological processes driving stream flow dynamics in these mountainous regions. Climatic variable data, including precipitation, temperature, and humidity, are obtained from reputable sources such as the Pakistan Meteorological Department (PMD) and the Water & Power Development Authority (WAPDA). Digital Elevation Model (DEM) is used which is downloaded from USGS Earth Explorer.

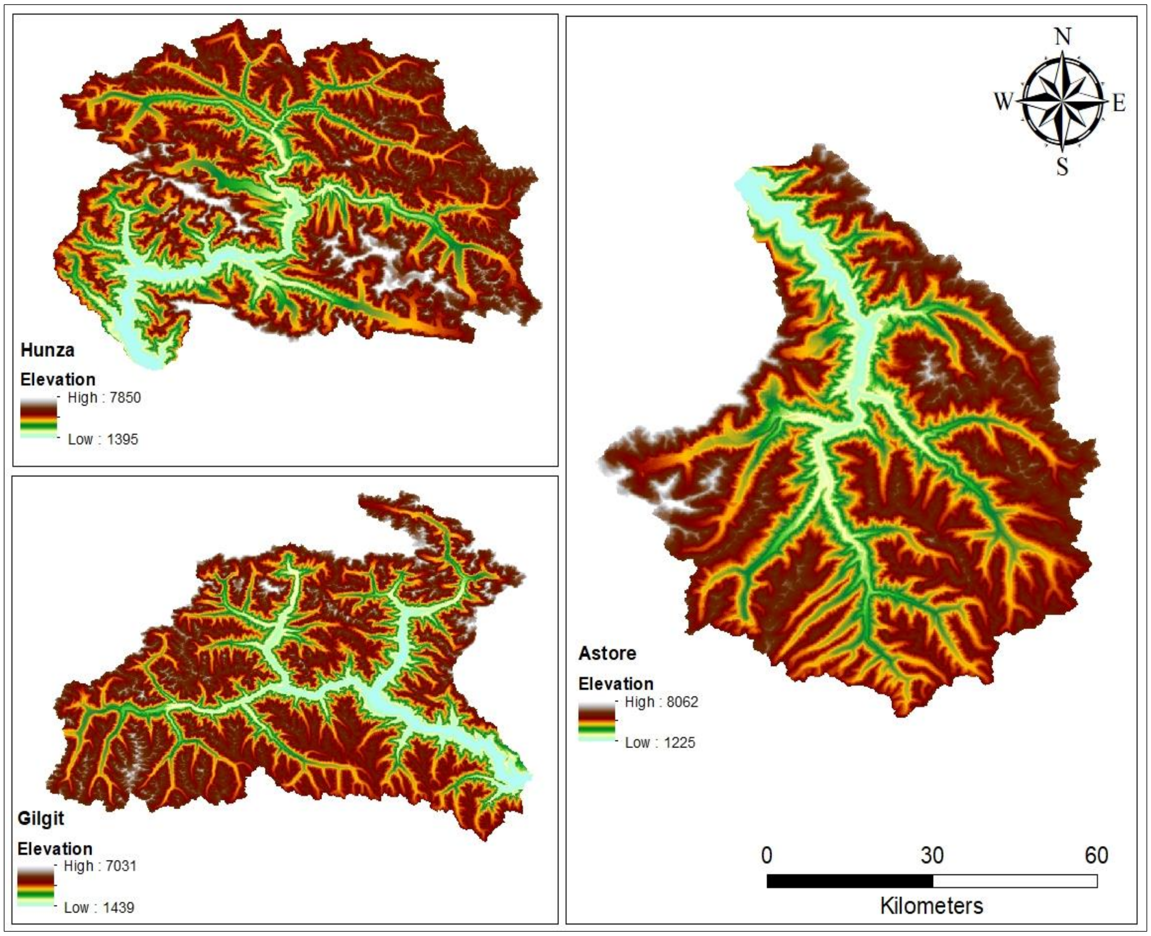


Figure 3.4 Digital Elevation Model (DEM)

These datasets provide valuable insights into the seasonal variations in meteorological conditions, which play a significant role in influencing stream flow through precipitation-runoff mechanisms.

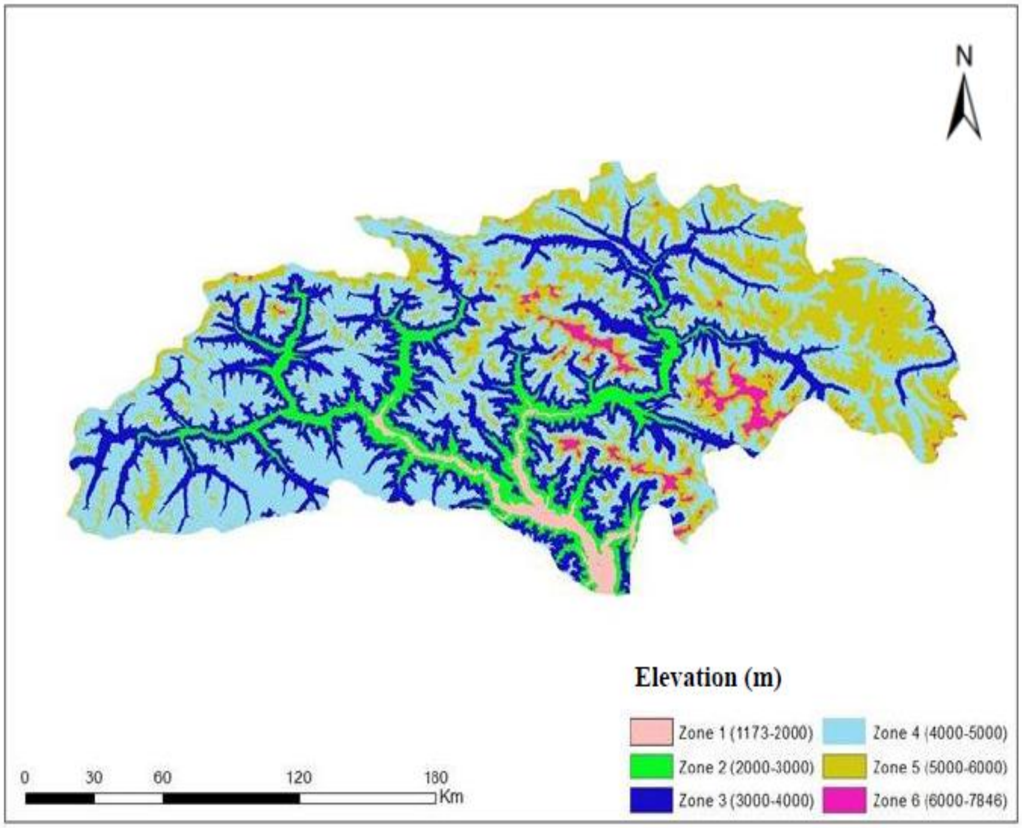


Figure 3.5 DEM showing area elevation zones.

In addition to climatic data, snow cover and glacier dynamics are crucial factors affecting stream flow in high-altitude regions like these regions. MODIS snow cover imagery is utilized to extract information on snow cover extent and glacier mass balance, offering insights into the contributions of snowmelt and glacier runoff to stream flow regimes. The integration of remote sensing data into the modeling framework enhances the spatial representation of snowmelt and glacier dynamics, contributing to more accurate stream flow predictions in data-scarce environments. Observed streamflow data, obtained from monitoring stations operated by WAPDA, serve as the ground truth for model calibration and validation. These datasets provide invaluable information on the temporal variability of the stream flow in the study area, allowing for assessment of model performance and refinement of prediction algorithms. However, ensuring the quality and consistency of observed streamflow data is paramount, requiring rigorous quality control and preprocessing procedures to identify and mitigate potential errors or inconsistencies. Preprocessing techniques such as data cleaning, outlier detection, and temporal aggregation are employed to improve reliability and accuracy of the input datasets. Missing data are interpolated using appropriate methods, while outliers are identified and either corrected or removed from the dataset to prevent them from adversely affecting model training and evaluation. Temporal aggregation is utilized to harmonize the temporal resolution of the input variables, facilitating the integration of disparate datasets into a cohesive modeling framework. Overall, the data collection and preprocessing phase of this thesis represents a critical step for ensuring the integrity and quality of input datasets used for stream flow prediction modeling in these regions. By leveraging a combination of climatic, remote sensing, and observed hydrological data, this research aims to develop robust prediction models capable of capturing the complex interactions between meteorological, glaciological, and hydrological processes driving stream flow dynamics in these scarcely gauged catchments.

CHAPTER 4

METHODOLOGY

The methodology employed in this research encompasses the development, training, and evaluation of ANN models and physical based hydrological model (SRM) for stream flow prediction in the Hunza, Gilgit and Astore rivers of Pakistan. Figure 4.1 indicates the overall methodology framework of this research work.

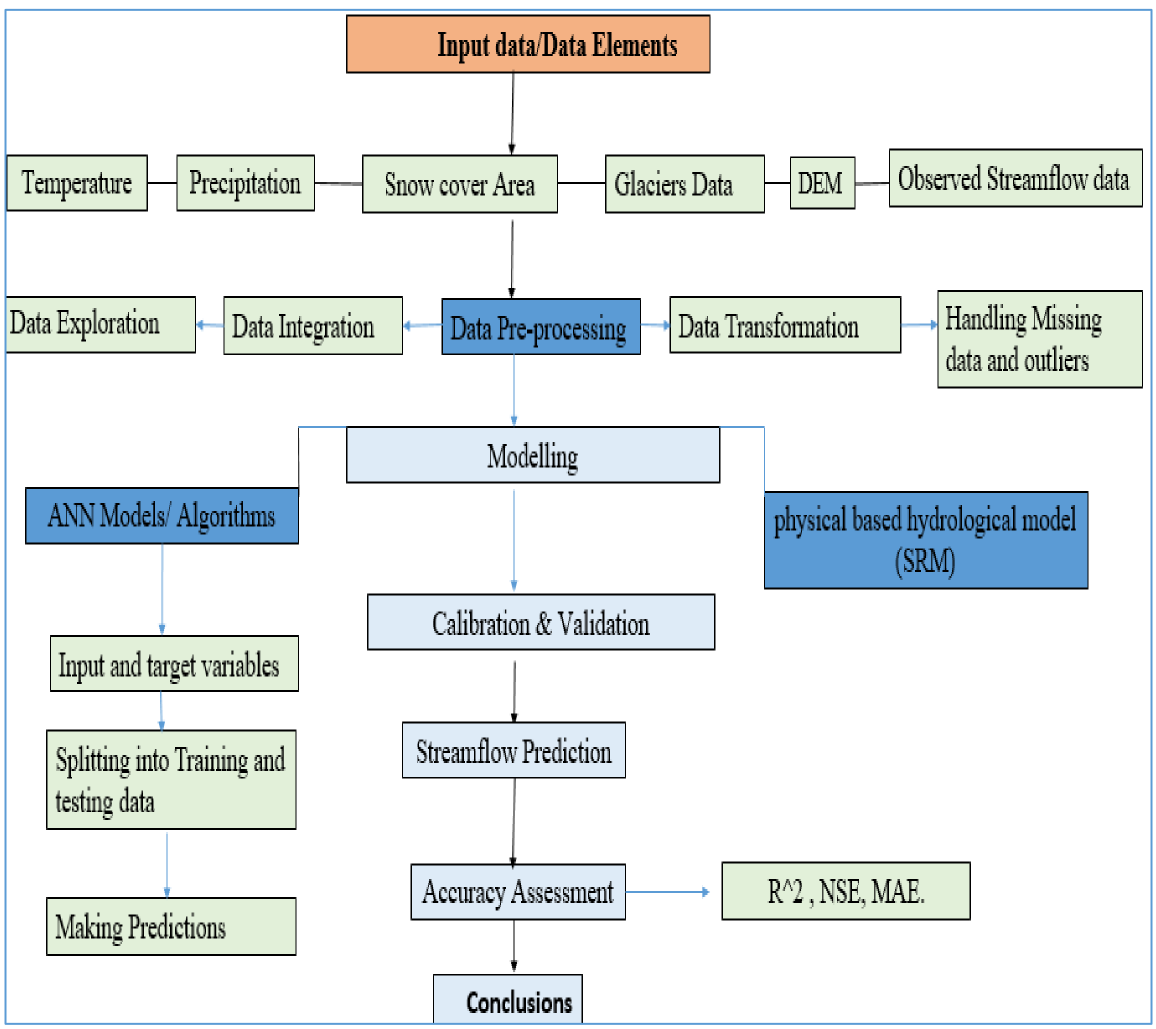


Figure 4.1 Methodology Flowchart

4.1 ANNs Architectures

The first step in the methodology involves the design of ANN model architecture, including specification of input variables, concealed layers, and the output layer. The selection of in-put variables is informed by domain knowledge and previous research findings, with climatic variables like temperature, precipitation and inflow, as well as snow cover and glacier dynamics, identified as key predictors of variability of the stream-flow in study area.

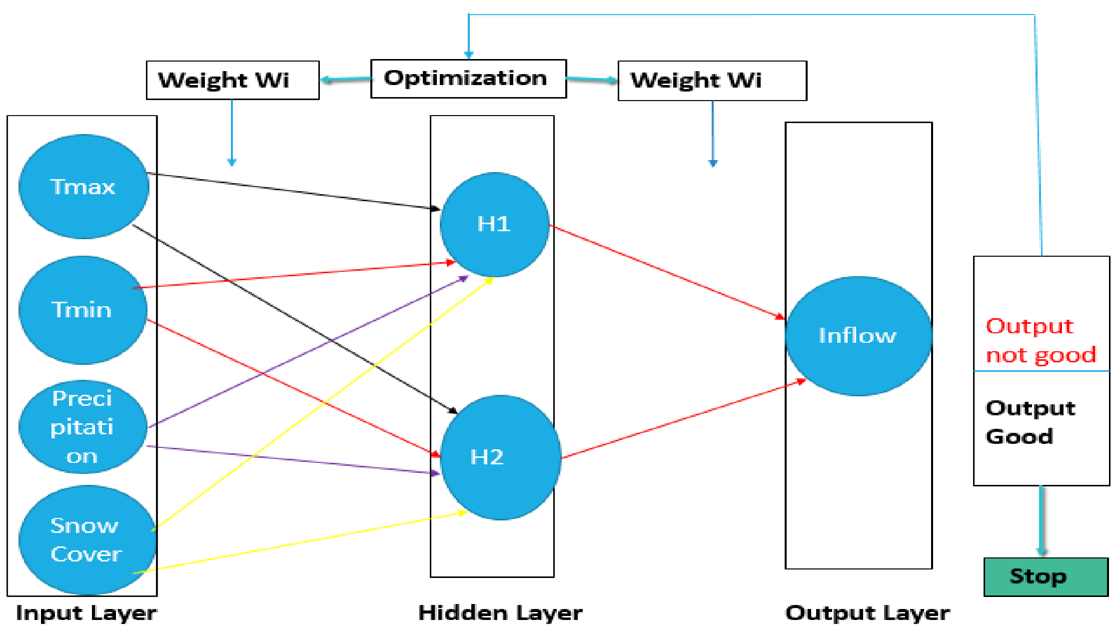


Figure 4.2 Architecture of ANNs

The number and size of hidden layers are determined through experimentation, with the goal of striking a balance between model complexity and generalization capacity. Once the ANN model architecture is defined, the next phase of the methodology entails data preprocessing and model training. Preprocessing techniques, including data cleaning, outlier detection, and temporal aggregation, are applied to ensure the quality and consistency of the input datasets. The ANN model is then trained using a portion of the available data, with optimization algorithms such as the Adam optimizer utilized to adjust model parameters and minimize prediction errors. Training data is divided into training, testing and validation to assess model performance and prevent overfitting. Finally, the methodology includes a comparative analysis of ANN models with other stream flow prediction methods, such as the Snow Melt Runoff (SRM) model. By comparing the function of ANN models with alternative approaches, the study purposes are to evaluate efficiency and the reliability of ANN models in capturing the complex hydrological processes driving stream flow in scarcely gauged mountainous catchments. Overall, the methodology provides a systematic framework for developing and evaluating ANN models for stream flow prediction, offering valuable insights into the hydrological dynamics of these regions. These are the following ANN models that are used in this research.

4.1.1 Keras sequential feedforward model, Single Layer Perceptron (SLP) Model

It is among the earliest neural networks to be introduced by Frank Rosenblatt in 1958. It is known as ANN. The basic purpose of is to compute the logical gates such as AND, OR and NOR which have binary input and output.

Perceptron functionalities are discussed below:

- It takes input from input layer.
- Weight and sum it up.
- Transmit this sum to non-linear function for producing output.

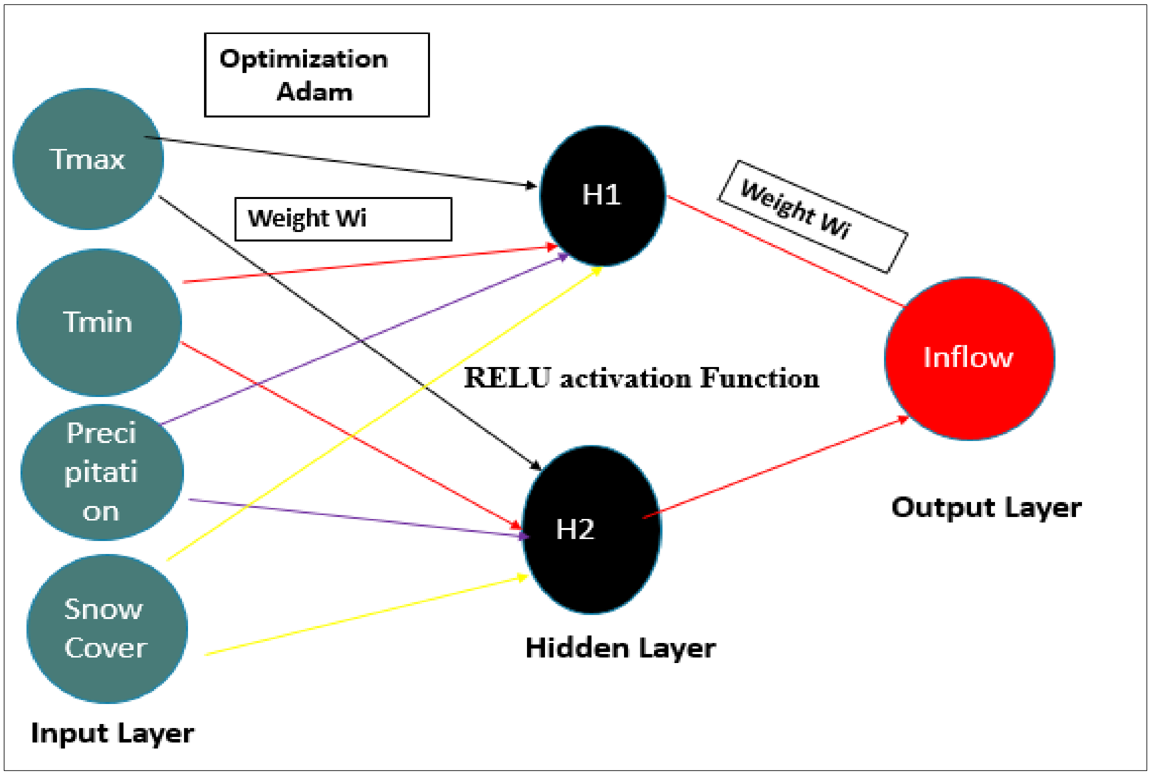


Figure 4.3 Architecture of Single Layer Perceptron (SLP) ANN Model

4.1.2 The Keras Sequential Feedforward with Backpropagation Model Multi-Layer Perceptron (MLP)

Once the input data is fed into the model it generates some results this process is called Feedforward and when these results is compared with the original values or target values this process is called Feedforward with back propagation model. The difference between the single-layer Perceptron model is that it has multiple number of layers the model has more chances to learn. The activation function that is used is RELU rectified linear unit is used that is best to use to determine the nonlinearity in the data. The optimizer used is ADAM. An ANN called an MLP is consist of multiple layers of neurons, which are the fundamental building blocks of computation. Figure 4.4 explains the architecture of multi-layer perceptron model of artificial Neural networks.

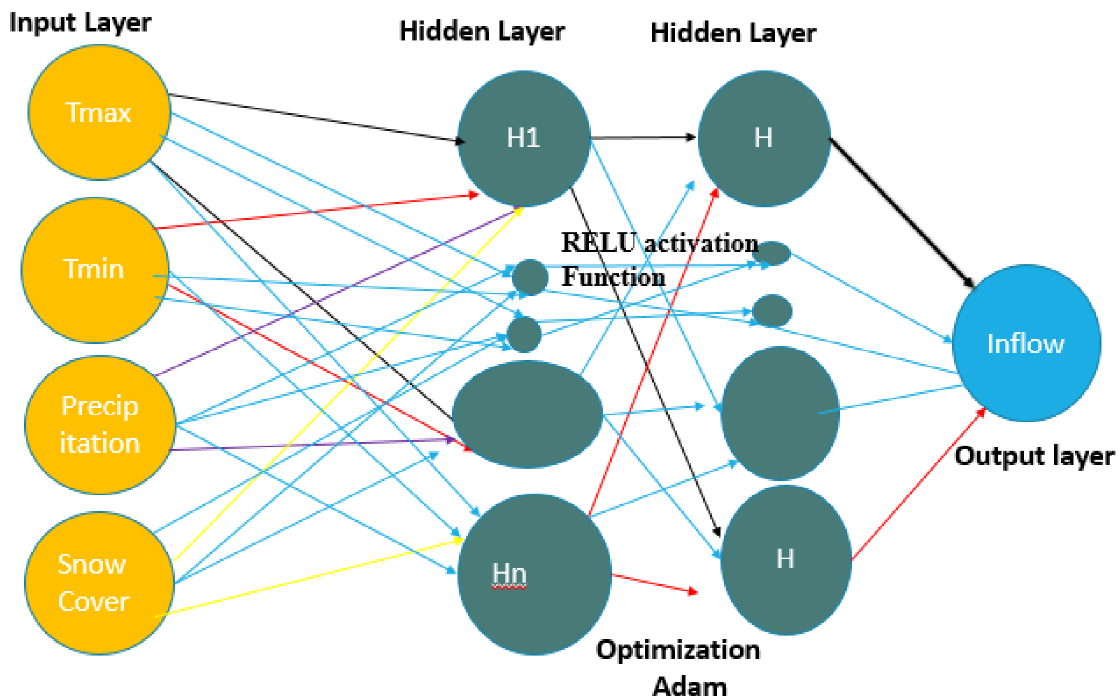


Figure 4.4 Architecture of Multi-Layer Perceptron (MLP) ANN Model

4.1.3 Radial Basis Function Network (RBF) Model

The third model of ANN is RBF network model. It Capture non-linear relationship and improve the ANN's predictive performance for time series data. For function approximation problems. It is different from other networks in 3-layer architecture, fast learning speed and universal approximation. In this Research Article we will discuss the architecture, working and use of as non-linear classifier of the Redial Basis Functions Neural Network. It has special functions, and the class consists of 3 layers input, hidden and output layer. It is different from other neural architectures, which has multiple layers and applying the non-linear activation functionalities. The first layer receives input and transmits it to the hidden layer for computation. This layer is different and the most powerful layer from the network. The last layer is used for prediction tasks such as regression and classification. Figure 4.5 displays the architecture of Radial Basis Function model of artificial Neural networks.

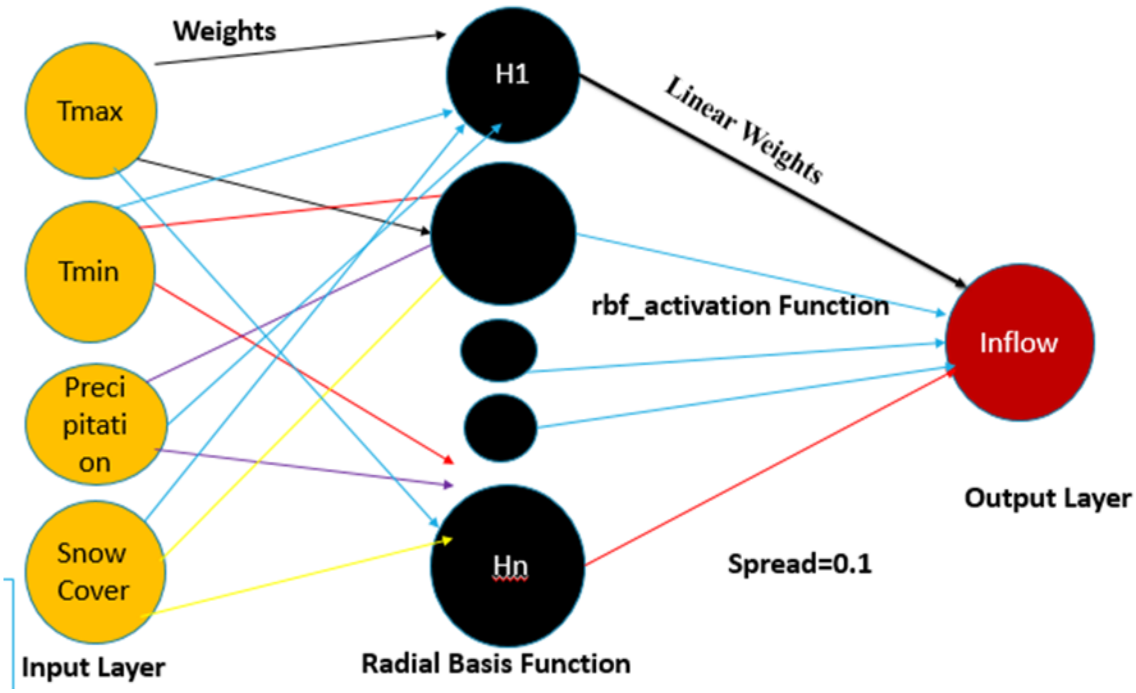


Figure 4.5 Architecture of Radial Basis Function Network (RBF) Model

4.1.4 Activation Function Rectified Linear Unit (ReLU):

It is extensively utilized in deep learning because of its straightforwardness and efficiency. When the input is positive it gets outputs otherwise result zero. It has zero gradient for negative inputs and does not differentiate at zero, It demonstrates effective performance in practical applications and assists in alleviating the vanishing gradient issue. Activation functions are essential for allowing multilayer perception (MLPs) to address intricate challenges effectively. It provides non-linearity that allows the network to adjust weights after occurring errors. In the absence of activation functions, the network would lack the capability to represent the complex relationships present within the data. The graphical representation of Rectified Linear Unit model is showing in Figure 4.6.

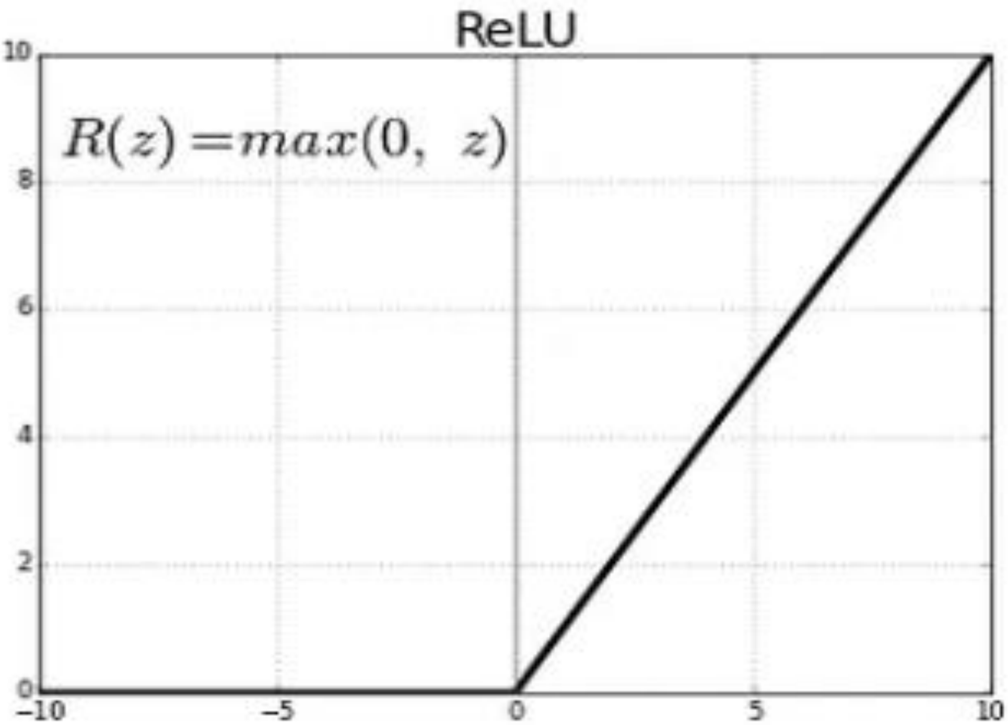


Figure 4.6 activation function ReLU.

4.2 Lose Function

Model performance is assessed across different temporal scales and spatial locations within the study area to capture the variability in stream flow dynamics. Sensitivity analysis is conducted to recognize the most significant variables and assess the robustness of the ANN models under different environmental conditions. The following lose function evaluates the performance of ANN:

Coefficient of determination (R^2), Nash Sutcliffe Efficiency (NSE)

R^2 is a measure which represents degree to which the independent variables explain the variability observed in dependent variables. The range is from 0 to1, when 1 it indicates will give a perfect model of data. The range of Nash-Sutcliffe efficiency from -∞ to 1 with 1 correspond give a match between model and the original values.

$$R^2 = 1 - \text{Unexplained Variation} \div \text{Explained variation}$$

4.2 Hydro Modelling

There are several appealing aspects to the idea of using regional hydrologic models to evaluate the effects of climatic change. First, there is no shortage of models that have been tested under various climatic and physiographic conditions, as well as models that are designed for use at different spatial scales and with different dominant process representations. This enables flexibility in determining and selecting the most suitable method to evaluate any region. Second, hydrologic models can be modified to fit the properties of the data that are currently available. The study purpose, model, and data accessibility have been the main determinants in choosing a model for a given case study among other variables. (Xu et al., 1999).

All models do, in fact, have applications in various fields. However, the simpler models, which have a smaller range of applications, can deliver adequate results at a significantly lower cost, provided that the objective function is appropriate. The more complex models, which have a wider range of applications, may be expected to deliver adequate results. Simple and physically based distributed-parameter models can be classified according to their intended uses as well as their degree of sophistication, which can range from low to high. The equivalent of selecting a suitable model is deciding when simple models can be used and when complex models must be used. (Xu et al., 1999).

4.2.1 Snowmelt Runoff Model (SRM)

To simulate and predict daily streamflow in mountain basins where snowmelt is a significant runoff factor, the Snowmelt-Runoff Model (SRM) was created. Most recently, it has been used to assess how a changing climate will affect seasonal snow cover and runoff. (Martinec, 1975) created SRM in little European basins. SRM has been used in ever-larger basins as a result of the development of satellite-based remote sensing of snow cover. (Martinec et al 1983).

Over the past decades, hydrologists have been actively exploring viable methods to model snowmelt and its influence on runoff (Pokhrel et al. 2014). As a result, two primary approaches, namely the energy-balance method and the degree-day method, have been introduced. (Zhang et al., 2014; Rashid et al. 2023). The energy-balance approach emerges as a highly comprehensive technique, offering a holistic means to model and evaluate surface flow through the intricate energy exchange among snow, soil, and air. Nonetheless, its extensive data requirements pose a limitation, making its application unfeasible in basins with inadequate data availability (Abudu et al. 2016). Degree-day base models are more practical than energy-balance models, especially in basins with little available data (Tahir et al. 2011). The most important factor in this process, according to some researchers' studies, is temperature. Models using the degree-day approach are well known for being straightforward and have been successfully used in several studies (Xu 1999; Xie et al. 2013; Nourani et al 2021).

Snowmelt runoff model (SRM) is a degree-day-based model, designed to simulate the impact of snowmelt, where it is a significant proportion of the water supply on watershed daily runoff (Martinec, 1975).

4.2.2 Structure of Model

By adding estimations of snowmelt and rainfall runoff, which are then combined with recession flow, the Snow melt Runoff Model (SRM) computes the catchment's daily runoff data, the model is described by following Equation.

$$Q_{n+1} = [Cs_n a_n (T_n + \Delta T_n) S_n + Cr_n P_n] \times \frac{A \times 10^4}{86{,}400}(1 - k_{n+1}) + Q_n K_{n+1}$$

In this equation

Q = Discharge per day measured in [m^3s-1]

C = Cs and Cr are runoff coefficients for snow and rain respectively

a = degree-day factor [cm $^{o}C^{-1}$ d^{-1}]

T = number of degree days [°Cd]

S = the ratio of the total area that is covered in snow

P = the precipitation that results in runoff [cm]

A = Total area of catchments [km^2]

K = recession coefficient

CHAPTER 5

RESULTS AND DISCUSSION

The section dedicated to consequences and discussion in this thesis a comprehensive analysis of the ANN models and physical based hydrological Snowmelt-Runoff-Model (SRM) for the stream-flow prediction in Hunza, Astore and Gilgit Valley regions of sub Indus basin of Pakistan. Drawing upon a diverse range of input datasets, including climatic variables, snow cover/glacier dynamics, and observed streamflow data, the ANN models are evaluated across different temporal scales and spatial locations to assess their effectiveness in capturing the complex hydrological processes driving stream flow variability. The first target is Astore river, Astore river has three metrological stations are the Rama station, Ratu station and Burzil station and have one-gauge station. The input data are split into seasonally (winter, monsoon and pre-monsoon) and annually (daily data).

5.1 Inflow prediction results at Astore River:

Stream flow prediction results are generated on Astore river and Table 1 presents the model result on Astore river.

Table 5.1 Station 1 Results of Astore river

	With Single Input Stations							
Models	**Annual**		**Winter**		**Monsoon**		**Pre-monsoon**	
Station 1	**R^2**	**NS**	**R^2**	**NS**	**R^2**	**NS**	**R^2**	**NS**
SLP	0.77	0.77	0.33	0.33	0.14	0.14	0.65	0.65

MLP	0.78	0.78	0.9	0.9	0.96	0.96	0.67	0.67
RBF	0.45	0.45	0.3	0.3	0.38	0.38	0.52	0.52

Table 5.1 provides the results of three different models (SLP, MLP, and RBF) used to analyse data at Station 1 along the Astore River across four distinct time periods: Annual, Winter, Monsoon, and Pre-monsoon. For each model, two performance metrics are presented: R^2 (coefficient of determination) and NS (Nash-Sutcliffe efficiency).

Table 5. 2 Station 2 Results of Astore river

Models	**Annual**		**Winter**		**Monsoon**		**Pre-monsoon**	
Station 2	$\mathbf{R^2}$	**NS**	$\mathbf{R^2}$	**NS**	$\mathbf{R^2}$	**NS**	$\mathbf{R^2}$	**NS**
SLP	0.52	0.52	0.83	0.81	0.42	0.42	0.59	0.59
MLP	0.57	0.57	0.83	0.81	0.81	0.81	0.6	0.6
RBF	0.45	0.45	0.29	0.29	0.33	0.33	0.73	0.73

Table 5.2 provides the results of three different models (SLP, MLP, and RBF) used to analyze data at Station 2 along the Astore River across four distinct time periods: Annual, Winter, Monsoon, and Pre-monsoon. For each model, two performance metrics are presented: R^2 and NS.

Table 5.3 Station 3 Results of Astore river

Models	**Annual**		**Winter**		**Monsoon**		**Pre-monsoon**	
Station 3	$\mathbf{R^2}$	**NS**	$\mathbf{R^2}$	**NS**	$\mathbf{R^2}$	**NS**	$\mathbf{R^2}$	**NS**
SLP	0.57	0.57	0.17	0.17	0.32	0.32	0.65	0.65
MLP	0.58	0.58	0.31	0.31	0.51	0.51	0.67	0.67

RBF	0.4	0.4	0.45	0.45	0.22	0.22	0.52	0.52

Table 5.3 provides the results of three different models (SLP, MLP, and RBF) used to analyze data at Station 3 along the Astore River across four distinct time periods: Annual, Winter, Monsoon, and Pre-monsoon. For each model, two performance metrics are presented: R^2 and NS. This table highlights the varying degrees of model performance depending on the time of year, with MLP showing the most consistent and robust results overall.

5.1.1 By Ensemble two stations data together

To enhance the model productivity and to improve the model accuracy apply the ensemble technique, ensemble is a technique is using multiple number of models together to improve the model accuracy and lower the error in place of using single model alone. Table 4 presents the prediction results by ensemble two stations data and their multiple models together.

Table 5.4 Prediction results of Astore river

Ensemble Two Stations Models Together		
Models	**R^2**	**NSE**
SLP	0.79	0.79
MLP	0.99	0.98
RBF	0.49	0.49

Table 5.4 presents the prediction results for the Astore River using an ensemble model that combines data from two stations. The performance of three different models SLP, MLP, and RBF is evaluated using two key performance metrics: R^2 and NSE.

5.1.2 By Ensemble all stations models together

Getting these results by ensembling all three models results that run sepretly on seasonall and anuualy data of each station. The results shows that the multi layer Perseptron model give good results. Having R^2 value that is 0.99 and NSE value is 0.98.

Table 5 presents the prediction results by ensemble all Astore river stations data and their multiple models togather.

Table 5.5 Prediction results of Astore river

Ensemble All Stations Together		
Models	**R^2**	**NSE**
SLP	0.79	0.79
MLP	0.99	0.98
RBF	0.59	0.59

Table 5.5 presents the prediction results for the Astore River using an ensemble model that combines data from all stations. The performance of three different models SLP, MLP, and RBF is evaluated using two key performance metrics: R^2 and NSE. This table highlights the varying degrees of model performance depending on the time of year, with MLP showing the most consistent and robust results overall.

5.1.3 Graphical Representation of Results

For graphical representation the scatter plot are used , the scatter plot shows the actual valuse on X-axis and the anticipated values on the Y-axis. A line represents the perfact estimate, the estimated values matche the orginal values. These scatter plots are ploted on Astore river prediction results on all three stations,that are Rama, Ratu and Burzil stations.

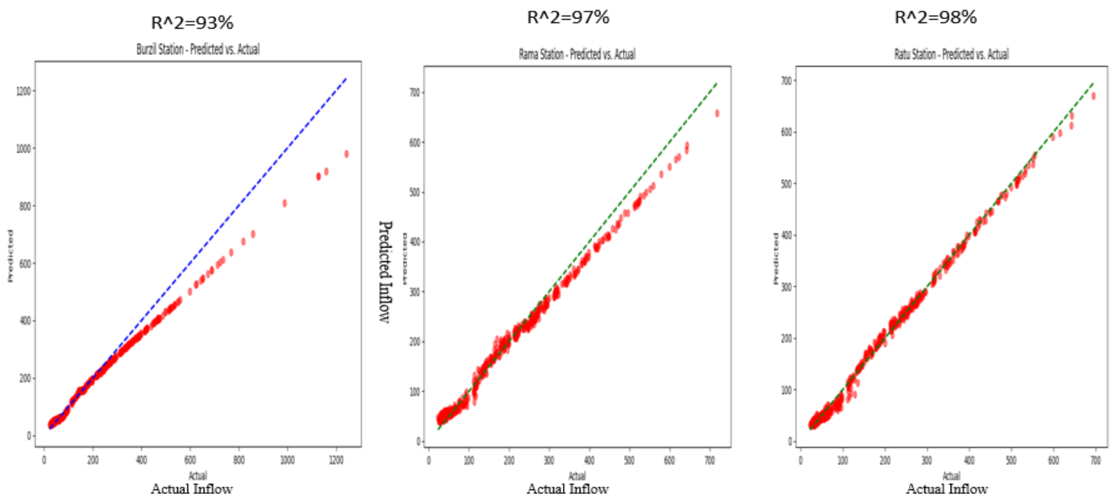


Figure 5.1 Scatter Plot on Observed and Predicted Inflow in Astore River.

Figure 5.1 presents the scatter plot on observed and predicted inflow in deferent stations of Astore river. The green line shows the actual inflow and red lines shows the predicted inflow. The plot suggests good agreement between actual and predicted values, as most points fall near the diagonal line, though some deviations may be present for larger inflow values.

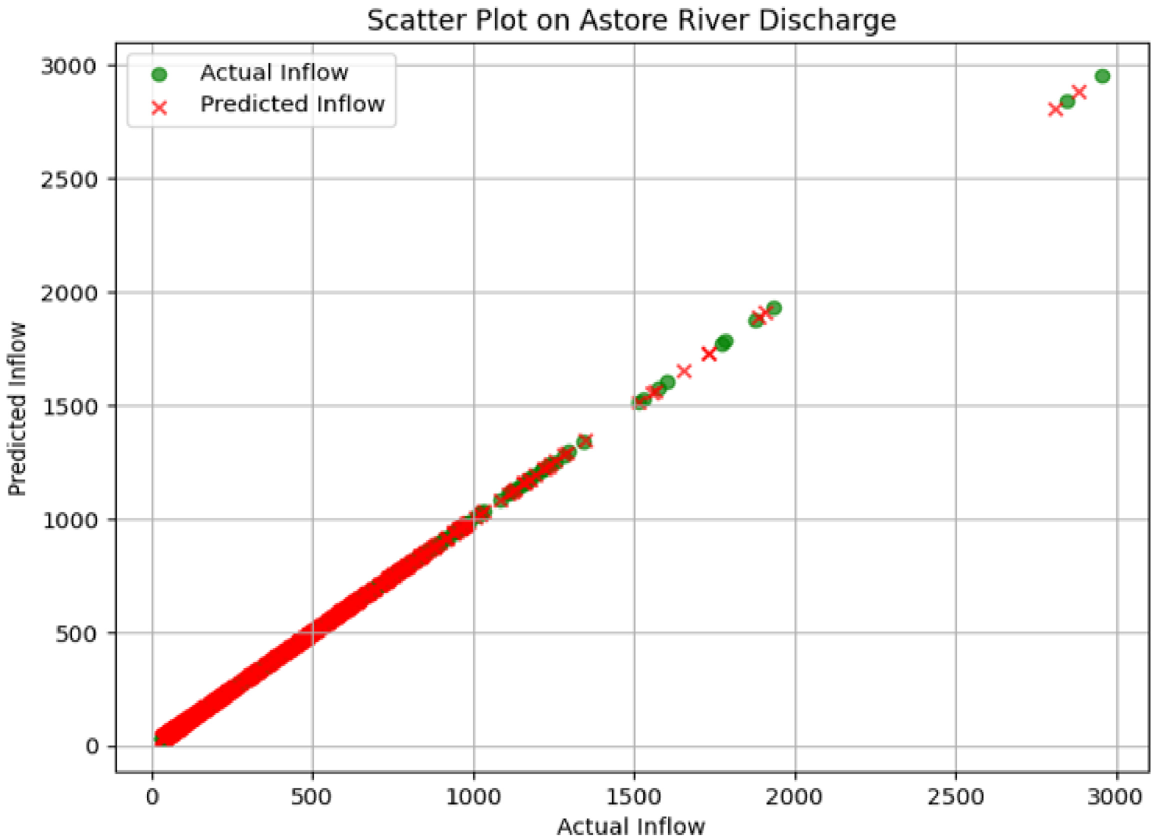


Figure 5.2 Scatter plot on Astore river all stations data

Figure 5.2 presents the scatter plot on observed and predicted inflow in Astore river. The green line shows the actual inflow and red lines shows the predicted inflow. The plot suggests good agreement between actual and predicted values, as most points fall near the diagonal line, though some deviations may be present for larger inflow values.

5.1.4 Hydrographs

5.1.4.1 Hydrograph on Burzil station Inflow data

Hydrographs are generated on the actual and predicted inflow data of all Astore river stations. A Hydrograph is a graphical representation of the flow rate or discharge over time at a specific point in the river. It's a way of displaying water level information over time. Figures 5.3,5.4,5.5 and figure 5.6 shows the Hydrograph on Burzil station ,Ratu station, Rama station and all stations ensemble results showing inflow data over time predicted vs Actual inflow.

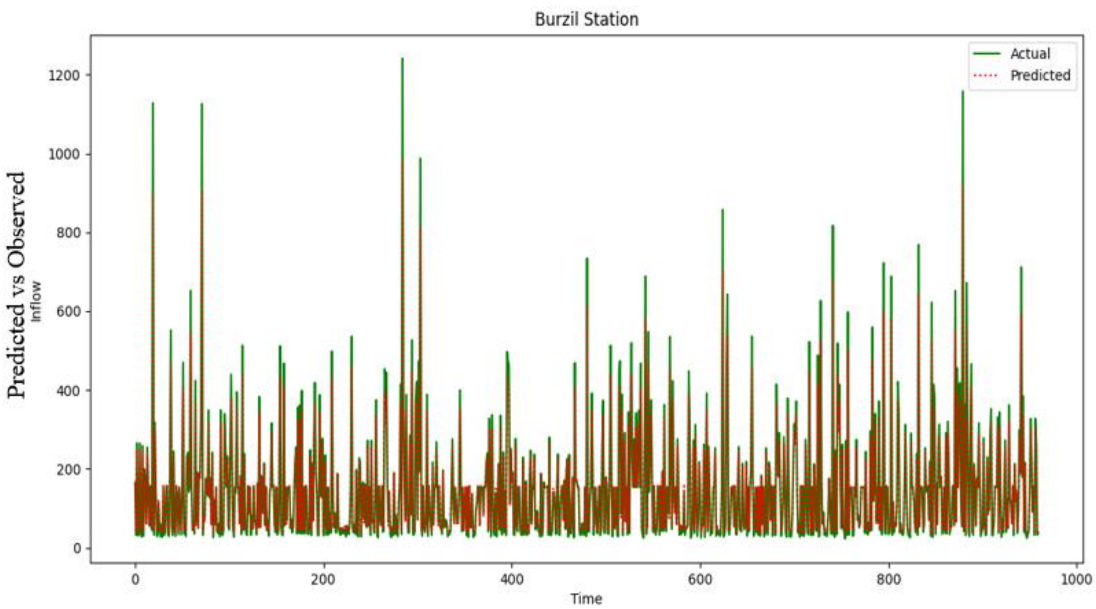


Figure 5.3 Hydrograph on Burzil station predicted vs Actual inflow

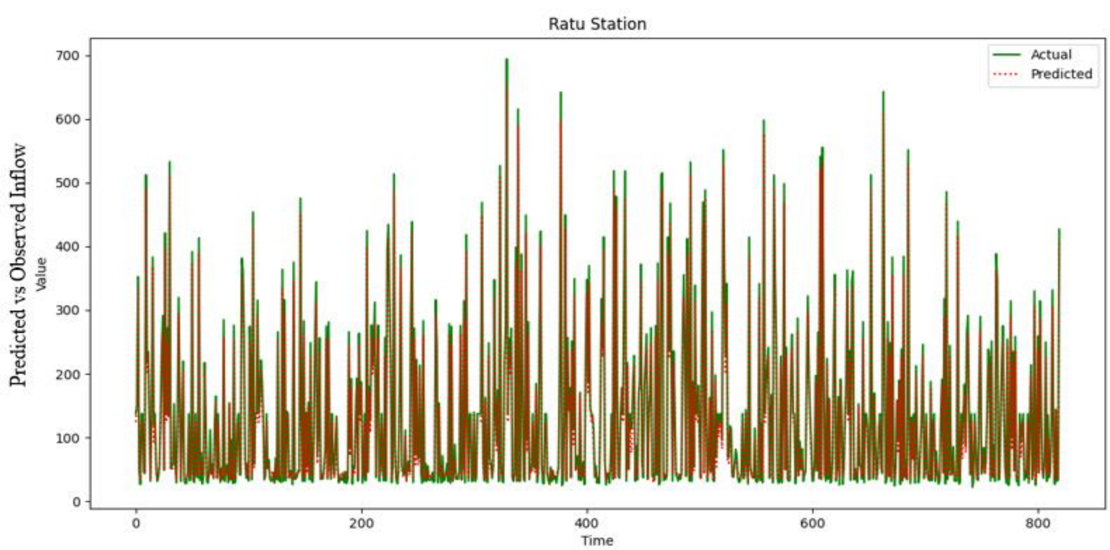


Figure 5.4 Hydrograph on Ratu station predicted vs Actual inflow

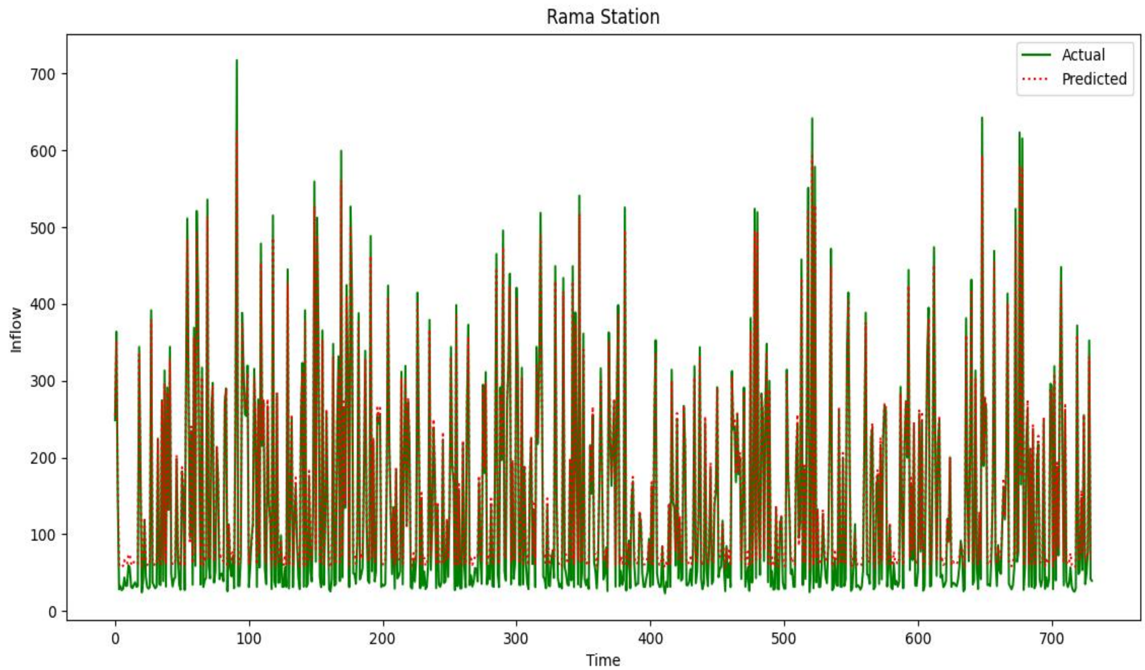


Figure 5.5 Hydrograph on Rama station predicted vs Actual inflow

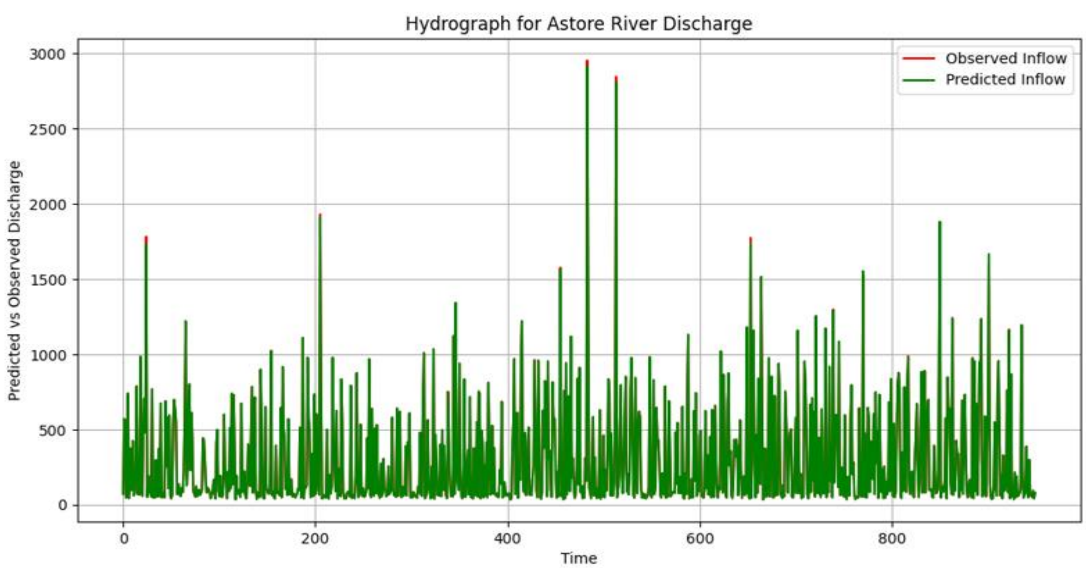


Figure 5.6 Hydrograph on Astore river predicted vs Actual inflow

5.2 Inflow Prediction on Gilgit River

Over the Gilgit River basin, these are the results generated by applying ANN models on annual and seasonal data.

Table 5.6 presents the ANN models results on Gilgit river discharge. The highest efficiency of MLP model ws observed with R^2 and NS coefficient values of 0.77 and 0.79, on annual basis. Similarly, the efficiency of MPL model found much better during different seasons.

Table 5.6 Prediction results of Gilgit River basin

	Annual		**Winter**		**Monsoon**		**Pre-monsoon**	
Models	**R^2**	**NS**	**R^2**	**NS**	**R^2**	**NS**	**R^2**	**NS**
SLP	0.69	0.53	0.55	0.55	0.45	0.45	0.65	0.65
MLP	0.77	0.79	0.81	0.81	0.96	0.96	0.67	0.67
RBF	0.45	0.45	0.63	0.3	0.38	0.38	0.52	0.52

Table 5.6 presents the performance metrics for three models (SLP, MLP, and RBF) in predicting river discharge across different seasons for the Gilgit River Basin. The metrics used are R^2 (coefficient of determination) and NS (Nash-Sutcliffe Efficiency), with higher values indicating better performance.

5.2.1 By Ensemble all models together

The Table 5.7 depicts the efficiency of MPL model based on ensemble and shows highest efficiency with R2 and NS coefficient values of 0.99.

Table 5.7 Prediction results of Gilgit river

Model	**R^2**	**NS**
MLP	0.99	0.99

5.2.2 Scatter Plot on Observed and Predicted Inflow of Gilgit River

For graphical representation the scatter plot are used , the scatter plot presents the actual valuse on the X-axis and predicted values on the Y-axis. There is a line which illustrate the perfact prediction where the predicted values are exactly matches the actual value.

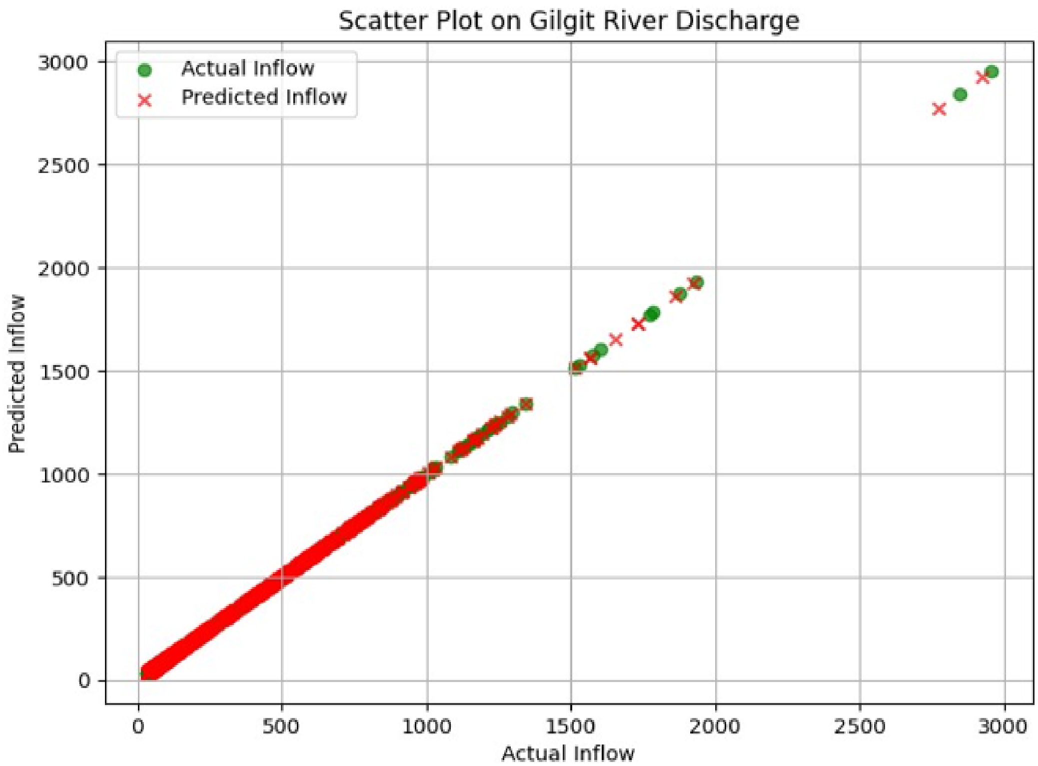


Figure 5.7 Scatter plot on Gilgit river predicted vs Actual inflow

Figure 5.7 present the scatter plot on observed and predicted inflow in Gilgit river. The green line shows the actual inflow and red lines shows the predicted inflow. The plot suggests good agreement between actual and predicted values, as most points fall near the diagonal line, though some deviations may be present for larger inflow values.

5.2.3 Hydrograph on Gilgit River Discharge

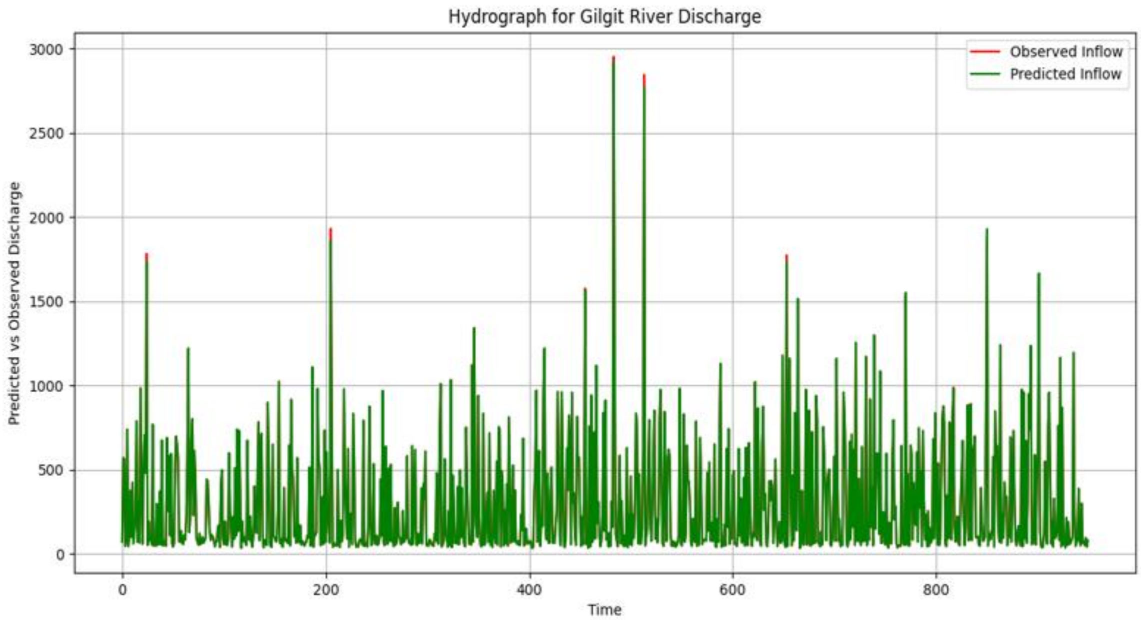


Figure 5.8 Hydrograph on Gilgit river Predicted Vs Actual inflow

Figure 5.8 present the Hydrograph on Gilgit river that represent the Predicted and Actual Inflow of different time series. The x-axis represents Time and y-axis represents Inflow. Hydrographs are generated on the actual and predicted inflow data of Gilgit river.

5.3 Inflow prediction Results of Hunza River

5.3.1 Inflow Prediction of Khunjerab Station

Table 5.8 and 5.9 present the prediction performance of three models (SLP, MLP, and RBF) for river discharge using data from a single input station on the Khunjerab River and Ziarat river. Performance metrics include R^2 and NS, where higher values indicate better predictive accuracy. The MLP model is the most effective for predicting river discharge.

Table 5.8 Khunjerab river prediction results

	With Single Input Stations							
Models	**Annual**		**Winter**		**Monsoon**		**Pre-monsoon**	
Station 1	R^2	**NS**	R^2	**NS**	R^2	**NS**	R^2	**NS**
SLP	0.67	0.67	0.63	0.63	0.74	0.74	0.65	0.65
MLP	0.78	0.78	0.91	0.91	0.92	0.92	0.67	0.67
RBF	0.45	0.45	0.53	0.53	0.38	0.38	0.82	0.72

5.3.2 Inflow Prediction of Ziarat Station

Table 5.9 Ziarat river prediction results

Models	**Annual**		**Winter**		**Monsoon**		**Pre-monsoon**	
Station 2	R^2	**NS**	R^2	**NS**	R^2	**NS**	R^2	**NS**
SLP	0.62	0.62	0.83	0.81	0.52	0.52	0.69	0.69
MLP	0.57	0.57	0.83	0.81	0.8	0.8	0.6	0.6
RBF	0.55	0.55	0.49	0.49	0.33	0.33	0.63	0.63

5.3.3 Inflow Prediction of Naltar Station

Table 5.10 summarizes the prediction performance of three models (SLP, MLP, and RBF) for the Naltar River at Station 3, evaluated over different time periods: Annual, Winter, Monsoon, and Pre-Monsoon.

Table 5.11 present the ensemble prediction results for three models (SLP, MLP, and RBF) when all stations are analysed together. R^2 Measures how well the model explains the variance in observed data. A value closer to 1 indicates better performance. The MLP model is the most effective for ensemble predictions across all stations, with exceptional accuracy and predictive skill.

Table 5.10 Naltar river prediction results

Models	**Annual**		**Winter**		**Monsoon**		**Pre-monsoon**	
Station 3	**R^2**	**NS**	**R^2**	**NS**	**R^2**	**NS**	**R^2**	**NS**
SLP	0.57	0.57	0.57	0.57	0.32	0.32	0.65	0.65
MLP	0.58	0.58	0.51	0.51	0.7	0.7	0.67	0.67
RBF	0.5	0.49	0.45	0.45	0.32	0.32	0.62	0.62

Table 5.11 Ensemble results

All Stations Together		
Models	**R^2**	**NSE**
SLP	0.79	0.79
MLP	0.99	0.98
RBF	0.49	0.49

5.3.4 Scatter Plot on Hunza River Discharge

Figure 5.9 shows the graphical representation of ANN results on Hunza River discharge through scatter plot. The green colour shows the predicted inflow and red shows actual inflow of different time series.

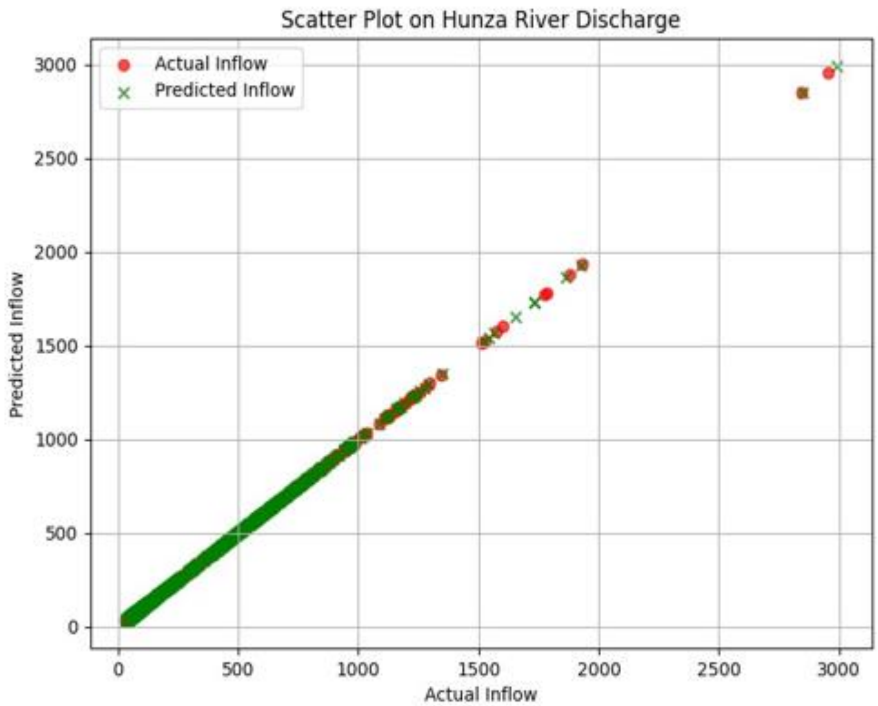


Figure 5.9 Scatter Plot on Hunza River showing actual and predicted inflow

5.3.5 Hydrograph of Hunza River Discharge

Figure 5.10 present the Hydrograph on Hunza River that represent the Predicted and Actual Inflow of different time series. The x-axis represents Time and y-axis represents Inflow. Hydrographs are generated on the actual and predicted inflow data of Hunza River. The green colour shows the observed inflow and red shows predicted inflow of different time series.

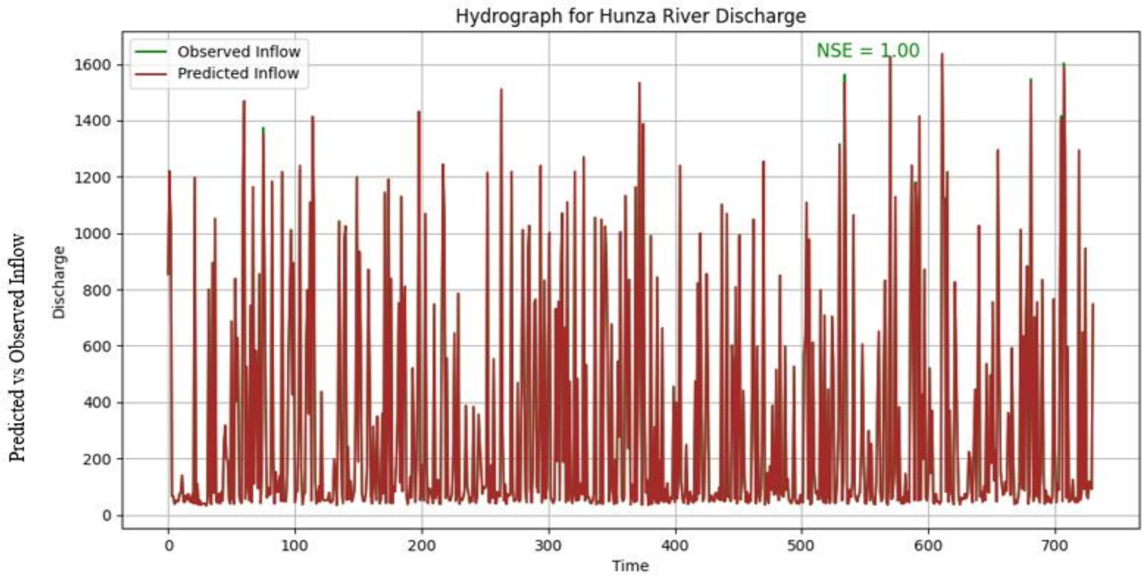


Figure 5.10 Hydrograph on Hunza River showing actual and predicted inflow

5.4 Snow Melt Runoff model

To simulate and predict daily streamflow in mountain basins where snowmelt plays a crucial role in runoff factor, the Snowmelt-Runoff Model (SRM) was created. The application of the Snow Melt Runoff (SRM) model represents a significant component of this thesis, aimed at exploring alternative approaches for stream flow prediction in the Hunza, Gilgit and Astore Valley regions of Pakistan. Furthermore, the ANN models outperform traditional hydrological models in capturing nonlinear relationships between input variables and stream flow, demonstrating their suitability for stream flow prediction in data-scarce mountainous catchments. Comparative analysis with alternative modeling approaches, such as the Snow Melt Runoff (SRM) model, further confirms the superiority of ANN models in capturing the complex interactions between meteorological, glaciological, and hydrological processes. Most recently, it has been used to assess how a changing climate will affect seasonal snow cover and runoff. (Martinec, 1975) created SRM in little European basins. SRM has been used in ever-larger basins as a result of the development of satellite-based remote sensing of snow cover. (Martinec et al., 1983) Over the past decades, hydrologists have been actively exploring viable methods to model snowmelt and its influence on runoff. As a result, two primary approaches, namely the energy balance method and the degree day method, have been introduced. Zhang et al., (2014) depicted the energy balance approach emerges as a highly comprehensive technique, offering a holistic means to model and evaluate surface flow through the intricate energy exchange among snow, soil, and air. Nonetheless, its extensive data requirements pose a limitation, making its application unfeasible in basins with inadequate data availability (Abudu et al., 2016). Degree day base models are more practical than energy-balance models, especially in basins with little available data. The most important factor in this process, according to some researchers' studies, is temperature. Models using the degree-day approach are well known for being straightforward and have been successfully used in several studies (Nourani et al., 2021).

The Snowmelt Runoff Model (SRM) is a degree-day-based model designed to simulate the effects of snowmelt on daily runoff in watersheds where snowmelt constitutes a significant portion of the water supply (Martinec, 1975). The SRM model, rooted in the physical principles governing snowmelt and runoff processes, offers a complementary perspective to the data driven approach of ANN models. By simulating the accumulation and melting of snowpack, as well as the subsequent release of meltwater into streams and rivers, the SRM model provides insights into the contributions of snowmelt to stream flow dynamics, particularly during the spring and summer months when snowmelt is prevalent.

5.4.1 Structure of the SRM model

By adding estimations of snowmelt and rainfall runoff, which are then combined with recession flow, the Snowmelt Runoff Model (SRM) computes the catchment's daily runoff data, the model is described by following Equation.

Qn+1 = [$Csnan$ (Tn+ ΔTn) Sn+ $CrnPn$]×A ×10486,400(1− kn+1)+$QnKn$+1

In this equation

Q = Discharge per day measured in [m^3s^{-1}]

C = Cs and Cr are runoff coefficients for snow and rain respectively

a = degree-day factor [cm $^{o}C^{-1}$ d^{-1}]

T = number of degree days [°Cd]

S = the ratio of the total area that is covered in snow

P = the precipitation that results in runoff [cm]

A = Total area of catchments [km^2]

K = recession coefficient

5.4.2 Analysis of Runoff Simulation

The resulting graphs show a positive correlation between simulated and observed stream flows. Its reliability as a tool for modeling and getting hydrological dynamics is strengthened by the model's ability to replicate observed runoff patterns, particularly the stated peak discharges during summer. Although acknowledged, the slight variations in runoff volumes do not diminish the model's overall effectiveness in capturing the essence of the produced runoff and its variation with the seasons. The snowmelt runoff model's output graphs provide a thorough representation of the model's effectiveness in modeling runoff during the calibration stage between 2000 and 2004. The comparison of measured and computed runoff provides an understanding of model's precision and the ability to correctly represent real-world hydrological processes in addition to serving as a visual validation of the model's effectiveness. The model performs satisfactorily across both calibration years, with only a slight discrepancy between the estimated and measured runoff volumes. Although the minor disparities in the levels of runoff may cause some concern, it's important to place these variations in context of the larger hydrological complexities.

5.4.2.1 Hunza River calibration results (2000-2004)

The snowmelt runoff model's output graphs provide a thorough representation of the model's effectiveness in modeling runoff at Hunza River during the calibration stage between 2000 and 2004. The comparison of measured and computed runoff provides an understanding of the model's precision and ability to correctly represent real-world hydrological process in addition to serving as a visual validation of the model's effectiveness. Figure 5.11 shows simulated inflow vs observed inflow during the 2000 to 2004 years' data.

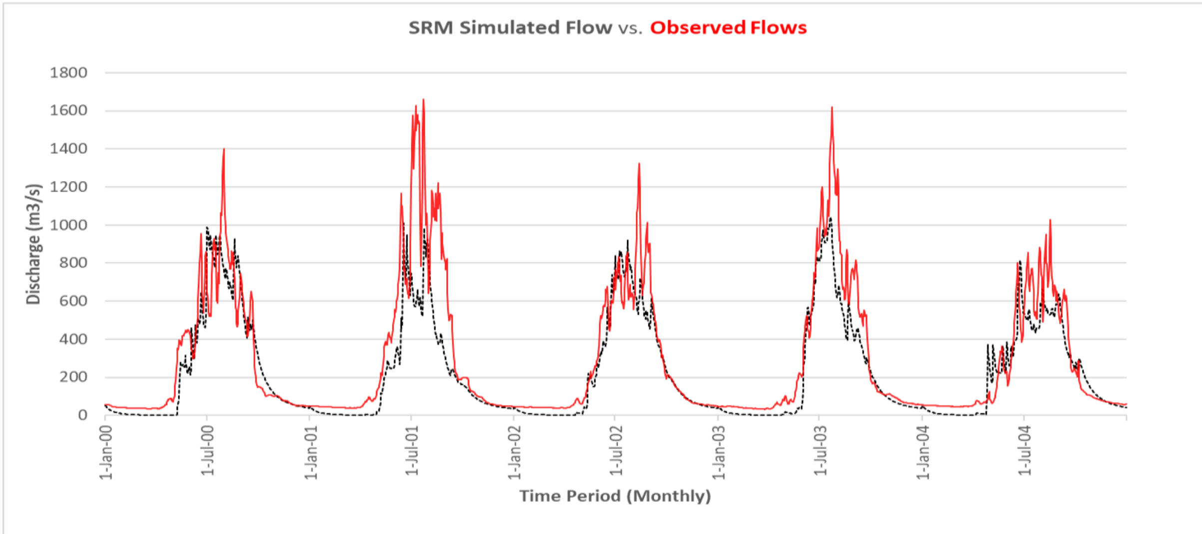


Figure 5.11 Hunza River calibration results

5.4.2.2 Hunza River validation results (2008-2012)

The snowmelt runoff model's output graphs provide a thorough representation of the model's effectiveness in modeling runoff at Hunza River during the validation stage between 2008 and 2012. Figure 5.12 shows simulated inflow vs observed inflow during the 2008 to 2012 years' data.

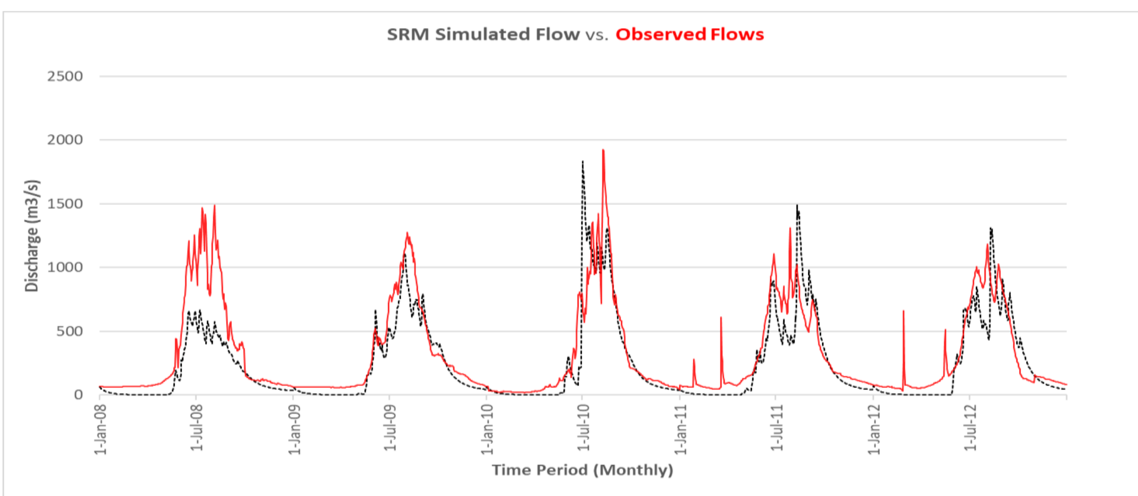


Figure 5.12 Hunza River Validation results

5.4.2.3 Model accuracy assessment

Table 5.12 presents a promising correlation between model predictions and empirical data. The calibration proficiency of the model is demonstrated by the relatively small volume differences during calibration and the close agreement between observed and simulated runoff values. The model performs satisfactorily in predicting runoff volumes across the various datasets.

Table 5.1 model accuracy parameters

Assessment Parameters	**Calibration (2000-2007)**	**Validation (2008-2012)**
R^2	0.78	0.88
NSE	0.83	0.81
DV%	33%	0.31%

5.4.2.4 Gilgit river calibration results (2000-2007)

The snowmelt runoff model's output graphs provide a thorough representation of the model's effectiveness in modeling runoff at Gilgit river during the calibration stage between 2000 and 2007. The comparison of measured and computed runoff provides an understanding of the model's precision and ability to accurately represent real-world hydrological process in addition to serving as a visual calibration of the model's effectiveness. Figure 5.13 shows simulated inflow vs observed inflow during the 2000 to 2007 years' data.

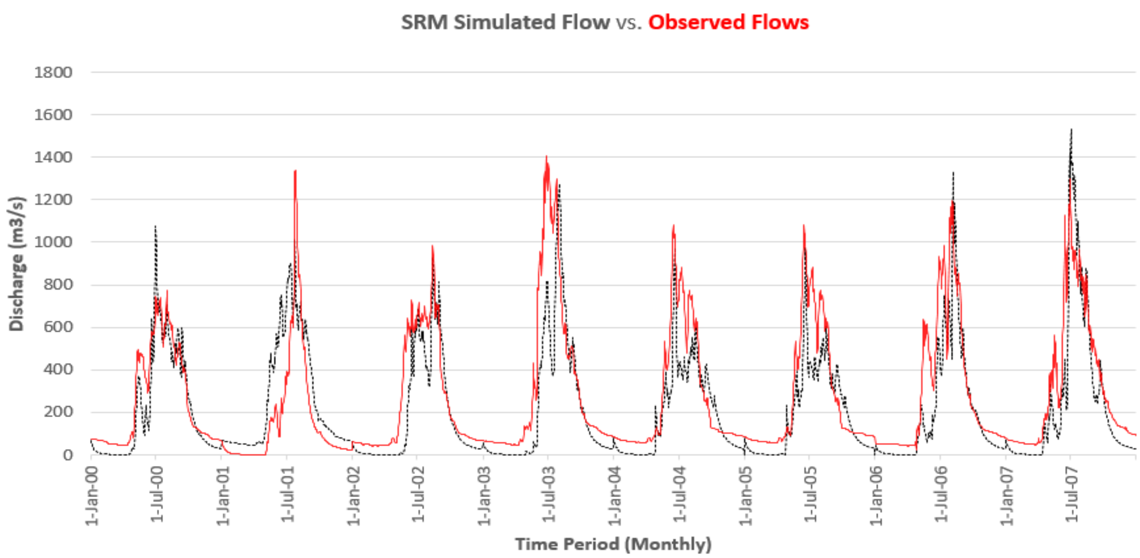


Figure 5.13 Gilgit river calibration results

5.4.2.5 Gilgit River validation results (2008-2012)

The small volume differences between computed and measured runoff 33% and 14% for calibration, and 0.31% and 0.12% for validation. show a promising correlation between model predictions and empirical data. The calibration proficiency of the model is demonstrated by the relatively small volume differences during calibration and the strong correlation between the observed and simulated runoff values. The model performs satisfactorily in predicting runoff volumes across the various datasets, even though slightly larger deviations are observed during validation. They nevertheless remain within reasonable bounds. Figure 5.14 shows simulated inflow vs observed inflow during the 2008 to 2012 years' data.

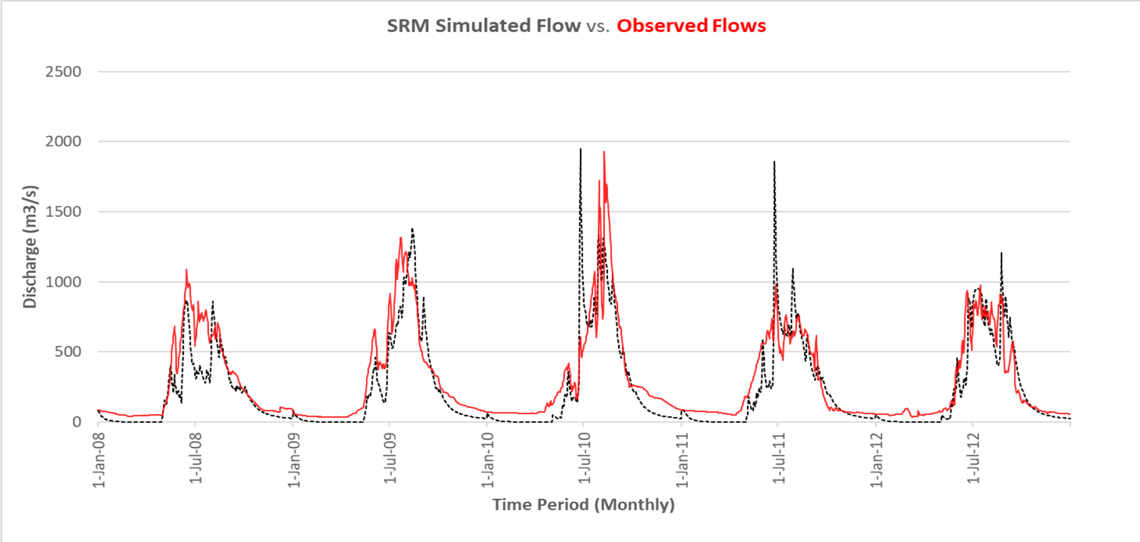


Figure 5.14 Gilgit river validatio**n** results

5.4.2.6 Model accuracy assessment

Table 5.13 presents a promising correlation between model predictions and empirical data. The calibration proficiency of the model is demonstrated by the relatively small volume differences during calibration and the strong correlation between the observed and simulated runoff values. The model performs satisfactorily in predicting runoff volumes across the various datasets.

Table 5.2 model accuracy Parameters

Assessment Parameters	**Calibration (2000-2007)**	**Validation (2008-2012)**
R^2	0.7814	0.89
NSE	0.7398	0.85
DV%	14%	0.12%

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

In conclusion, this thesis has undertaken a comprehensive investigation by using ANN into the stream flow prediction and Snow Melt Runoff (SRM) in Hunza, Gilgit and other Regions of Pakistan. Through an extensive review of literature, data collection, methodology development, and model application, this research has made significant contributions to our comprehension of hydrological dynamics in mountainous catchments with limited data availability. The research findings demonstrate the efficacy of ANN models in capturing the complex relationships between climatic variables, snow cover/glacier dynamics, and stream flow variability, achieving high predictive accuracy across different temporal scales and spatial locations within the study area. Apply some pre-processing techniques to clean the data and remove the out lairs in the data also handling some missing values in the data sets. Available climatic data and inflow data from different gauge stations are split into two ways like annual daily data and seasonal data. In seasonal data there is monsoon data pre-monsoon and winter data. Using both models one by one the data is partitioned into 80% in training the model and 20% in testing and making predictions on the 20% testing data. Then apply three models of ANNs. The first model is single layer perceptron model and second one is kerras sequential multi-layer perceptron model. The third model that is used is the Radial base function model. The consequences display multi-layer perceptron model which excites the best as giving the higher R^2 and NSE values. To improve the model performance, apply the ensemble technique, in which they used multiples models together in place of using one model alone. Moreover, the application of the SRM model has provided complementary insights into the contributions of snowmelt to stream flow dynamics, enhancing our understanding of seasonal variability and hydrological response mechanisms in high-altitude regions. The SRM model runs on study areas to predict the streamflow. The comparative analysis between ANN and SRM models has highlight the benefits and drawbacks of each technique, highlighting the importance of integrating data-driven and physically based modeling techniques for robust stream flow prediction in mountainous catchments and results shows that the ANNs models perform the best with data scarce limitation. By leveraging innovative modeling approaches and integrating diverse datasets, this research has advanced our knowledge of hydrological processes and provided valuable insights for the management of water resources in the context of climate change adaptation strategies in the study areas. However, this study has also encountered several challenges and limitations, including data scarcity, complex terrain, uncertainties in climate change projections, and model uncertainties. To overcome these obstacles, researchers, decision-makers, and local communities must work together to improve data collection, improve model representation, and develop adaptive management strategies. Furthermore, ethical considerations and stakeholder engagement are essential for ensuring the relevance, applicability, and ethical integrity of research findings in addressing water resource challenges in the study area.

In light of these findings, future research endeavors should focus on addressing the identified challenges, refining modeling techniques, and integrating interdisciplinary approaches to enhance our understanding of hydrological processes in mountainous catchments. By fostering collaboration and knowledge exchange, we can endeavor to manage water resource management and build climate resilience in the Hunza, Gilgit and

Astore Valley regions and beyond. Ultimately, this thesis contributes to the broader goal of advancing scientific knowledge and informing evidence-based decision-making to address water resource challenges in mountainous regions amidst a changing climate.

REFERENCES

1. **Abbas, M. A., & Saad. (2014).** Application of artificial neural network model for managing and forecasting water reservoir discharge. *Journal of Water Resource and Protection, 6*(12), 1147-1157.
2. **Adamowski, J., Chan, H. F., Prasher, S. O., & Sharda, V. N. (2012).** Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. *Journal of Hydroinformatics*.
3. **Adnan, M., Nabi, G., Kang, S., Zhang, G., Adnan, R. M., Anjum, M. N., & Ali, A. F. (2017).** Snowmelt runoff modelling under projected climate change patterns in the Gilgit River Basin of Northern Pakistan. *Polish Journal of Environmental Studies*, *26*(2).
4. **Alqurashi, M. (2021).** Two monthly streamflow forecasting models: Support vector machine (SVM) and artificial neural network (ANN).
5. **Azmat, M., Shafique, M., & Ahmed, I. (2015).** Estimation of water resources availability and mini-hydro productivity in high-altitude scarcely gauged watershed. *Water Resources Management, 29*(4), 1159-1173.
6. **Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., & Mekonnen, D. F. (2019).** Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia. *Journal of Hydrology*, *569*, 612-626.
7. **Ha, S., Lee, M. J., & Lee, H. (2021).** Predicting monthly streamflow using deep neural networks in the Yangtze River Basin. *Water, 13*(16), 2165. https://doi.org/10.3390/w13162165
8. **Hayat, H., & Ahmad, S. R. (2019).** Snowmelt-runoff model and RCP scenarios simulating current and future river-flows in the Karakoram and Himalayan regions of Pakistan.
9. **Martinec, J. (1975).** Snowmelt-runoff model for stream flow forecasts. *Hydrology Research, 6*(3), 145-154. https://doi.org/10.2166/nh.1975.0009

10. **Mukhopadhyay, B., & Khan, A. (2015).** A reevaluation of the snowmelt and glacial melt in river flows within Upper Indus Basin and its significance in a changing climate. *Journal of Hydrology, 527*, 119-132. https://doi.org/10.1016/j.jhydrol.2015.05.055
11. **Nayak, P. C., Sudheer, K. P., & Jain, S. K. (2007).** Rainfall–runoff modeling through hybrid intelligent system. *Water Resources Research, 43*(7). https://doi.org/10.1029/2006WR005858
12. **Pokhrel, B. K., Chevallier, P., Andréassian, V., Tahir, A. A., Arnaud, Y., Neppel, L., & Budhathoki, K. P. (2014).** Comparison of two snowmelt modelling approaches in the Dudh Koshi basin (eastern Himalayas, Nepal). *Hydrological Sciences Journal, 59*(8), 1507-1518. https://doi.org/10.1080/02626667.2014.931176
13. **Priddy, K. L., & Keller, P. E. (2005).** *Artificial Neural Networks: An Introduction* (p. 68). SPIE Press.
14. **Rashid, M. U., Ahmed, W., Islam, I., Petrounias, P., Giannakopoulou, P. P., & Koukouzas, N. (2023).** Impact of climate change on the stability of the Miacher slope, Upper Hunza, Gilgit Baltistan, Pakistan. *Climate, 11*(5), 102. https://doi.org/10.3390/cli11050102
15. **Tahir, A. A., Chevallier, P., Arnaud, Y., Neppel, L., & Ahmad, B. (2011).** Modeling snowmelt-runoff under climate scenarios in the Hunza River basin, Karakoram Range, Northern Pakistan. *Journal of Hydrology, 409*(1-2), 104-117. https://doi.org/10.1016/j.jhydrol.2011.07.022
16. **Xu, C. Y. (1999).** Climate change and hydrologic models: A review of existing gaps and recent research developments. *Water Resources Management, 13*, 369-382. https://doi.org/10.1023/A:1008117901058
17. **Yeleliere, E., Cobbina, S. J., & Duwiejuah, A. B. (2018).** Review of Ghana's water resources: The quality and management with particular focus on freshwater resources. *Applied Water Science, 8*, 1-12. https://doi.org/10.1007/s13201-018-0852-7.
18. **Zhang, G., Xie, H., Yao, T., Li, H., & Duan, S. (2014).** Quantitative water resources assessment of Qinghai Lake basin using Snowmelt Runoff Model (SRM). *Journal of Hydrology, 519*, 976-987. https://doi.org/10.1016/j.jhydrol.2014.07.028

19. **Xie, H., Zhang, G., Yao, T., Li, H., & Duan, S. (2013, December).** Quantitative water resources assessment of Qinghai Lake basin using Snowmelt Runoff Model (SRM). In *AGU Fall Meeting Abstracts* (Vol. 2013, pp. GC33B-1122).