Leveraging Remote Sensing and AI Algorithms for Predicting Inflow and Water Levels in Transboundary Reservoirs on Eastern Reivers of Pakistan

Zaheen Fatima [1], Akif Rahim [2,*]

[1] Centre for Integrated Mountain Research (CIMR), University of the Punjab, Lahore Pakistan.
[2] Irrigation Department FRAU, Government of the Punjab, Pakistan
* Correspondence: rahim@fzp.czu.cz

Abstract

This study integrates remote sensing technologies and artificial intelligence (AI) algorithms to predict inflow and water levels in transboundary reservoirs, a crucial endeavor for managing shared water resources in regions characterized by complex hydrological dynamics. Effective management of these reservoirs is essential for ensuring water and food security, mitigating natural disasters, and supporting economic stability, particularly in countries like Pakistan, which often faces uncertainties regarding reservoir levels and inflows during flood seasons from adjacent transboundary sources. The primary objective of this research is to capture variations in transboundary catchments using remote sensing and AI-based modelling approaches. The focus is on the Bhakra Dam, situated on the Sutlej River in Himachal Pradesh, India, where heavy rainfall, snow melt, and dam operations significantly influence the dam's capacity, leading to potential flooding in Pakistan. This study conducts an inflow, outflow, and storage analysis using various AI algorithms, including Linear Models (LM), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGB), Partial Least Squares (PLS), Random Forest (RF), Ridge Regression (RG), and Gaussian Processes (GP). Historical data from 2014 to 2020 regarding inflow, outflow, water levels, and storage from Indian dams was sourced from the Central Water Commission of India. Additionally, daily rainfall data from the Global Precipitation Measurement (GPM) v6 and land surface temperature (LST) and emissivity data from MODIS were utilized. A single C-band synthetic aperture radar (C-SAR) from Sentinel-1 was employed to assess surface area and extent of the dam lake. The performance of the machine learning algorithms was evaluated using the R package Hydrostat, employing a comprehensive set of statistical metrics, including Nash-Sutcliffe Efficiency (NSE), R-squared (R^2), Knowledge Graph Embedding (KGE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Residual Standard Deviation (RSR). To avoid redundancy, these metrics are reported succinctly. The Model Efficiency Index (MEI), with values approaching 1 indicating strong agreement between observed and predicted data, was used as a key indicator of model performance.

Among the tested models, Extreme Gradient Boosting (XGB) demonstrated the highest accuracy for predicting inflow, Random Forest (RF) excelled in outflow prediction, and Ridge Regression (RG) was most effective for storage estimation. The results demonstrate a high MEI, indicating the accuracy of AI algorithms for predicting water levels, storage, inflow, and outflow. Furthermore, a correlation model efficiency matrix was utilized to analyze relationships between variables such as level differences, area differences, temperature, and average rainfall, revealing a near-perfect positive correlation (0.99) between storage and level differences. This study underscores the potential

of integrating remote sensing and AI technologies to enhance water resource management in transboundary contexts.

Potential areas of research with a specific emphasis on comparing their performance against machine learning algorithms involves comparing the predictive capabilities of machine learning algorithms with those of physically-based models for reservoir inflow forecasting. Additionally, research efforts could be directed towards leveraging remote sensing data in regions with limited data availability, particularly in developing countries.

Keywords: Model Efficiency Index (MEI); Global Precipitation Measurement (GPM); Support Vector Machine(SVM); Random Forest (RF); K-Nearest Neighbors (KNN); Root Mean Squared Error (RMSE)

1. Introduction

Reservoirs play a vital role in water conservation by functioning as storage places for excess runoff water during periods of flooding. Usually, the construction of dams includes the integration of controlled outputs to fulfill the desired purpose. An outlet is a channel via which excess water is discharged. The main objective of these reservoirs is to fulfill the needs of hydropower production, domestic water supply, flood management and mitigation measures, and conservation of aquatic ecosystems. As per a study conducted by [1]. While there are numerous benefits related to food security, water supply, and electricity supply, it is crucial to recognize the presence of certain detrimental environmental repercussions. These encompass changes in river transportation, adjustments in sediment and nutrient transport, variations in water temperature, and the reduction of biodiversity based on. [2]. For Pakistan, climate change is mainly recognized for the hydrological changes in the country, affecting the inflow and water levels of transboundary reservoirs, which are, in turn, critical for water resource management and communities downstream. [3]Reservoirs and dams are built to effectively handle flood control and provide the regulated discharge of substantial amounts of water downstream, so maintaining the safety of rivers and canals. Reservoir inflow and level forecasting entails the application of meteorological and hydrological data in different models to potentially predict stream flows. Prior to a flood occurrence, it is imperative to carry out accurate forecasts of reservoir input. Precise prediction about future consumption and water levels in dams is essential management of reservoirs. [4]. Future studies should concentrate on creating comprehensive frameworks that merge sophisticated predictive tools for forecasting critical flood characteristics, such as water level peaks, flood timing, and inundation patterns, during catastrophic events like Probable Maximum Floods. By integrating hydrological forecasts with structural vulnerability assessments, including the potential for piping and overtopping failures, these frameworks can improve flood risk evaluations, enhance early, and bolster disaster preparedness in transboundary river basins. [5]

Hazard maps showing the probability of flooding are important for evaluating and managing flood risks. Several research projects have effectively used various factors such as climate data, characteristics of watershed, soil properties, and information on land use and cover obtained through remote sensing to improve flood forecasting. Combining these different sets of data helps to promote sustainable development and improve strategies for reducing the impact of flood disasters. [6].

The use of machine learning in predicting reservoir inflow has been yielding more accurate forecasts of extreme inflow events in both single-step and multi-step forecasting scenarios. While deep learning has enhanced reservoir inflow forecasting, its applications have been largely confined to real-time management and short-term planning. Although

deep learning has shown promise in short-term hydrologic forecasting, its attempts to predict long-term reservoir inflows have been less successful due to increased uncertainty and complex long-term relationships. [7]

Machine learning has far-reaching implications across various scientific disciplines due to its ability to evolve and adapt. Unlike conventional algorithms that rely solely on pre-programmed rules and lack enables it to learn and refine its performance. By leveraging a dynamic and responsive framework, machine learning significantly enhances the speed and precision of decision-making processes. [8]

In the past few decades, the application of Forecast Informed Reservoir Operation (FIRO) has been executed on Mendocino Lake, situated in the state of California. The FIRO methodology is a strategy for decision-making to improve reservoir operation and management with the consideration of hydrologic forecasting. It helps to decide about the ideal volume of water, so it gets released from the lake reservoir, making water management techniques more flexible. [9] [10].

Physical based models such as Soil and Water Assessment Tool (SWAT) [11] the Long Term Hyetograph based Hydrological Impact Assessment model at watershed scale (Watershed-Scale-L-THIA) [12] and Hydrological Simulation Program-Fortran (HSPF) [13] Hydrological Simulation Program-Fortran (HSPF) where from the reservoirs contribute inflow discharge, each include conventional methods. It should be noted that, statistical base has been incorporated into earlier datasets applied in these simulations [14]. These models require large datasets in hydro-geomorphological characteristics of the watersheds including very details about dynamics and features. Moreover, the models use calibrated parameters and changes in calibration attempts can yield different results among various models. In addition, these models require major computational resources and costs a lot. According to limited availability from transboundary reservoirs. [15].

Remote sensing data is the preferred mode of applying such an approach due to data sparse nature and climate change related transboundary reservoir operation complexities. Remote sensing refers a technological method that involves mounting sensors on airplanes or satellites [16] to view and identify various objects on Earth without physically interacting with them. It was 1960 that the technique got initiated for the first time. As per the study carried out by. [17]. This technique relies on a range of radiations, such as visible light, infrared, and electromagnetic waves, which are produced or reflected from the target in order to detect and identify distant objects on the Earth or its surface. Remote sensing data is used in many applications such as disaster assessment, environmental assessments, land and resource surveys, monitoring land use and land cover changes, tracking urban dynamics, meteorological observations, and satellite surveillance. [18]The technique uses visible, infrared, and other electromagnetic waves to detect and identify objects on earth or another planet at a distance. As far as applications are considered, Remote sensing data is used in disaster assessment, environmental impact assessments, land and resource surveys or mining etc., monitoring of land use and land cover changes (Forest, Agriculture), tracking urban dynamics (Urban growth Analysis), meteorological observations and also satellite surveillance. [19] Remote sensing-based utilization data has greatly improved our ability to study surface of the Earth features. As an example SIRPL demonstrated its high utility in analyzing inaccessible areas (e.g. military zones, severe weather conditions or transboundary area e.g. reservoirs). The sharing of remote sensing data in these localities is under the control of countries who may choose to establish restrictions for access and an amount of shared data.

Recent development in the ash field of artificial intelligence technologies makes a significant benefit for better accuracy and reliability in analysis and interpretation of remote sensing data. The problem became even more important with big, complicated &

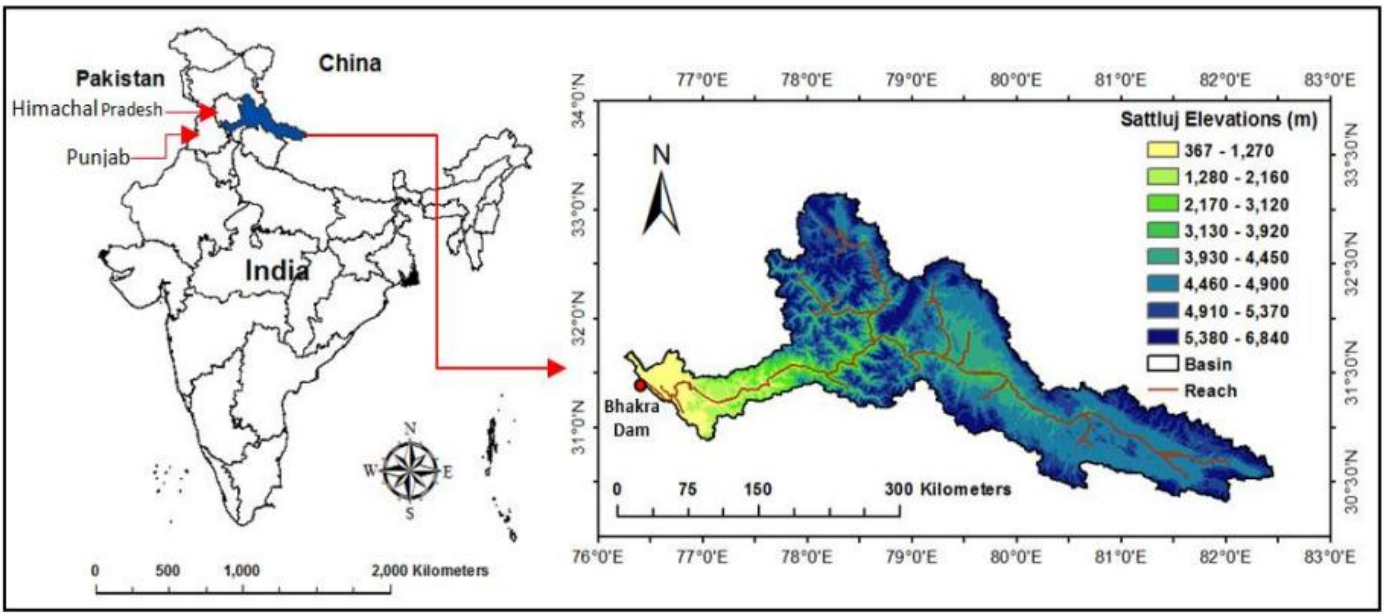
varied datasets acquired through remote sensing sources. Recently Machine Learning(ML) and Deep Learning Neural Network(ANN) Algorithms have gained prominence over the years. These algorithms have shown to successfully address the challenges and performed better that physical-based models for predicting monthly reservoir inflow for a tropical region. This is facilitated by using impulsive data-driven algorithms and ML modeling techniques that are keying from the relevant data. Machine learning algorithms are simple, lightweight, easy to implement and gradually reduce the complexity with time. [20] [21]. Several machine learning (ML) algorithms have been used in hydrological applications for water resource management observed in the study of [22] Effectively, the algorithms have established themselves with data and system analysis on a massive scale. According to [23] [24] [25]. For example, SVM was applied in the forecasting of lake water level [26] based on the daily reference modeling by climatic data to a very dry area [27]According to [28]. Artificial Neural Networks (ANN) have been utilized in various studies to estimate runoff coefficient, [29] predict river flow [30] forecast water demand under climate change, [31] calculate of wastewater temperature, [32] prognosticate groundwater levels, [33] and predict rainfall runoff retaliation and suspended sediment modeling using Genetic Programming [34], [35] Additionally, ANN has been applied to determine the optimal operation of aquifer reservoir systems, [36] groundwater modeling [37] and approximate crop yield according to. [38].

2. Study Area

Mansarowar lake situated in Tibet at an elevation of 4572 m is the point of origination for Satluj river. The river extends in north direction from 30° N to 33° N latitudes and in east direction from 76°E to 83°E longitudes. River flows in South-West direction having approximate length of 1,448 km enters India from Shipkila situated in Himachal Pradesh. From Himachal Pradesh river flow south -westerly direction and reaches into plain region of Punjab where India's highest concrete dam is located. It is straight concrete gravity dam having height 225.55m (740 ft) above the deepest foundation level. Bhakra dam height above river bed is 167.64 m (550ft). The reservoir generated by Bhakra dam is having 168.35 sq km of area and gross storage capacity of 9621 million cum. Bhakra dam live storage capacity is 7191 million cum. Total catchment area of this reservoir upstream of Bhakra dam is 56,980 k. The reservoir formed by the Bhakra Dam is Govindsagar, having length 96.56 km. Figure 1 indicates the location and DEM of Satluj River Basin. The reservoir can be depleted upto dead storage level El. 1462 (445.62 m MSL) to fulfil the demand of irrigation and power generation.

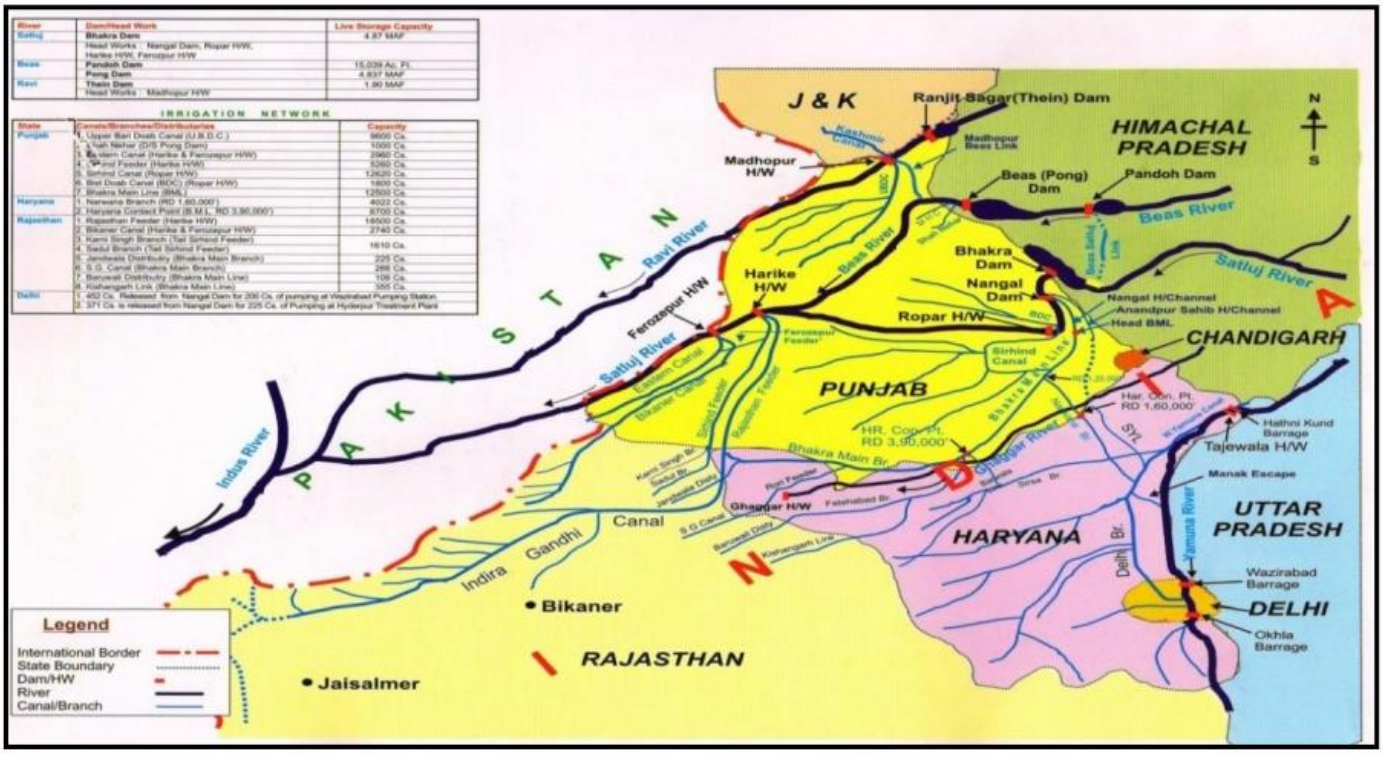
The Bhakra dam has a maximum capacity to be filled up to the reservoir level of El. 1690 (515.11 m MSL) and can be depleted down to the dead storage level of El. 1462 (445.62 m MSL). The reservoir has the ability to store the entire surplus flow of an average year, with a capacity of 6 million acre-feet. The Bhakra dam has sufficient capacity to accommodate the expected silting during the reservoir's lifespan. Figure 2 and 3 display the locations of various dams, canals, and links.

The basin exhibits intricate hydro meteorological conditions that pose challenges for accurate modeling. The basin experiences significant fluctuations in climatic conditions. The upper reaches of the Satluj Basin ranges consistently have snow, while the sub-mountainous areas at the foothills of the Satluj river valley have a sub-tropical climate.



186

Figure 1. Showing DEM location of Satluj River Basin 187



188

Figure 2. Location of Dams, head works, canals, rivers Satluj, Beas 189

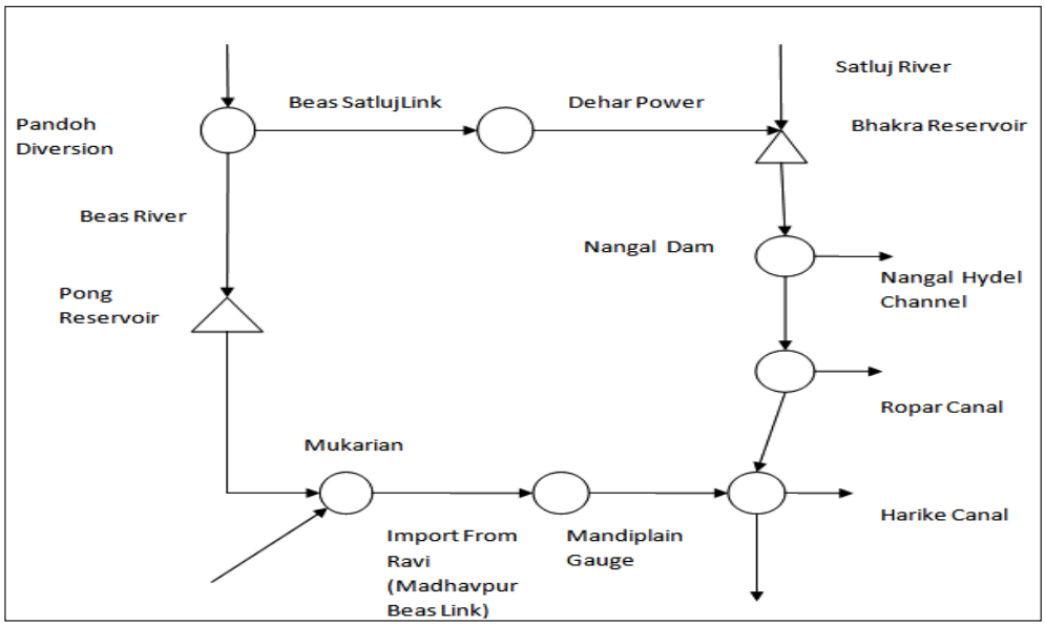


Figure 3. Schematic diagram showing Dam canal Links

2.1. *Spillways & Galleries*

There is an overflow spillway located below the top of the dam that is 79.25 meters long. Its purpose is to allow the passage of water. The flood water is regulated by four spillway gates, each measuring 15.24 meters in length and 14.48 meters in height. In addition to the spillway, there are sixteen river outlets with dimensions of 2.64 m X 2.64 m each, which are equipped with radial gates.

The outlets are situated in two tiers, with a total of eight outlets. These outlets are specifically located in the central area of the spillway. The dam will discharge extra water from the reservoir to fulfill irrigation needs. The excess water may cause flooding. The gates and river outlets have the capacity to discharge 8212 cubic meters per second of floodwater. At the crest of the dam, the height is 9.14 meters.

A road with a width of m is constructed, and galleries are built inside the dam for the purposes of drainage and inspection. The dam has galleries inside that span a length of 5 kilometers.

The following table, Table 1, provides information on the area in thousand acres and the capacity of the Bhakra Dam reservoir.

A software has been developed to calculate the volume of water in million acre-feet at different dam elevations. Conduct research that can accurately determine the level of a reservoir and the amount of storage capacity it has. This code maintains the status of a reservoir by updating it based on the input and outflow of water.

The time has been allocated to plot the elevation against storage and area curves have been plotted and are shown in Figure 4 and 5 Figure respectively.

Table 1. Bhakra Dam elevation verses area & storage capacity data

Elevation in feet	Area in thousand acres	Capacity in million acre feet	Elevation in feet	Area in thousand acres	Capacity in million acre feet
1150	0	0	1460	14.5	1.94
1180	0.21	0.002	1480	15	2.255
1200	0.65	0.011	1500	17.66	2.53
1220	1.5	0.033	1520	19.35	2.95
1240	2.35	0.073	1540	21.35	3.36

1260	3.25	0.123	1560	23.35	3.8
1280	4.43	0.203	1580	25.5	4.265
1300	5.8	0.305	1600	28.15	4.83
1320	7.18	0.435	1620	30.3	5.41
1340	8.25	0.585	1640	34.06	6.05
1360	9.33	0.755	1660	37.05	6.78
1380	10.3	0.95	1680	40.15	7.575
1400	11.25	1.17	1700	41	7.8
1420	12.18	1.4	1720	41.68	8
1440	13.35	1.665	1740	43.4	8.4

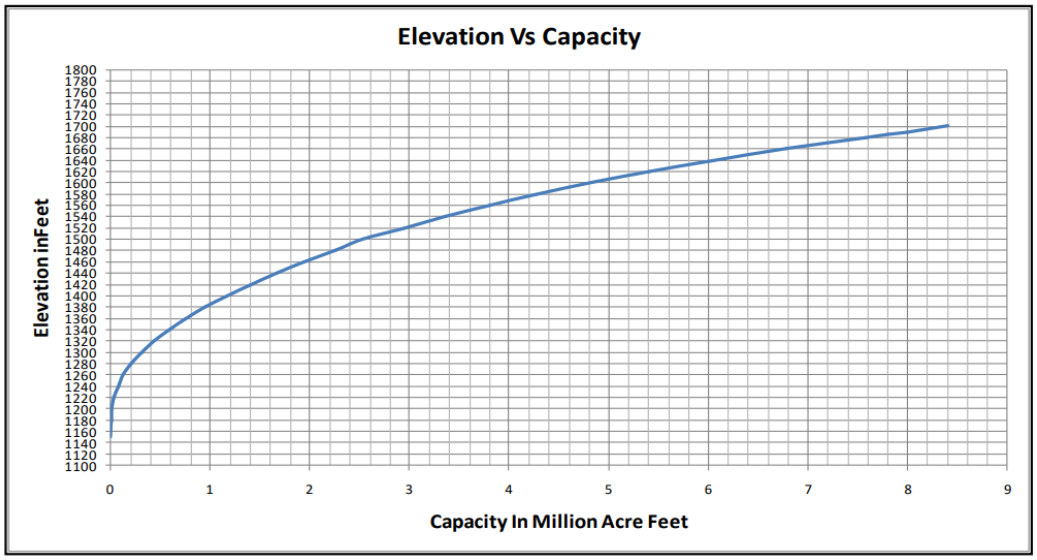


Figure 4. Curve Showing Elevation verses Capacity

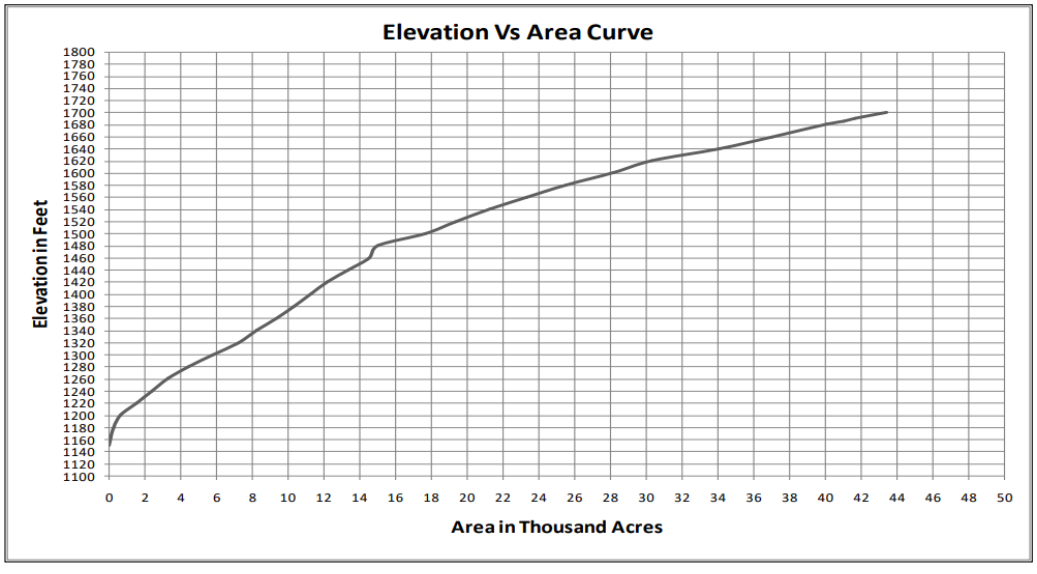


Figure 5. Curve Showing Elevation verses Area

2.2. *Nangal Dam & Nangal Hydel Channel*

The Bhakra dam is situated on the Satluj river, with the Nangal dam located 13 km downstream. The Nangal dam has a height of 29 m and consists of 26 bays, each measuring 9.14 m in length. It has the capacity to handle a flood discharge of 9910 cumecs. The dam diverts 12,500 cusecs of water to the Nangal Hydel Channel and another 12,500

cusecs to the Anandpur Sahib Hydel Channel, in order to meet the needs of irrigation and power generation. The Nangal dam serves as a reservoir to store water released from the Bhakra dam, and it also supplies water to the Nangal Hydel Channel and Anandpur Sahib Hydel Channel. The Nangal Hydel Channel, which is 61.06 km long, begins at the left bank of the Satluj river from the Nangal Dam. After power generation, water is released from the Ganguwal and Kotla power generating units on the Nangal Hydel Channel for irrigation purposes.

2.3. *Beas Satluj link*

The Beas-Satluj connection canal connects the Beas river to the Satluj river. Pandoh is a diversion dam constructed on the Beas River, located 21 km upstream of Mandi in the state of Himachal Pradesh. The dam is classified as an earth-cum-rock fill dam and has a height of 76.20 meters. The structure consists of five bays equipped with radial gates, which are used to control the water flow. The water from the Beas River is redirected to the Satluj River through a tunnel with a diameter of 7.62 meters and a length of 13.1 kilometres. This tunnel, which is lined with concrete, extends to Baggi and has a capacity to transport a discharge of 254.85 cubic meters per second. An 11.8 km concrete lined open canal is used to transport the Beas water from Baggi to Sundernagar for the purpose of balancing reservoir. The discharge capacity of the Sundernagar Hydel Channel is 240.7 cubic meters per second. The channel discharges into a balancing reservoir with a storage capacity of 370 hectare meters. The balancing reservoir ensures the necessary supply to fulfil the current demand at the Dehar power plant. The reservoir has a length of 2130 meters and a breadth of 449.88 meters. The water from the balancing reservoir is once again conveyed through the Sundernagar Satluj tunnel, which has a diameter of 8.53 meters and a length of 12.35 km. The tunnel, which is built with concrete, originates from the Sundernagar Balancing Reservoir and discharges its water into the Surge Shaft. From there, three penstocks transport the water to the Dehar power plant. The tunnel has a capacity to transport 403.53 cubic meters per second of water. A surge shaft with a diameter of 22.86 m and a height of 125 m has been installed to manage abrupt surges in the penstock caused by a rapid decrease in power demand or a complete shutdown of the producing units. The Dehar power station is situated on the right bank of the Satluj River, upstream of the Bhakra dam. This concept involves harnessing the potential energy of falling water from a vertical distance of 320 meters to generate power. The Dehar power station consists of six generating units, each with a capacity of 165 MW. The water discharged from Dehar is used to enhance the electricity capacity of Bhakra.

Table 2. Characteristics of Bhakra Nangal Project

Dam Type	Straight Concrete Gravity Dam
Height of dam above deepest foundation	225.55 m (740 ft)
Height of dam above river bed	167.64 m (550 ft)
Length at top of the dam	518.16 m
Width at top of the dam	9.14m (30ft)
Length at bottom of the dam	99.06 m (325ft)
Base width of dam	190.5m (625 ft)
Catchment Area upstream of dam	56,980 sq km
Maximum Reservoir Elevation	1680 ft
Spillway crest level	El. 1645.211 ft
Dead storage level of dam	1462 ft
Gross reservoir capacity at El. 1685 ft.	9621 million cum.
Live storage capacity at El. 1685 ft.	7191 million cum.
No. of Penstocks	10

3. Materials & Methods

Historic hydrological data: (2014-2020) of Inflow, Outflow, Level, and Storage of Indian dams collected from Central Water Commission of India.

3.1. *Rainfall in Satluj River Basin at Bhakra Region*

The bar chart in Figure 6 displays the precipitation levels in this area from 1983 to 2018. The mean annual precipitation of this area from 1983 to 2018 is 1120 mm. The highest amount of rainfall occurs annually from June to September. The mean precipitation accumulated over this time period is 810 mm. The remaining portion of the year is expected to have an average rainfall of 310 mm. Figure 6 illustrates a declining pattern in precipitation from 1983 to 2018. The rate of decrease in rainfall from June to September is 2.77 mm/year, whereas for the rest of the year it is 2.57 mm/year. There is a noticeable decline in rainfall in the Satluj river basin in Bhakra Dam Region. The bar chart depicted in Figure 6 illustrates the amount of rainfall in this region from 1983 to 2018. The average annual precipitation in this region from 1983 to 2018 is 1120 mm. More precisely, this precipitation takes place between the months of June and September, and it represents the highest yearly rainfall amount. The mean precipitation received during this time frame is 810. Figure 6 indicates that the remaining part of the year is projected to have an average rainfall of 310 mm.

Displays the precipitation levels in this area from 1983 to 2018. The mean annual precipitation in this area during the period from 1983 to 2018 is 1120 mm. Specifically, this rainfall occurs from June to September, the largest amount of rainfall occurs annually. The average amount of rainfall received during this period is 810. The remaining portion of the year is expected to get an average rainfall of 310 mm, as shown in Figure 6. From 1983 to 2018, there has been a noticeable decline in rainfall. This tendency is particularly evident during the month of June. In September, there is a decrease of 2.77 mm per year, while for the rest of the year. Additionally, the rainfall trend is seeing a decline at a rate of 2.57 mm/year. This decline in rainfall is evident. Recorded in the Satluj river basin in the Bhakra Dam Region.

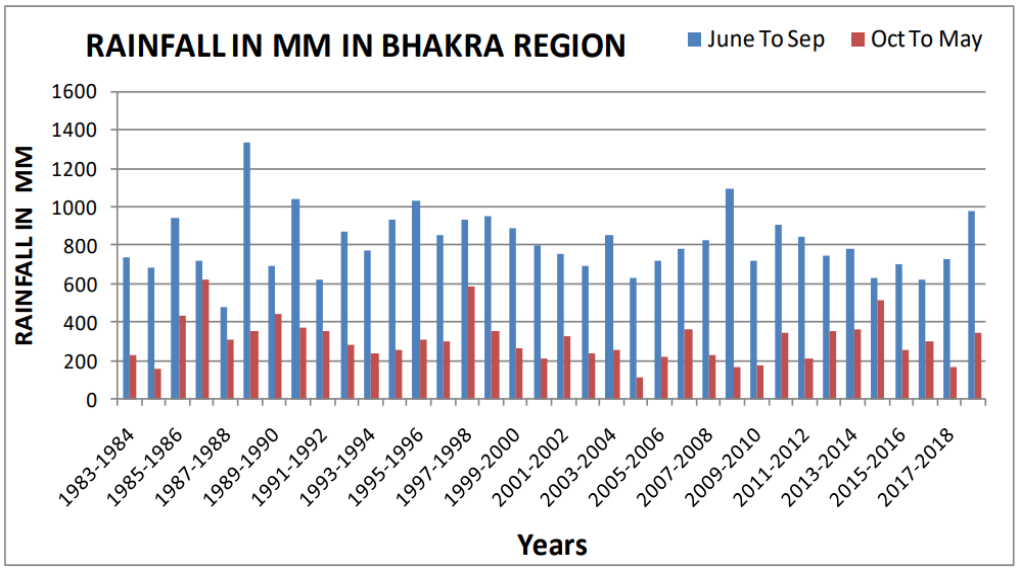


Figure 6. Bar chart showing rainfall from duration 1983 - 2018

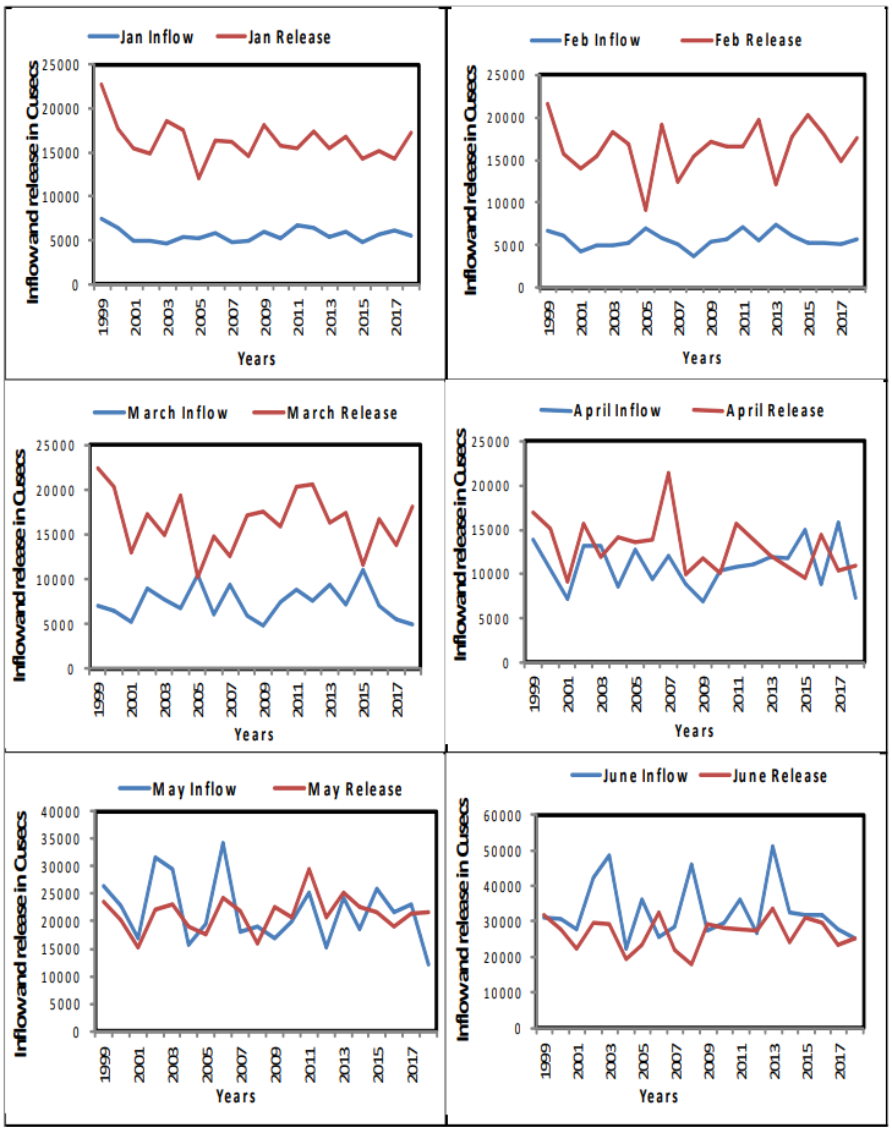


Figure 7. Graph showing inflow and release from June-December for duration 1999 – 2018 at Bhakra

3.2. *Monthly Inflow and Release for Duration 1999 -2018*

The monthly input and output of water from the Bhakra dam throughout the period from 1999 to 2018 may be found in Table 3. The data shown in Table 3 clearly shows that the inflow fluctuates significantly throughout different months, while the release is contingent upon factors such as demand, storage availability, and inflow. This information can be seen in Table 3. The average inflow throughout the month of January has remained consistent over the previous 20 years, with just minor variations. There is a downward trend in the data, with larger releases during this time period and an average release of 16268 cusecs according to the requested requirements.

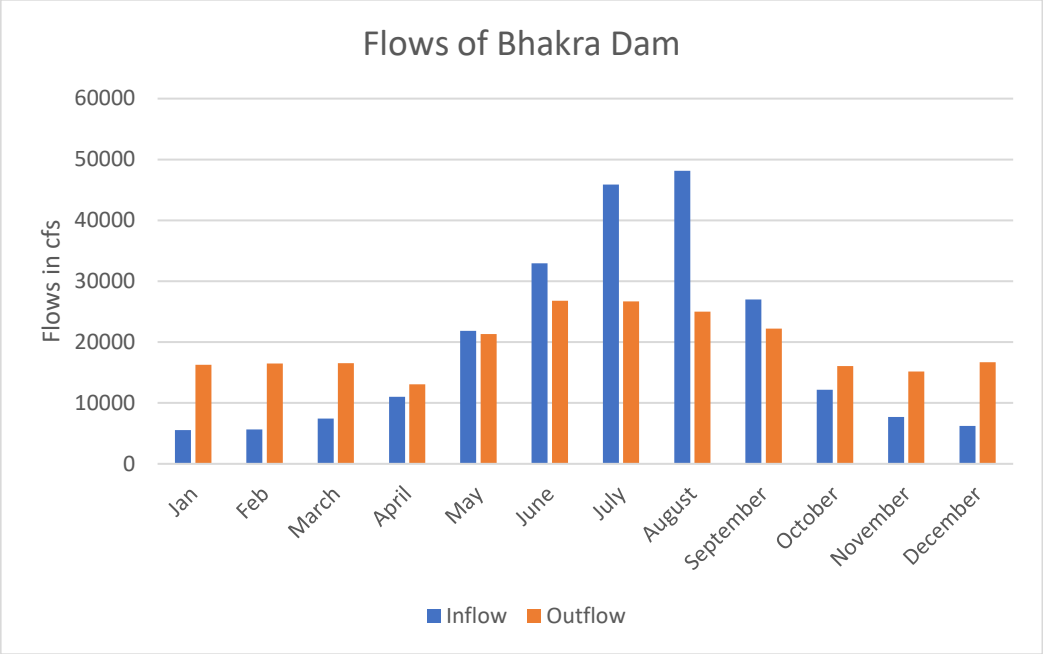


Figure 8. Average monthly Inflow and Release for Duration 1999-2018

The data from the past 20 years indicates a slight increase in inflow during the months of February, March, and April, but this increase is not significant. In January, February, and March, the average inflow is 41 cusecs, which is considerably low compared to the monthly release during these months. However, in April, the average inflow significantly increases to 10994 cusecs, with a release of 13083 cusecs. In May and June, the trend shows a decrease in inflow, although the inflow is still higher than the discharge. In July, August, and September, the trend of inflow is positive and exceeds the release during these months. In October, November, and December, the trend remains positive, but the inflow decreases considerably compared to the release during these months.

3.3. *Maximum and Minimum Reservoir Level for the Duration 1975-2019*

The shown data represents the highest and lowest annual reservoir levels in feet from 1975 to 2019. Figure 9 illustrates that the reservoir's maximum level is limited at an elevation of 1680 feet. The water level of El. 1462 ft has been reduced to the point where it is almost empty. This has been done to safeguard the area downstream and minimize any negative impacts. The reservoir level can be increased up to an elevation of 1690 feet in collaboration with Chairman BBMB. According to Figure 9, the water level consistently exceeds the storage level El.1680 ft annually. From 1975 onwards, there has been no notable alteration in the maximum or lowest values, as observed annually with a value of 3.9. from 2019 forward, except for a few years. Figure 9 displays the annual minimum and maximum levels of the reservoir.

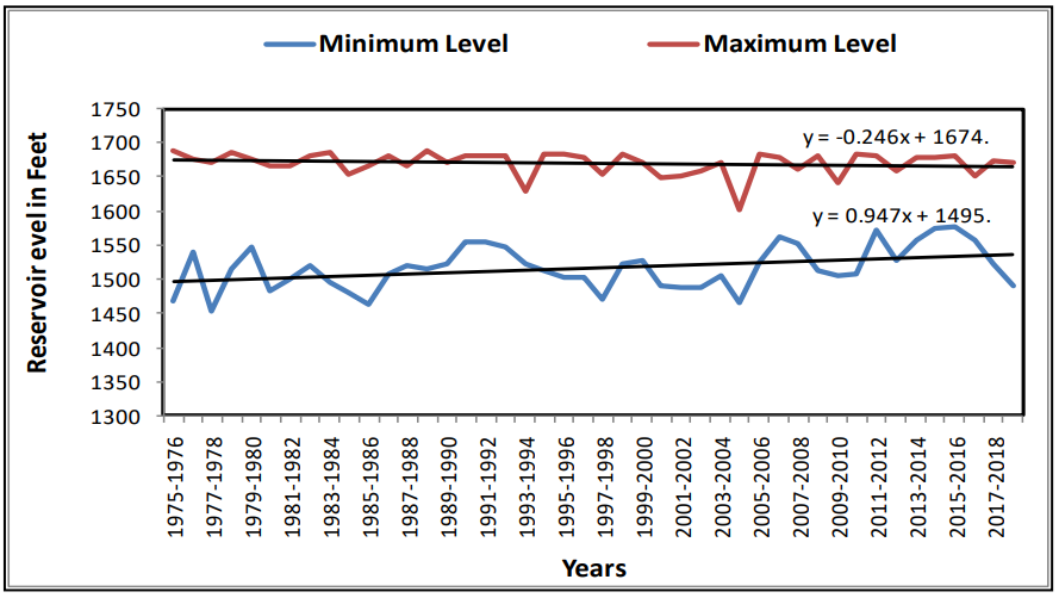


Figure 9. Reservoir yearly Minimum and Maximum level

3.4. *Climatic Data*

3.4.1. Late Daily Rain Data of GPM

The Global Precipitation Measurement (GPM) V6, available on Google Earth Engine (GEE), is a satellite mission conducted by NASA and the Japanese Aerospace Exploration Agency (JAXA). Its purpose is to collect advanced observations of rain and snow on a global scale every three hours. Launched on February 27th, 2014, the GPM Core Observatory satellite carries state-of-the-art instruments that have set a new standard for measuring precipitation from space. The data provided by GPM is used to consolidate precipitation measurements from various partner satellites across the globe, enabling the quantification of rainfall and snowfall patterns worldwide.

The late daily data is broadcasted daily, but its findings are only revealed after undergoing error removal, which typically takes 2-3 days. We rely on this corrected late daily production of GPM for our research. In contrast, the early daily data contains errors, which is why we decided to use the late daily data of GPM in our study.

3.4.2. MODIS Land Surface Temperature and Emissivity (MOD11)

The Land Surface Temperature (LST) and Emissivity daily data are obtained using two different algorithms. The generalized split-window algorithm retrieves data at a resolution of 1km pixels, while the day/night algorithm retrieves data at a resolution of 6km grids. In the split-window algorithm, emissivities in bands 31 and 32 are estimated based on factors such as land cover types, atmospheric column water vapor, and lower boundary air surface temperature. These factors are divided into manageable sub-ranges to optimize the retrieval process. On the other hand, the day/night algorithm retrieves daytime and nighttime LSTs and surface emissivities by analyzing pairs of day and night MODIS observations in seven TIR bands. The resulting product includes LSTs, quality assessment, observation time, view angles, and emissivities.

3.4.3. SAR Process & Analysis (Synthetic Aperture Radar)

The Synthetic Aperture Radar (SAR) on board the Sentinel-1A/B satellites, operated by the European Space Agency (ESA), offers continuous imaging in all weather conditions and at any time of day or night. It operates at C-band frequencies and has four distinct imaging modes, each with its own spatial resolution and coverage. SAR is an active data technology that emits its own energy and measures the amount of energy reflected back.

One advantage of SAR is its ability to operate at wavelengths that are not affected by cloud cover or lack of illumination, allowing it to collect data under any weather conditions, day or night. The following figures display the results of SAR analysis conducted on a randomly selected area. [39]

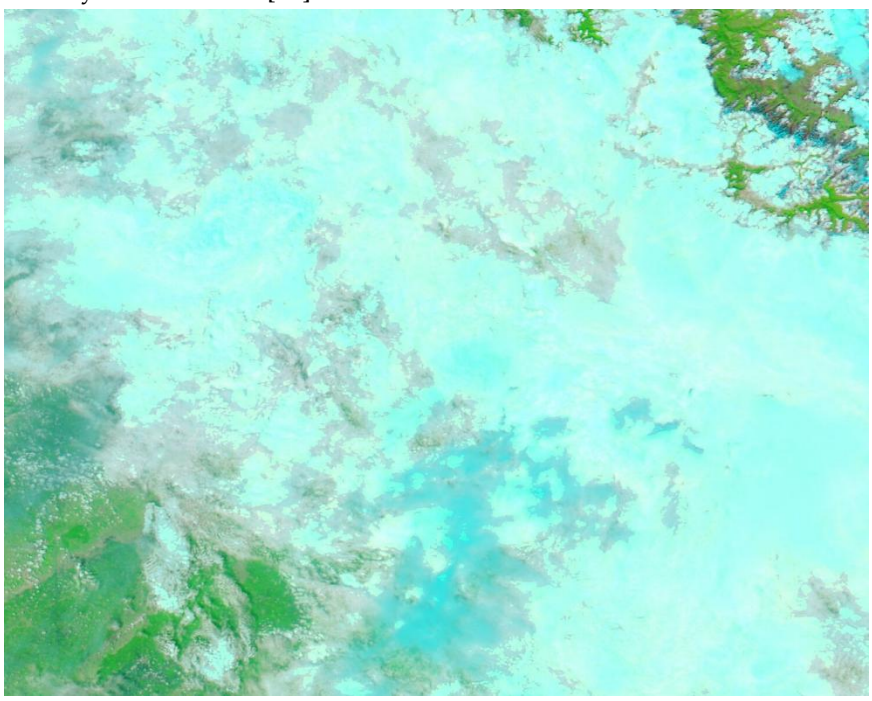


Figure 10. Before SAR Analysis Cloud Cover

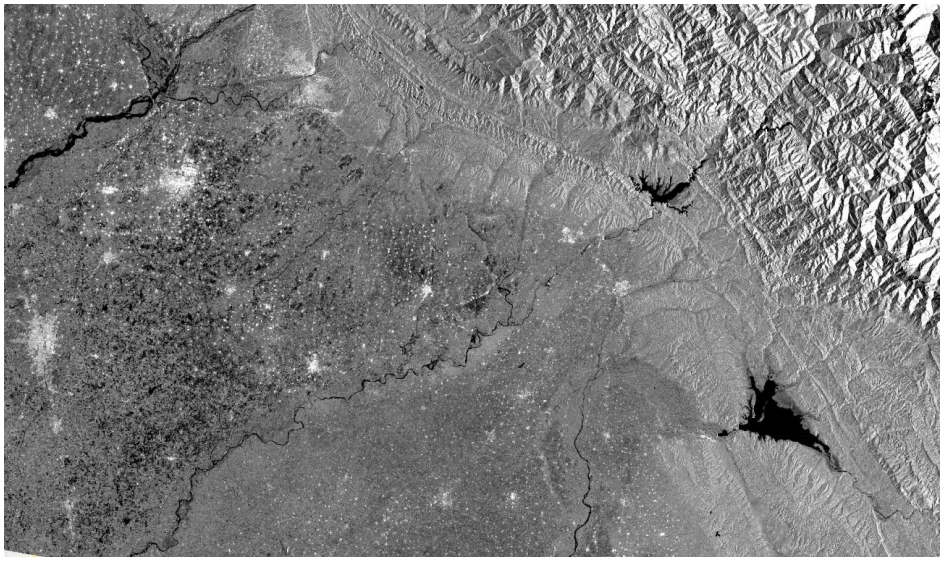


Figure 11. After SAR Analysis Cloud Cover Remove

This is the Process of SAR Analysis later this Bhakra Dam SAR Analysis also shown.

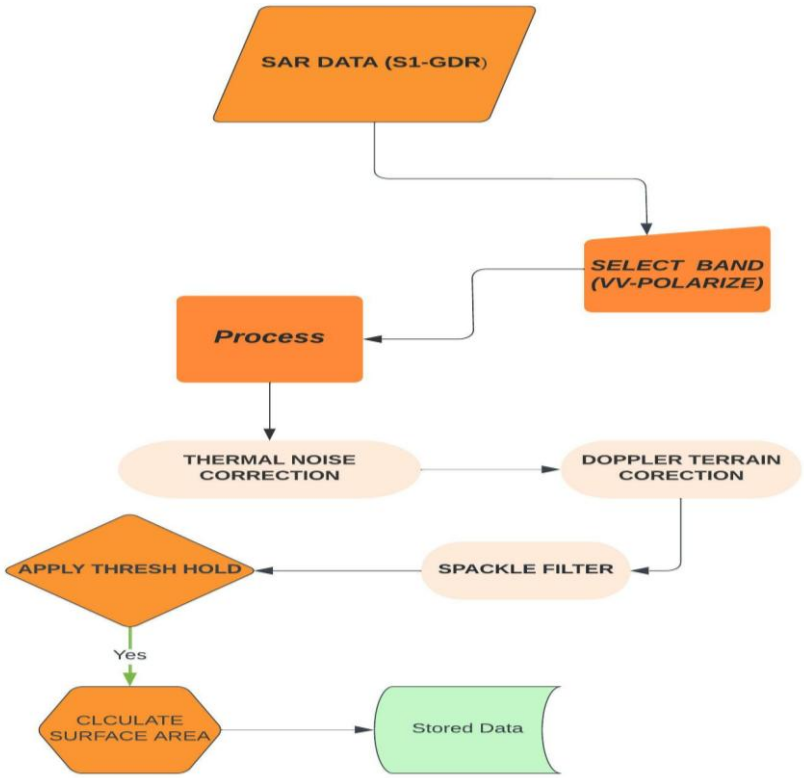


Figure 12. SAR Process Flowchart

This flow chart outlines the procedure for handling SAR (Synthetic Aperture Radar) data, specifically S1-GDR data. The process commences with the input of SAR data labeled as S1-GDR, followed by the selection of the VV-polarize band. The data then undergoes thermal noise correction and Doppler terrain correction, followed by the application of a spackle filter. Subsequently, a threshold of -0.16 is applied to the filtered data, resulting in the computation of the surface areaLastly, the surface area calculated is then stored with the processed data. This systematic methodology provides correct and effective processing of SAR data and involves the multiple corrections and filters before it is stored. One of the first things to do is to eliminate thermal noise and in this way, undesired interference due to external environmental factors is removed. Unchecked, thermal noise may significantly affect the quality of SAR backscatter data leading to false classification of surface. The signal is enhanced by means of the corrective action against thermal noise, thereby enhancing its signal clarity, which results in more accurate later analysis including the detection of water bodies. Then, the Doppler terrain correction is used to fix the spatial distortion due to the interaction of the radar with the topography of differing heights, which made the SAR images properly aligned with geographical coordinates. [40]

A speckle filter is achieved to reduce the grainy distortion seen in SAR imagery, which is an effect of the SAR signal coherency. This process of filter does not significantly reduce the level of detail in the image, but does improve the clarity of it. Backscatter coefficients are filtered and a threshold value of -0.16 is applied to distinguish between water and non-water areas. The method supported by the previous studies of SAR information to map the reservoir enables the correct determination of the reservoir surface area [41] [42].

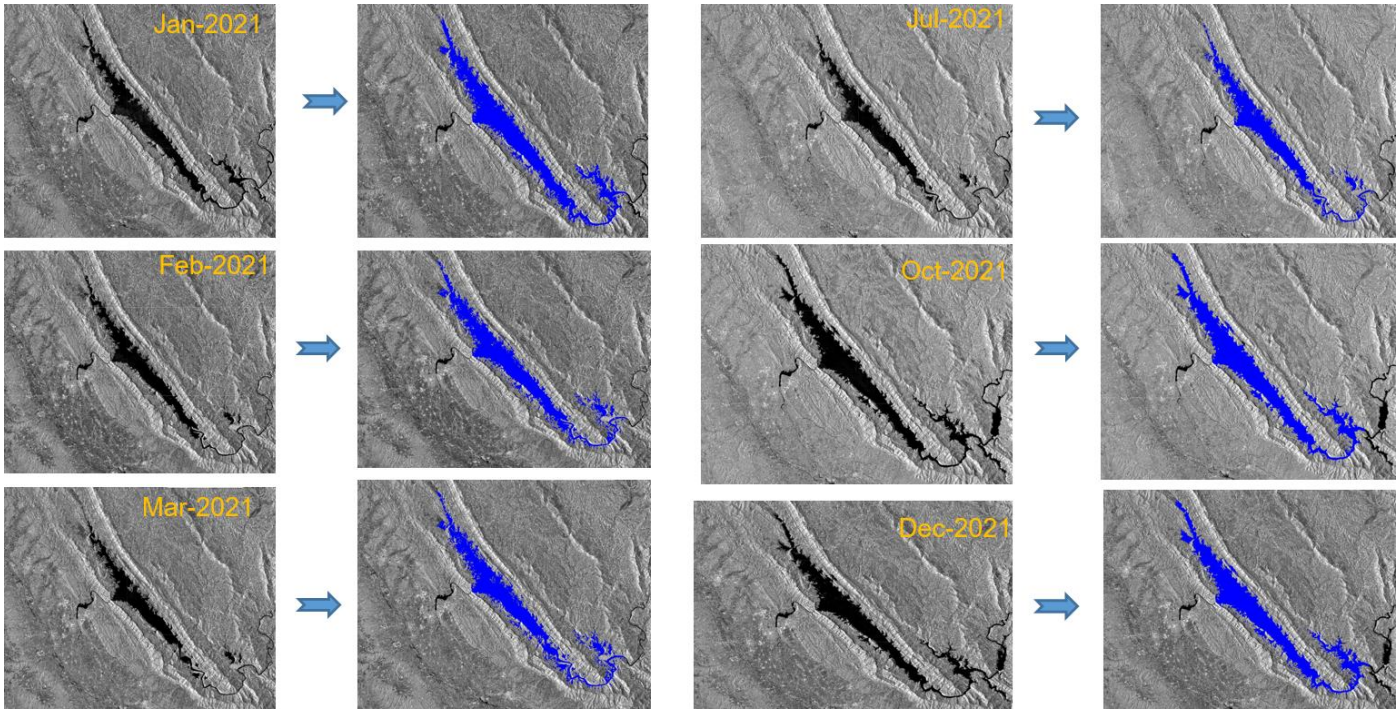


Figure 13. SAR Analysis on Bhakra Dam

3.5. *Basic Concept of Reservoir Operation*

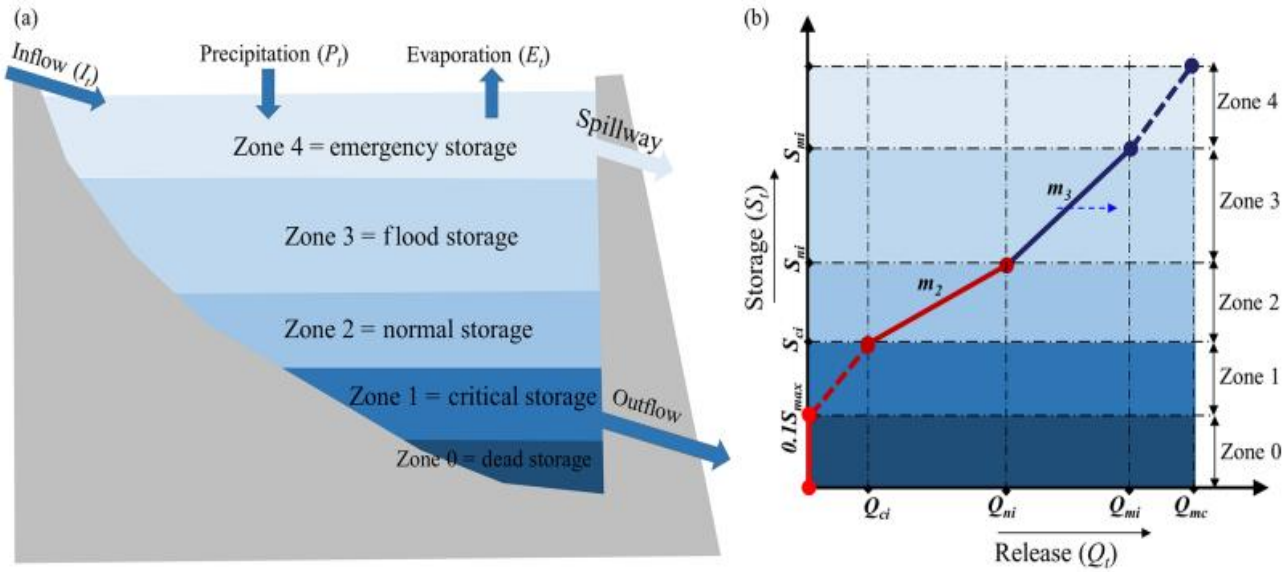


Figure 14. Representation and improved parameterization of reservoir operation

The diagram indicates the segmentation of a reservoir into four different zones that are characterized by inflows and outflows. Furthermore, it describes a reservoir-release character which is piecewise-linear in nature. Parameters m1 and m 2 determine steepness of the release curve and they change monthly. The blue arrow upwards indicates that inflow to the reservoir may also affect the determination of the release in Zone 3 [1].

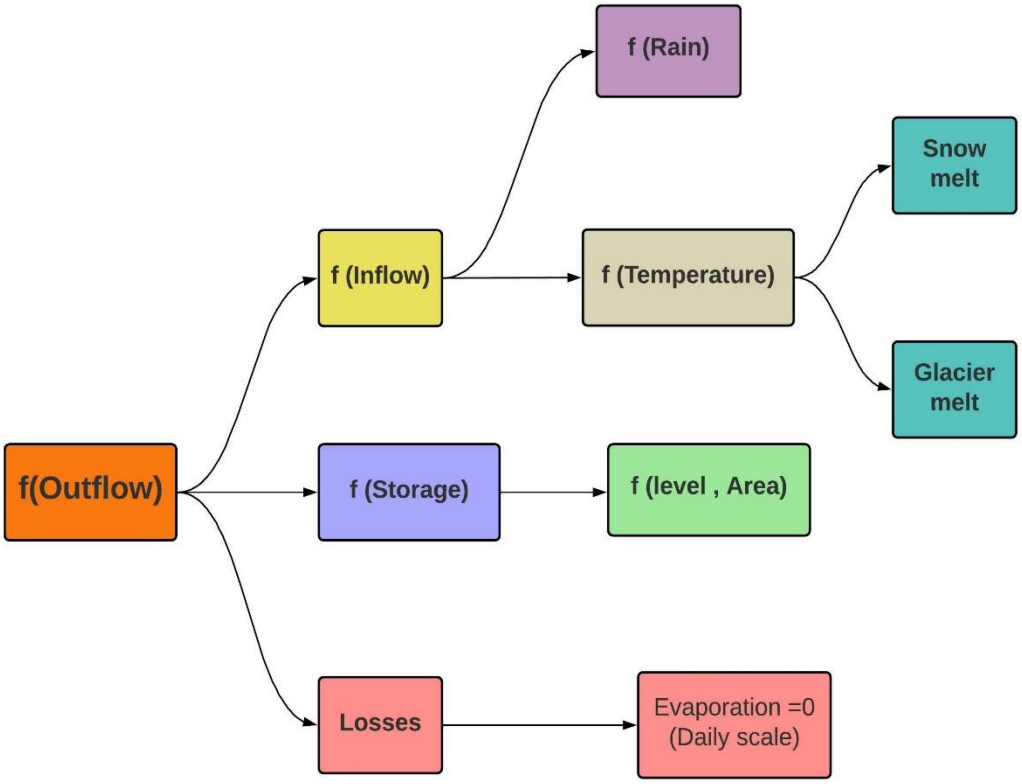


Figure 15. Procedure for determining the outflow of a water reservoir system

The process of outflow determination of a water system is demonstrated in this flow chart. Three major factors affect the outflow, and these are the inflow, storage, and losses. Inflow depends on the interaction between precipitation and temperature, and the tem-pera-temperature also influences snowmelt and glacier melt. The level and the area of a water body determine storage. The losses are considered where evaporation will be set to zero on a daily basis. All of this leads to the calculation of the outflow giving the full image of the dynamics of the water system. The outflow basic formula is given below.

$$f(Q)= f(I) – f(S) – \text{Losses (Evaporation)}.$$

3.6. *Machine Learning Algorithms*

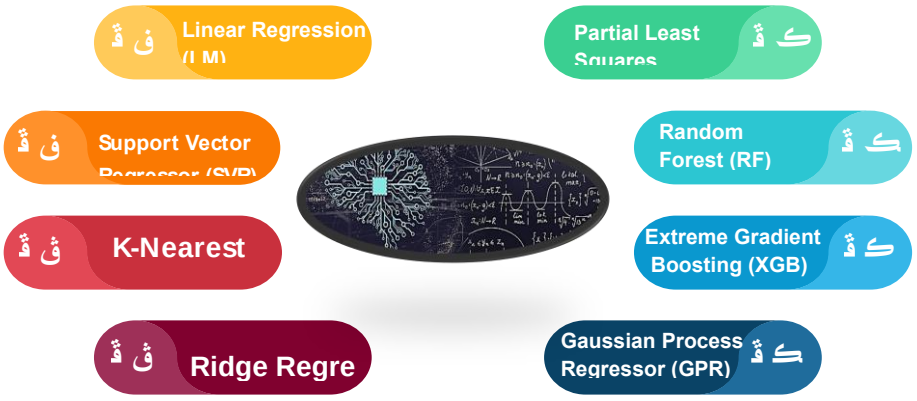


Figure 6. Machine Learning Algorithms

3.6.1. K-Nearest Neighbors (KNN)

KNN is a nonparametric method. The popularity of KNN has been boosted by the fact that the algorithms are fairly general and simple. The system can capture all the existing issues or cases and put them into new groups based on a similarity measure (Zamri et al., 2022). KNN algorithm is a supervised learning approach that anticipates the classification of unlabeled data on the basis of the attributes and labels of the training data (Bzdok et al., 2018). The KNN algorithm labels data-sets based on the use of k nearest training data points (neighbors) nearest to the test query. It uses a similar training model to the experimented query (Ma-hesh, 2020). The principle applied in KNN is the nearest neighbor and classify them according to its simi-larity. The KNN typically function under the assumption of the Euclidean dis-tance measure to identify data points nearest to a particular group (Zamri et al., 2022). A flexible choice among the popular distance measures is an-other alternative distance measure, the Minkowski distance with a distance parameter q. It generalises the Eu-clidean distance. KNN algorithm is an easy and frequently utilized ma-chine learning algorithm in the classification task due to its very flexible and understandable structure (Mahesh, 2020). The technique has been conversant with its application in the resolution of regression and classification assignments including data of var-ying sizes, the amount of labels, the magnitude of noise, and the range and context (Zhang et al., 2017). In KNN regression, k is the weighted average of the nearest k neighbors of a sample point to provide the prediction of that point. Weights of the neighbours are then set by the distances of the sample point to its neighbours, with the nearest neighbours receiving the higher weights. There may be a need to reduce the di-mensions by resolving the problem, as (Lange et al., 2020) also indicates. Support Vector Machine (SVM)

3.6.2. The support vector machine (SVM)

SVM is highly esteemed as a reliable prediction system due to its main objective of minimizing the upper bound on the generalization error, rather than solely focusing on the training error (Vapnik,1995). Furthermore, the solution is globally optimal in often achievable situations, whereas competing machine learning methods like ANN may only converge to local minima. The Support Vector Machine (SVM) algorithm transforms the input variables into a feature space of greater dimensions, denoted as $\phi(x)$. The map is typically implemented implicitly via a kernel function, also referred to as the kernel trick. The kernel function exhibits similarity to the covariance function in Gaussian process. Support Vector Machine (SVM) is a machine learning method that discovers the most effective hyperplanes inside the feature space to accurately differentiate between different classes. The optimization of the separation between the hyperplanes, known as the margin, is how this is accomplished. By employing the kernel technique, Support Vector Machines (SVM) may accurately identify data points that cannot be separated by a linear boundary in the original input space. Regression problems Support Vector Machines (SVM) seek to minimize an objective expression consisting of a loss term that grows past a set limit and a L2 regularization term. Ideally, the choice of the kernel function must be based on the natural make-up of the incoming data, as well as, their relation to the output. Also, one should add that the SVM model is sparsely rep-resented as a linear combination of a specified subset of the training data. This sub-population is projected into the feature space and is called the support vectors (support vec-tors) (Xu et al., 2021).

3.6.3. Gaussian Process Regression (GPR)

Gaussian process regression (GPR) is a Bayesian method used in the case of kernel regression and has been reported to perform extremely well in various benchmark problems. A Gaussian Process (GP) is a system of random variables correlated with some locations in time and space and that are distributed in a multivariate Gaussian fashion. A Gauss-ian Process (GP) belongs to a mean function and a covariance function, which

identifies covariance between all pairs of random variables that denote the quantity of interest at different locations or time points. It is expected that the two functions must reflect the prior knowledge that the user had of the overall pattern and the degree of consistency exhibited by the target function, respectively. The utilization of the covariance function bears resemblance to the kernel approach employed in Support Vector Machines (SVM) (Rasmussen & Williams, 2006), since it implicitly transforms the inputs into features $\phi(x)$. The Gaussian Process (GP) is employed in kriging methodologies within the realm of geo statistics, where the average and covariance are commonly represented as functions of geographic coordinates. In the domain of machine learning, the explanatory variables constitute the independent variables of the mean and covariance functions. Gaussian Process Regression (GPR) can precisely estimate intricate and non-linear connections between the target and input variables. Gaussian process regression is a statistical technique that measures the level of uncertainty linked to the predictions. Unlike the SVM, which relies on a limited number of support vectors, GPR achieves precise predictions for new data points by computing a linear combination of all the training data points. The weights for this combination are calculated using the covariance function. The disadvantage with Ground penetrating radar (GPR) is that; the cost of computing the covariance matrix as well as performing it might prove to be prohibitive when using a large dataset. To solve this problem and to make GPR more scalable to large datasets, scholars have come up with several approximation algorithms (Liu et al., 2020).

3.6.4. Linear Regression (LR)

Linear Regression (LR) is a statistical process that is applied to explain the relationship between a single response variable and one or more explanatory factors. It is a linear methodology that seeks to come up with the most appropriate line that fits this relationship. Linear regression develops a linear relationship between variables of input and output. Linear function is a direct relationship between input variables and an output variable which is represented by a straight line. According to Michele et al. (2020), the linear function is obtained with the help of optimization criteria.

Any regression machine learning model is often founded on what is commonly referred to as the regression equation that can be represented as: $Y = XB + e$.

Where Y is the dependent variable, and X is the independent variable, the coefficients of the regression are to be estimated, denoted as B, and the er-rors or residuals are denoted by e.

3.6.5. Ridge Regression

Ridge regression is a statistical approach used to analyze multiple regression models when the data is affected by multi collinearity. Multi-collinearity gives rise to an issue where the basic linear regression model (least square estimates) loses its unbiasedness and the variance increases to the extent that the projected values deviate significantly from the true value. The primary benefit of employing ridge regression is to mitigate the issue of overfitting. It operates similarly to linear regression, but incorporates an additional term (α) to mitigate overfitting. The objective of any machine learning model is to extrapolate the underlying pattern that has to be predicted. In other words, the model should perform optimally on both the training and test data. Overfitting is the phenomenon where a trained model demonstrates good performance on the training data but performs badly on the testing dataset. Ridge regressions have been extensively utilized for the prediction of streamflow, and their results have been determined to be equivalent to cutting-edge machine learning techniques (Lima & Lall, 2010; Chokmani et al., 2008; Xiang et al., 2020).

3.6.6. XG Boost Regressor (XGB)

XGBoost is a relatively new machine learning approach, which can scale to large datasets. Chen and Guestrin developed it in 2016. Gradient Boosting Decision Tree (GBDT) is an iterative model that employs the gradient de-scent algorithm to optimise the loss function in the parameter space to solve the optimisation problem. XGBoost is, however, an enhanced GBDT. XGBoost is predominantly used largely because of its regularised boosting process. Another common characteristic of XGBoost, over and above standard GBDT, is the regularization term, which may be considered a penalty on the complexity level of the model. With the complexity attribute of each regression tree, one can easily control the complexity of the model to prevent overfitting. The objective function, denoted by Obj, is com-prised of an error and a penalty that, in fact, can be represented by the following formulas:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \epsilon \mathcal{F} \quad (1)$$

Where K is the number of trees, f denotes the functional space of F, and F represents the collection of potential CARTs. The goal function of the aforementioned model is expressed as:

$$obj(\theta) = \sum_{i}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (2)$$

The objective is Obj. Loss function, represented as l, is the difference between the observed and the predicted value of the model yi. Ω is the penalty; γ and λ are two tuning parameters. T is the count of leaf nodes and ω is the proportion of leaf nodes. XGBoost uses the Newton method to address the optimization problem in the function space and permits tailoring of the objective function to the needs of the user. A second-order Taylor expansion needs to be used to calculate the error function (Zheng et al., 2019; Ni et al., 2020). Therefore, the objective function should be twice differentiable, which is a requirement of the objective version. XGBoost is a training model that combines the outputs of every iteration of regression tree and learns based on the errors committed during previous iterations. It is also a training process which depends on residuals. XGBoost is also among the most popular machine learning models with an impressive computa-tion speed and high simulation precision, which contributes to its popularity in hydrology research (Ji et al., 2021).

3.6.7. Random Forest (RF)

The RF model, introduced by Breiman (2001), employs Breiman's "bagging" concept to create an ensemble of decision trees with regulated variance. The RF model has been extensively utilized for regression and forecasting tasks. The RF model utilizes random binary trees and exploits a subset of observations through the bootstrapping technique. This involves sampling a random subset of the training dataset from the raw dataset to construct the model. The dataset that is not included in the model development is referred to as "out-of-bag" (OOB) and is used to analyze generic problems.

The concept, known as Random Forest, emerged as an expansion of CART (classification and regression tree) with the aim of enhancing prediction accuracy (Liaw and Wiener, 2002). The Random Forest (RF) technique, proposed by Breiman in 2001, is a non-parametric and quasi-unsupervised algorithm that belongs to the decision tree family. It utilizes a collection of statistically independent trees to provide predictions for regression and classification tasks. The system combines multiple decision trees and aggregates their predictions. The method exhibits optimal performance when the quantity of observations surpasses the quantity of variables. Furthermore, it is regarded as dynamic since it can easily adjust to different ad-hoc learning challenges and effectively generate rankings of varying importance. To mitigate the statistical instability caused by even minor changes in the training dataset, several decision trees with bootstrap aggregation are employed (Mushtaq et al., 2022). Random forests are a highly efficient tool for making predictions.

Due to the Law of Large Numbers, they do not exhibit overfitting. By including appropriate randomness, they become precise classifiers and repressors. Moreover, the framework, in connection to the potency of the different predictors and their associations, provides understanding into the random forest's predictive capability (Breiman, 2001).

4. Model Evaluation

This research utilizes the R-Pkg of Hydrostat to assess the efficacy of Machine Learning Algorithms.

4.1. *Nash-Sutcliffe Efficiency (NSE)*

The Nash-Sutcliffe efficiency (NSE) is a quantitative measure employed to evaluate the effectiveness of a model or simulation.

The Nash-Sutcliffe efficiency (NSE) is a statistical metric that calculates the ratio of the residual variance (also known as "noise") to the measured data variance (sometimes known as "information") (Nash and Sutcliffe, 1970). NSE quantifies the level of agreement between the observed and simulated data by assessing how well they match with the 1:1 line. The NSE is calculated using equation 1.

$$\mathrm{NSE} = 1 - \left[\frac{\sum_{i=1}^{n} \left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{n} \left(Y_i^{obs} - Y^{mean}\right)^2} \right] \tag{3}$$

The variables utilized in this assessment include Yi obs (the ith observed value), Yi sim (the ith simulated value), Y mean (the mean of observed data), and n (the total number of observations). The NSE (Nash-Sutcliffe Efficiency) is a metric that runs from negative infinity to 1.0, where a value of 1.0 represents the highest level of performance. Acceptable values fall within the range of 0.0 to 1.0. Values that are less than or equal to 0.0 show that the observed value is a more accurate predictor than the simulated value, which is regarded undesirable.

4.2. *NSE is recommended for two major reasons*

(1) It is endorsed by ASCE (1993) and Legates and McCabe (1999), and is advised for utilization.

(2) The NSE (Nash-Sutcliffe Efficiency) is widely used to provide detailed information on reported values. Sevat and Dezetter (1991) concluded that NSE is the most effective objective function for accurately representing the overall fit of a hydrograph. Legates and McCabe (1999) proposed a modified version of NSE that is less affected by extremely high values, but it was not chosen due to its limited usage and lack of reported values. [43]

4.3. *Pearson's correlation coefficient (r) and coefficient of determination (R2)*

Pearson's correlation coefficient (r) and coefficient of determination (R2) are statistical measures used to quantify the strength and direction of the linear relationship between two variables.

Pearson's correlation coefficient (r) and coefficient of determination (R2) quantify the level of collinearity between simulated and measured data. The correlation coefficient, which ranges from -1 to 1, measures the strength of the linear relationship between observed and simulated data. A value of r = 0 indicates no linear relationship, while r = 1 or -1 indicates a perfect positive or negative linear relationship, respectively. Similarly, R2 represents the proportion of the variance in measured data that can be explained by the model. R2 ranges from 0 to 1, with higher values indicating less error variance. Typically,

values greater than 0.5 are considered acceptable (Santhi et al., 2001, Van Liew et al., 2003). However, it is important to note that r and R2 are sensitive to extreme values (outliers) and do not account for additive and proportional differences between model predictions and measured data. (Legates and McCabe, 1999).

4.3.1. Knowledge Graph Embedding Model (KGE)

Knowledge Graph Embedding (KGE) models aim to acquire the embedding and vector representation of nodes and edges through supervised learning. This is achieved by projecting them into a continuous low-dimensional space. These vectors, with a few hundred dimensions, offer memory efficiency. The vector space assigns semantic meaning to each point, representing a concept, similar to word embedding.

- Its Prerequisites are as follows
- Basics of Machine learning and Neural Networks.
- Understanding of mathematical relationships, 3D distance calculation, trigonometry, and functions.

A cup of coffee and your favorite music to help you stay focused because there will be so much math involved. A Knowledge Graph Embedding (KGE) should possess sufficient expressiveness to capture KGE qualities that pertain to the capacity to describe unique logical patterns for relations. KGE has the capability to incorporate or eliminate certain properties as per requirements. [44]

4.3.2. Model Evaluation Statistics (Error Index)

Mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) are commonly used error indices for model evaluation. These indices provide valuable information about the errors in the units (or squared units) of the variable of interest, which helps in analyzing the results. A perfect fit is indicated by RMSE, MAE, and MSE values of 0. According to Singh et al. (2004), RMSE and MAE values that are less than half the standard deviation of the measured data can be considered low, and either of these indices is suitable for model evaluation. A standardized version of RMSE is selected.

4.3.3. MSE Observations Standard Deviation Ratio (RSR)

RMSE is a frequently employed statistical measure for error indices. (Chu and Shirmohammadi, 2004; Singh et al., 2004; Vasquez-Amábile and Engel, 2005) have discussed the topic. It is widely accepted that a lower root mean square error (RMSE) indicates better model performance. However, only Singh et al. (2004) have provided a guideline to determine what is considered a low RMSE based on the standard deviation of observations. Following Singh et al.'s recommendation, a model evaluation statistic called the RMSE-observations standard deviation ratio (RSR) was developed. RSR standardizes the RMSE by using the observation standard deviation and incorporates both an error index and additional information suggested by Legates and McCabe (1999). RSR is calculated by dividing the RMSE by the standard deviation of the measured data, as shown in the equation.

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\left[\sqrt{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right)^2}\right]}{\left[\sqrt{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{mean}\right)^2}\right]} \quad (4)$$

The RSR metric combines error index statistics and incorporates a scaling/normalization factor to make it applicable to different constituents. RSR ranges from 0, indicating perfect model simulation with no residual variation (RMSE), to a large positive value. A lower RSR indicates a lower RMSE and better model simulation performance.

$$\text{Collinearity Index} = 50\%(\text{Training Collinearity}) + 50\% \text{ (Testing Collinearity)}$$

$$\text{Error Index} = 50\%(\text{Training Error}) + 50\% \text{ (Testing Error)}$$

4.3.4. Model Efficiency Index(MEI)

$$\text{MEI} = f(\text{Collinearity Index} \mid \text{Error Index})$$

where,

Collinearity Index= f(NSE)

Error Index = f(RSR)

Higher the value of MEF indicates the best model

4.4. *Systematic Working of ML Algorithms*

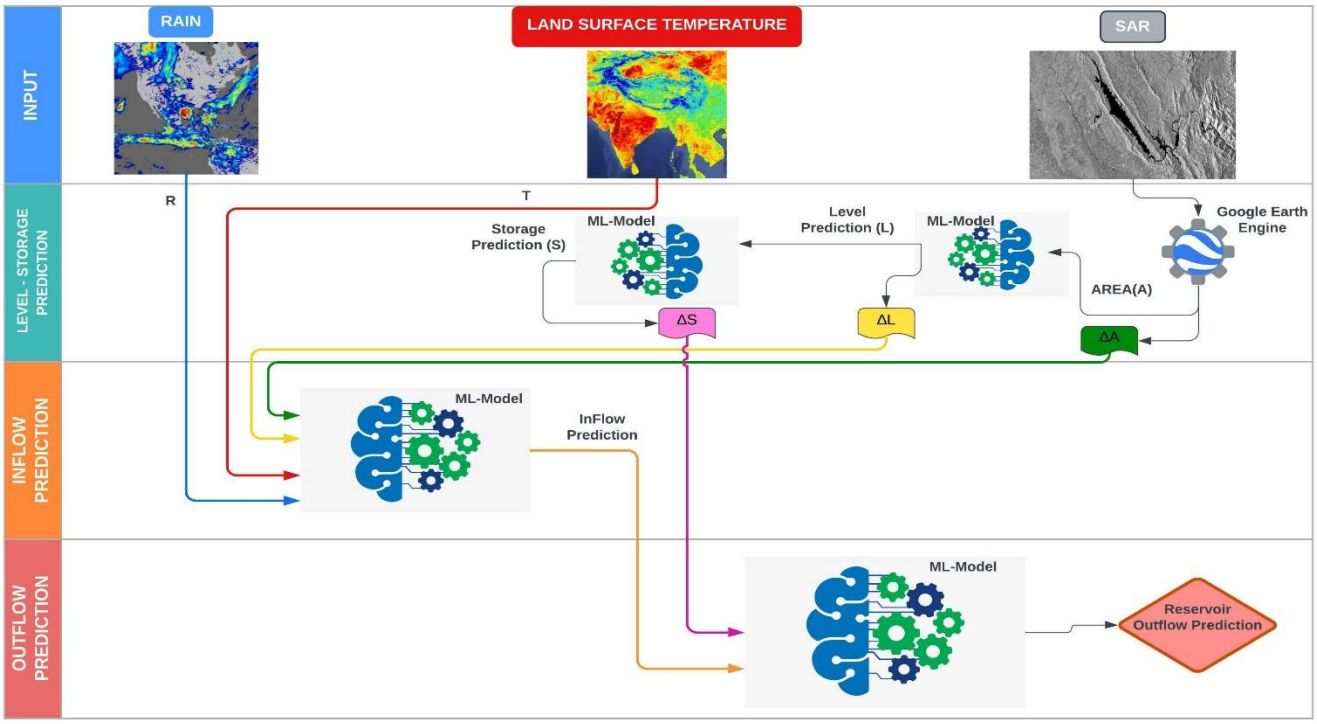


Figure 7 Systematic Work Flow Diagram

This systematic approach enables the integration of remote sensing and environmental data with advanced ML algorithms, resulting in more accurate and dynamic predictions of reservoir inflow and outflow, which are essential for effective water resource management and flood risk mitigation.

Primary inputs include Rainfall (R), Land Surface Temperature (LST), and Synthetic Aperture Radar (SAR) data. Machine learning models are utilized to generate predictions for water level and storage by analyzing the relationships among environmental factors. The models calculate the variation in storage (1) and water level (1). Google Earth Engine is used to process SAR data to calculate the change in the surface area (ΔA) of the reservoir as an outcome of the temporal changes. The changes in level (ΔL), storage (ΔS), and area (ΔA) are calculated, and they are vital characteristics to the following machine learning mod-ule, which predicts the inflow in the reservoir. The inflow forecasts, and other variables calculated, are then inputted into a later machine learning model to predict reservoir outflow.

4.4.1. Input Stage

The methodology begins with collection of input information, such as rainfall (R), land surface temperature (T) and SAR imaging. These are essential in the understanding of the hydrological processes that influence the reservoir.

4.4.2. Level-Storage Prediction

The first analytics process involves predicting the storage (S) in the reservoir using a machine learning (ML) algorithm. Using this algorithm, the data on rainfall and tem-perature is taken into account to estimate the amount of water in the reservoir.

4.4.3. Level Prediction:

At the same time, the machine learning model predicts the water level (L) in the reservoir as well. This projection is affected by the differences in storage (ΔS) and is essential to understand the current state of the reservoir.

4.4.4. Area Calculation:

The surface area (A) of the reservoir is calculated with the help of Google Earth Engine that works with the SAR data to give accurate results. Also considered in the study is any change in the area (ΔA).

4.4.5. Inflow Prediction:

It is projected to inflow into the reservoir using another machine learning model. This forecast is based on the precipitation data and the predicted stor-age levels which will help estimate the amount of water flowing into the reservoir.

4.4.6. Out Flow Prediction

Finally, the predictions gained after the steps above become the input of another machine learning model that determines the outflow of the reservoir. This outflow estimation uses the information of inflow, storage and water level to provide a comprehensive understanding of amount of water that will be released out of the reservoir.

Overall, our workflow is an effective way of combining various data sources and machine learning techniques to predict the behavior of the reservoir, ensuring that water management decisions are made..

5. Results

5.1. *Data Analysis*

In this process plotted Correlation matrices to select the suitable parameters for the input of Machine Learning Algorithms.

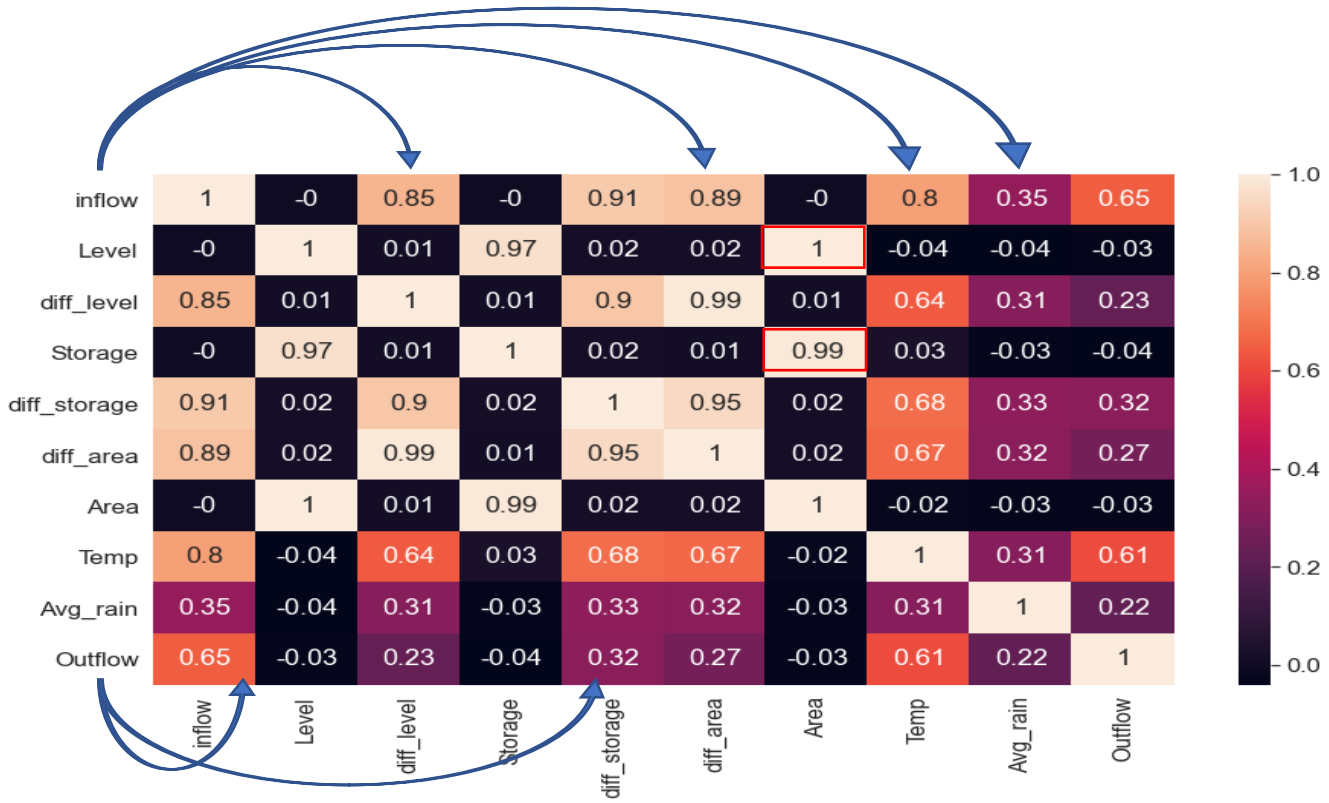


Figure 8. Correlation Model Efficiency Matrix

The model efficiency index, MEI, is a thorough measure that assesses how well a model performs by taking into account both collinearity and error indicators. A greater MEI suggests a stronger fit of the model to the observed data, which is crucial for guaranteeing accurate predictions in reservoir operations.

This matrix displays the inflow relationship between variables such as level difference, area difference, temperature, and average rainfall. The high results indicate a significant link between these variables.

The second essential variable in this matrix is the outflow, which exhibits a strong correlation with both inflow and the difference in storage. Therefore, we analyze these variables and their relationship to assess the outflow, which is the variable of interest. However, we exclude temperature from the outflow analysis because inflow already incorporates temperature data, and the daily nature of this data results in a negligible evaporation rate, close to zero, which has no impact on the outflow.

The relationships between various factors and inflow and outflow in reservoirs are significant. Inflow is strongly linked to differences in level, storage, area, and temperature, emphasizing their importance in predicting inflow. Outflow, on the other hand, is closely connected to inflow and differences in storage. Levels and storage are highly correlated, as are differences in level and area, showcasing their impact on reservoir dynamics. While temperature and average rainfall also play a role, temperature is not directly analyzed for outflow as its influence is already accounted for in inflow. Understanding these relationships can help in developing more targeted and efficient machine learning models for reservoir management.

Inflow shows strong positive correlations with variables such as difference in level (0.85), difference in storage (0.91), difference in area (0.89), and temperature (0.8). This highlights that changes in reservoir level, storage, and area, as well as temperature, are significant contributors to predicting inflow.

Outflow is most strongly correlated with inflow (0.65) and difference in storage (0.61), indicating that these two factors are critical for accurately forecasting outflow.

Level and storage display a near-perfect correlation (0.97), and the difference in level and difference in area are also highly correlated (0.99), underscoring the interconnectedness of these parameters in reservoir dynamics.

Temperature and average rainfall show moderate correlations with inflow and outflow, but temperature is excluded from direct outflow analysis because its effect is already captured through inflow, and daily evaporation rates are negligible.

The matrix helps identify which variables are most influential for modeling and predicting outflow, supporting more targeted and efficient machine learning models for reservoir management.

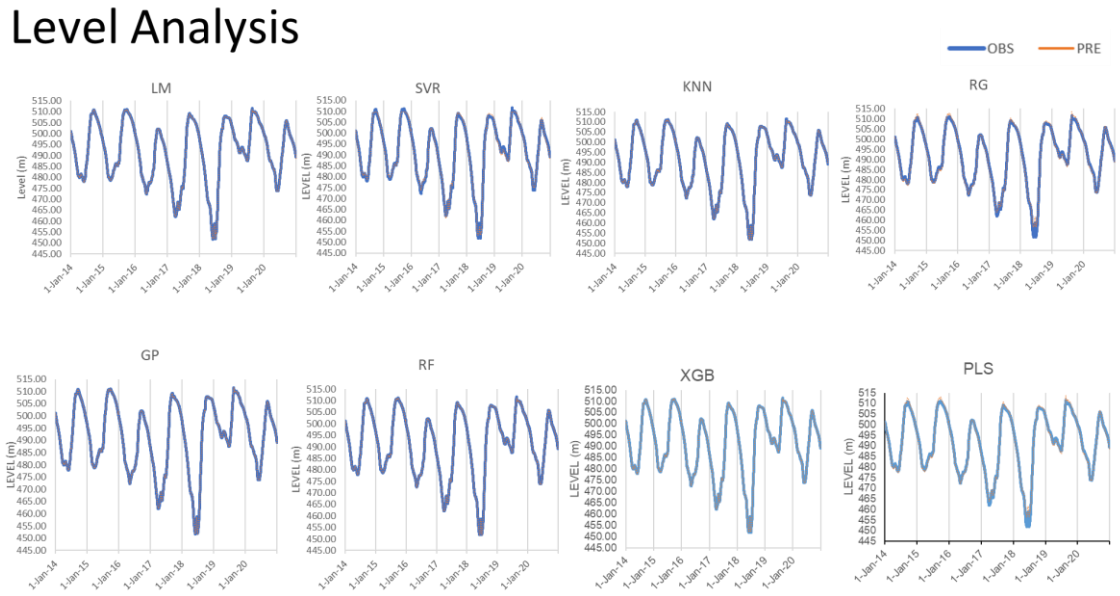


Figure 9. Level Analysis Results

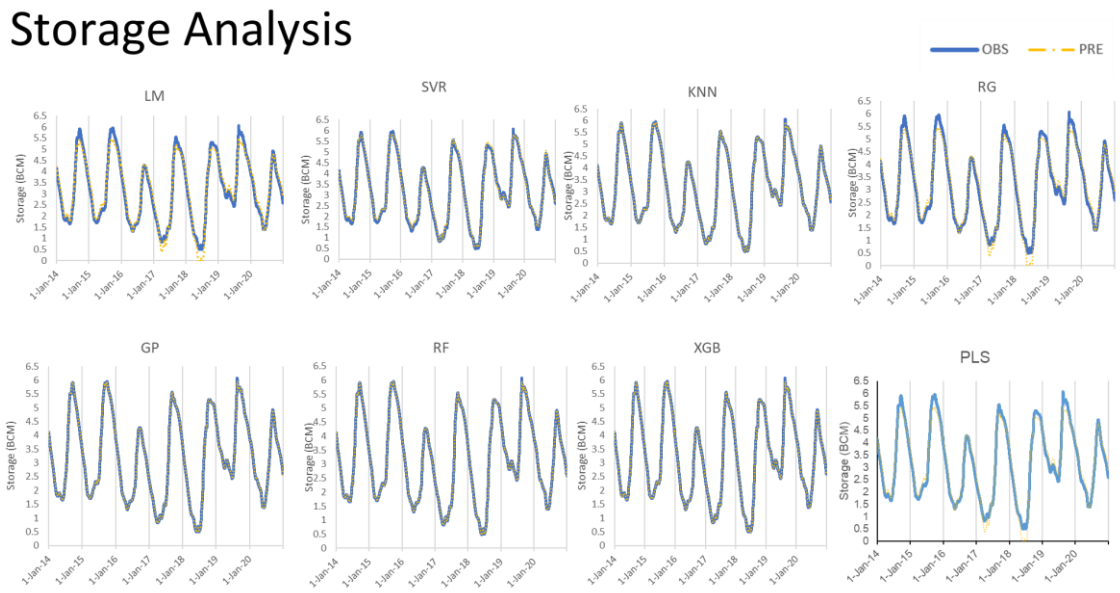


Figure 10. Storage Analysis Results

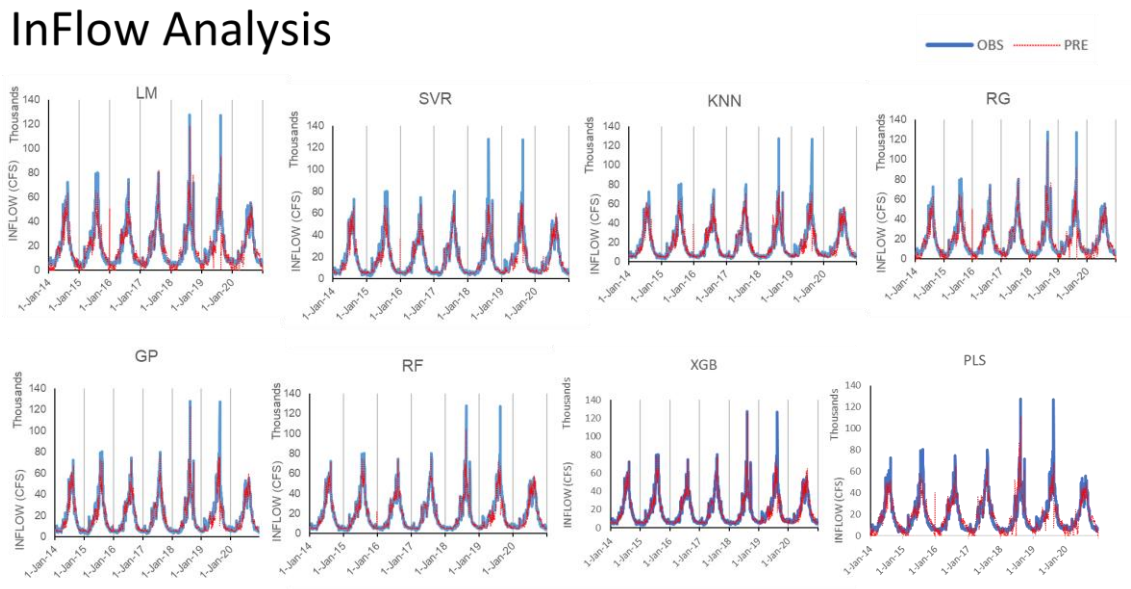


Figure 11. Inflow Analysis Results

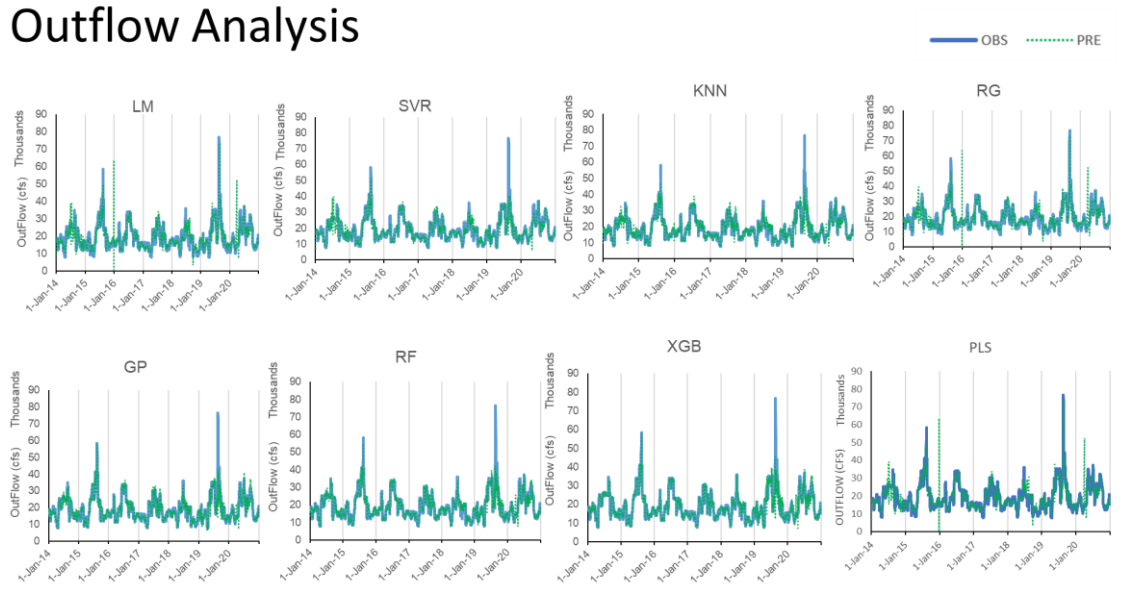


Figure 12. Outflow Analysis Results

5.2. *Best Fit Model Selection Based on Results*

5.2.1. Best Fit Model for Level

In the Table Highlighted Models & Model Efficiency Index Results are best fit Results. The Level Analysis Table assesses the accuracy of multiple machine learning models in forecasting reservoir water levels, utilizing a range of statistical measures, including Nash-Sutcliffe Efficiency, R-squared, Kling-Gupta Efficiency, Mean Absolute Error, Mean Squared Error, Root Mean Squared Error, and Residual Standard Deviation, for both training and testing data sets. The Model Efficiency Index serves as a comprehensive metric to gauge model performance, where higher scores indicate a stronger correlation between predicted and actual values.

KNN demonstrated flawless performance across all primary evaluation metrics (NSE, R^2, KGE = 1) for both the negligible error rates (MAE, MSE, RMSE, RSR was exceptionally high, indicating a nearly perfect fit. As a non-parametric approach, KNN relies on identifying the most similar training examples in the feature space to make predictions. Its this case implies that the reservoir level data exhibit robust local trends, which KNN is well-suited to capitalize on. [45]

XGB is a robust method in ensemble learning that constructs numerous decision trees in a sequence, focusing on minimizing errors at each stage. Its capacity to represent intricate, non-linear connections makes it especially well-suited for analyzing hydrological time series data. Achieving virtually flawless predictions, as evidenced by optimal values of NSE, R^2, and KGE, and negligible errors in MAE precision in both model training and validation. The exceptional performance of XGB in predicting be attributed to its ability to its overall resilience in modeling complex hydrological systems. [46]

Random Forest (RF) demonstrates exceptional performance across various metrics such as NSE, to 1, while MAE, MSE, RMSE, and RSR are close to 0. Additionally, RF exhibits an unlimited MEI, underscoring its dependability in predicting levels. Given its composition of decision trees, RF effectively prevents overfitting and can effectively manage intricate interactions among features often observed in reservoir level behavior. [47] [48]

The GP model demonstrates exceptional accuracy (NSE, R^2, KGE = 1; MAE, MSE, RMSE, RSR = 0) even with an infinite MEI, suggesting that it can replicate the underlying dynamics with remarkable fidelity. As a probabilistic, non-parametric approach, GP offers not only precise forecasts but also quantifiable uncertainty assessments, and its adaptability enables it to effectively capture gradual patterns in reservoir level data.

The selected models, namely KNN, XGB, RF, and GP, consistently provided the most accurate predictions for reservoir levels. This is denoted by their great statistical measures and elevated MEI values. Recent research establishes that the models are skillful to examine nonlinear, high-dimensional hydrological data. These results demonstrate the significance of applying advanced machine learning methods to accurate and reliable forecasting of water levels during reservoir control. The most notable models were the KNN, XGB, RF, and GP, which gave very accurate estimates of reservoir levels. This was demonstrated by their good statistical results as well as the highest MEI values. Recent research has also validated their ability to be successful in addressing complex, nonlinear, and multidi-mensional hydrological data. The results indicate the significance of applying advanced machine learning methods to make accurate and reliable predictions of the water levels in reservoir operations. [49]

Table 3. Level Analysis Table

Model	LEVEL (m)														MEI
	Training							Testing							
	NSE	R2	KGE	MAE	MSE	RMSE	RSR	NSE	R2	KGE	MAE	MSE	RMSE	RSR	
LM	1	1	1	0.07	0.01	0.1	0.01	1	1	1	0.08	0.01	0.11	0.01	100.00
SVM	1	1	1	0.82	0.83	0.91	0.06	0.99	0.99	0.99	0.88	0.96	0.98	0.11	11.71
KNN	1	1	1	0.01	0	0.02	0	1	1	1	0.02	0	0.04	0	INF
XGB	1	1	1	0	0.05	0	0	1	1	1	0	0	0.05	0	INF
PLS	0.99	0.99	1	0.92	1.33	1.15	0.08	0.99	1	0.94	0.77	0.75	0.87	0.09	11.65
RF	1	1	1	0	0	0.01	0	1	1	1	0.01	0	0.03	0	INF
RG	0.99	0.99	1	0.92	1.33	1.15	0.08	0.99	1	0.94	0.77	0.75	0.87	0.09	11.65
GP	1	1	1	0	0	0	0	1	1	1	0	0	0	0	INF

5.2.2. Best Fit Model for Storage

Table Highlighted Models and Model Efficiency Index Results exhibit best fit Re-sults.

K-Nearest Neighbor (KNN), Extreme Gradient Boosting (XGB) and Random For-est (RF) showed high predictive performance of the reservoir storage level in addition to inflow and outflow. These algorithms take advantage of spatiotemporal patterns in the data, where KNN uses similarity based on proximity and XGB and RF uses complex interactions between features. They are collectively useful in giving a sound estimation of storage changes that are important in the management of reservoir and in the provision of water se-curity.

Table 4. Storage Analysis

Model	STORAGE (BCM)														MEI
	Training							Testing							
	NSE	R2	KGE	MAE	MSE	RMSE	RSR	NSE	R2	KGE	MAE	MSE	RMSE	RSR	
LM	0.97	0.97	0.98	0.22	0.07	0.26	0.17	0.96	0.98	0.87	0.2	0.05	0.23	0.2	5.22
SVM	1	1	0.98	0.1	0.01	0.11	0.07	0.99	0.99	0.99	0.09	0.01	0.1	0.09	12.44
KNN	1	1	1	0	0	0	0	1	1	1	0	0	0.01	0.01	200.00
XGB	1	1	1	0	0	0.01	0	1	1	1	0	0	0.01	0.01	200.00
PLS	0.97	0.97	0.98	0.22	0.07	0.26	0.17	0.96	0.98	0.87	0.2	0.05	0.23	0.2	5.22
RF	1	1	1	0	0	0	0	1	1	1	0	0	0.01	0.01	200.00
RG	0.97	0.97	0.98	0.22	0.07	0.26	0.17	0.96	0.98	0.87	0.2	0.05	0.23	0.2	5.22
GP	1	1	1	0.01	0	0.01	0.01	1	1	1	0.01	0	0.01	0.01	100.00

5.2.3. Best Fit Model for Inflow

In the Table Highlighted Models & Model Efficiency Index Results are best fit Results.

XGB performed better in extreme gradient boosting than other machine learning algo-rithms in predicting the inflow in reservoirs. XGB is superior in capturing the complexity and non-linearity of relationships and interactions in hydrological processes that are influenced by different environmental factors such as precipitation, temperature, and land surface alterations. By successively improving prediction errors and combining less powerful models, XGB effectively identifies minor patterns and associations in the input data, leading to accurate inflow predictions. This aspect is especially useful when dealing with transboundary reservoirs, whereby inflow trends may be unpredictable and influenced by climat-ic and human influences upstream.

Recent research on XGB as an effective machine learning model that predicts reservoir inflow is justified due to its ability to observe complex, non-linear interactions and interdependences among multiple environmental variables, such as rainfall land surface properties. [50] [51] [52].

Table 5. Inflow Analysis Table

Model	INFLOW (CFS)														MEI
	Training							Testing							
	NSE	R2	KGE	MAE	MSE	RMSE	RSR	NSE	R2	KGE	MAE	MSE	RMSE	RSR	
LM	0.89	0.89	0.92	3999.27	28077413	5298.81	0.33	0.79	0.82	0.81	4863.93	51709021	7190.9	0.45	2.15
SVM	0.92	0.93	0.92	2476.94	19504080	4416.34	0.27	0.84	0.87	0.82	3920.64	40237097	6343.27	0.4	2.63
KNN	0.94	0.94	0.95	2350.93	16829738	4102.41	0.25	0.84	0.86	0.85	4040.81	41186821	6417.7	0.4	2.74
XGB	0.99	0.99	0.99	906.95	1732059	1316.08	0.08	0.83	0.85	0.87	4088.15	41902121	6473.18	0.41	3.71
PLS	0.85	0.85	0.89	4653.68	37649105	6135.89	0.38	0.78	0.81	0.76	4954.66	54807797	7403.23	0.47	1.92
RF	0.99	0.99	0.97	1119.98	3473336	1863.69	0.12	0.85	0.86	0.87	3871.06	38314429	6189.87	0.39	3.61
RG	0.89	0.89	0.92	4004.52	28100262	5300.97	0.33	0.79	0.82	0.81	4850.21	51492771	7175.85	0.45	2.15
GP	0.96	0.96	0.96	2124.52	11576632	3402.44	0.21	0.85	0.87	0.85	3865.08	38657980	6217.55	0.39	3.02

5.2.4. Best Fit Model for Outflow

In the Table Highlighted Models & Model Efficiency Index Results are best fit Results.

The Random Forest (RF) algorithm has proven to be the most effective model for predicting outflow. RF is a machine learning technique that creates multiple decision trees and combines their outputs to enhance prediction accuracy and prevent overfitting. Its strength lies in its ability to handle complex relationships between variables, making it suitable for modeling reservoir outflow, which is influenced by various factors such as inflow, reservoir storage, operational releases, and other hydrological elements. RF's capacity to process large amounts of data and automatically select important features enables it to pinpoint key predictors for outflow, leading to dependable and understandable forecasts crucial for water resource management and flood prevention. [53] [54] [55].

Table 6. Outflow Analysis Table

Model	OUTFLOW (CFS)														MEI
	Training							Testing							
	NSE	R2	KGE	MAE	MSE	RMSE	RSR	NSE	R2	KGE	MAE	MSE	RMSE	RSR	
LM	0.83	0.83	0.88	1699.65	7645300	2765.01	0.41	0.91	0.92	0.85	1549.87	7719650	2778.43	0.31	2.42
SVM	0.91	0.91	0.92	1359.22	4104768	2026.04	0.3	0.79	0.81	0.77	1516.09	17079132	4132.69	0.46	2.24
KNN	0.94	0.97	0.94	1086.53	2858781	1690.79	0.25	0.78	0.08	0.75	1886.24	17930092	4234.39	0.47	2.39
XGB	0.99	0.99	0.98	527.64	516523.9	718.7	0.11	0.7	0.72	0.73	2145.11	24138391	4913.08	0.54	2.60
PLS	0.83	0.83	0.88	1699.65	7645300	2765.01	0.41	0.91	0.92	0.85	1549.87	7719650	2778.43	0.31	2.42
RF	0.98	0.98	0.96	629.26	886433.3	941.51	0.14	0.76	0.79	0.75	1899.62	19219948	4384.06	0.49	2.76
RG	0.83	0.83	0.87	1706.33	7646230	2765.18	0.41	0.9	0.92	0.85	1569.54	7767202	2786.97	0.31	2.40
GP	0.96	0.96	0.97	912.74	1643856	1288.13	0.19	0.66	0.68	0.72	2020.45	27537534	5247.62	0.58	2.10

6. Discussions

The application of machine learning algorithms such as K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGB), Random Forest (RF), and Gaussian Processes (GP) has demonstrated significant efficiency in predicting various aspects of reservoir management, including reservoir levels, storage, inflow, and outflow.

KNN is particularly effective in predicting reservoir levels and storage due to its simplicity and ability to handle non-linear relationships in data. Studies have shown that KNN can outperform traditional models in various predictive tasks, including estimating reservoir oil viscosity, which suggests its robustness in handling complex datasets [56]. Additionally, KNN has been utilized in developing reservoir operating rules, indicating its versatility in hydrological applications [57]. The model's reliance on historical data allows it to capture patterns effectively, making it a valuable tool in reservoir management.

XGB has emerged as a powerful algorithm for predicting reservoir inflow, showcasing its high accuracy and efficiency across multiple studies. It has been reported that XGB can achieve predictive accuracies exceeding 90%, with some studies indicating up to 97% accuracy in specific applications [58]. This model's ability to learn complex patterns and relationships in data makes it particularly suitable for inflow forecasting, where accurate predictions are critical for effective reservoir operation [59]. Furthermore, XGB's performance in various predictive scenarios highlights its adaptability and effectiveness in hydrological contexts [58].

Random Forest (RF) has also shown strong performance in predicting reservoir outflow. Its capacity to handle non-linear data and interactions among variables makes it a robust choice for hydrological modeling [60].RF has been recognized for producing accurate and stable predictions, which is essential for managing reservoir operations effectively. The model's ensemble approach allows it to mitigate overfitting, enhancing its

predictive capabilities in dynamic environments such as reservoirs. Studies have indicated that RF can outperform traditional statistical methods, further solidifying its role in modern reservoir management practices [60].

Gaussian Processes (GP) provide a probabilistic prediction method which not only gives point estimates but also quantifies uncertainty, which is important in reser-voir management. Although particular examples of GP as applied to reservoir predic-tions were not given, the overall benefits of GP in complex system modeling hint at its possible relevance in this area. The versatility and strength of GP models qualify them to be applicable in situations where how much the pre-dictions are uncertain is of the same significance as the predictions themselves.

Finally, the combination of KNN, XGB, RF, and GP into reservoir manage-ment practices have demonstrated positive outcomes in forecasting the reservoir levels, storage, inflow, and outflow. The strengths of each algorithm are different in the table, and they can contribute to the better choice and efficiency of work with water resources..

7. Recommendations

Recent studies have investigated the effectiveness of the different machine learning models in forecasting various attributes of the reservoir management. The K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGB), Random Forest (RF) and Gaussian Process (GP) models have yielded encouraging outcomes in terms of predicting the levels of the reservoirs. KNN, XGB, and RF have demonstrated satisfactory efficiency in their predictive capabilities of reservoir storage, and XGB has demonstrated competence especially in predicting reservoir inflow in all the models discussed. Moreover, Random Forest (RF) has been found to be effective at predicting reservoir outflow as opposed to the other models that have been tested.

These results indicate that the selection of the most appropriate model depends on the aspect of reservoir management that is under prediction. As an example, when the main emphasis is made on estimating the level of the reservoir, all four models (KNN, XGB, RF, and GP) may be viewed as potential solutions. Nevertheless, KNN, XGB, and RF can be the most desirable options to predict the reservoir storage.

In terms of predicting the in-flow of the reservoir, the XGB model stands out as the most efficient one, whereas RF is the one to be suggested to predict the out-flow of the reservoir. It is noteworthy that the results of these models can differ with regard to the dataset in question, the size of the training sample, and the complexity of underlying relationships between the input variables and the target one. Therefore, it is recommended to experiment with multiple models and compare their performance to identify the most suitable one for a given problem.

8. Conclusion

This study aims to provide insights and recommendations for enhancing reservoir management practices, facilitating informed decision-making processes, formulating effective plans, and optimizing the allocation of reservoir storage for flood control purposes. The precise prediction of reservoir inflow is of utmost importance in ensuring efficient and prompt flood management, mitigating potential harm, and minimizing the likelihood of failing to satisfy water requirements downstream. Subsequent investigations may prioritize the development of ensemble models and the subsequent comparison of their efficacy with machine learning methods in the realm of reservoir inflow prediction. Moreover, conducting a comparative analysis of machine learning algorithms and physically-based models in terms of their predicting capabilities for reservoir inflows will provide a full evaluation of the respective benefits associated with these forecasting methodologies. The use of remote sensing data in regions with limited data availability, particularly in

developing nations, has significant potential and warrants further exploration in future research endeavors. Additional topics of research might include the investigation of the impact of water management practices on lake resources, with a particular focus on the inclusion or exclusion of recreational values as control variables in optimization processes. Furthermore, there is potential for exploring frameworks aimed at improving reservoir water management, as well as studying arctic terrestrial ecosystems.

9. Future Research

For better Results huge amount of data and deep learning modules can accurate more results.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. S. R. M. E. Fuad Yassin et al, "Representation and improved parameterization of reservoir operation in hydrological and land-surface models," Hydrol. Earth System Science, 2019.
2. C. J. M. P. B. Vörösmarty et al, "Global threats to human water security and river biodiversity, Nature,," 2010.
3. M. W. Izhar Ahmad, "Climate change-induced spatiotemporal variations of land use land cover by using multitemporal satellite imagery analysis," Journal of Water and Climate Change, 2024.
4. O. B. B. D. Manizhe Zarei et al, "Machine learning algorithms for forecast informed reservoir operation (FIRO) to reduce flood damages," Nature, 2021.
5. I. W. M. A. A. e. a. Ahmad, "Hydrological risk assessment for Mangla Dam: compound effects of instant flow and precipitation peaks under climate change, using HEC-RAS and HEC-GeoRAS.," SN Appl. Sci, 2023.
6. M. A. S. A. I. e. a. Waseem, "Urban flood risk assessment using AHP and geospatial techniques in swat Pakistan," SN Appl. Sci., 2023.
7. Z. A. C. A. O. e. a. Zachary C. Herbert, "Long-term Reservoir Inflow Forecasts: Enhanced Water Supply and Inflow Volume Accuracy Using Deep Learning," Journal of Hydrology, 2021.
8. H. M. e. a. Muhammad Muqeet, "Enhanced cellulose nanofiber mechanical stability through ionic crosslinking and interpretation of adsorption data using machine learning," International Journal of Biological Macromolecules, vol. 237.
9. Delaney C. et al., "Forecast informed reservoir operations using ensemble streamfow predictions for a multi-purpose reservoir in Northern California," Water Resources, 2020.
10. Xiang Z. et al., "Urban drought challenge to 2030 sustainable development goals," Total Environment, p. 693, 2019.
11. Rocha J. et al., "Impacts of climate change on reservoir water availability, quality and irrigation needs in a water scarce Mediterranean region (southern Portugal)," Total Environment, 2020.
12. Ryu J. et al., "Development of a watershed-scale long-term hydrologic impact assessment model with the asymptotic curve number regression equation," Water, p. 8, 2016.
13. A. E. G. S. &. U. B. Albek M.et al., "Ensemble streamfow projections for a small watershed with HSPF model," Environ. Sci. Pollut, 2019.
14. Hong J. et al., "Development and evaluation of the combined machine learning models for the prediction of dam infow," Water, 2020.
15. B. B. L. Ficklin D. L. et al., "SWAT hydrologic model parameter uncertainty and its implications for hydroclimatic projections in snowmelt-dependent watersheds," Hydrology, 2014.
16. Y. Li et al., "“Deep learning-based object tracking in satellite videos: A comprehensive survey with a new dataset,," Remote Sens.Mag., vol. 10, 2022.
17. B. N. R. a. L. C. R. A. F. Goetz et al., "“Remote sensing for exploration; An overview," Econ. Geol.,, 1983.
18. D. Pasetto et al., "Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends," Methods Ecol. Evol.,, 2018.
19. L. X. Z. Y. Y. G. Y. a. B. A. J. L. Ma Y et al., "Deep learning in remote sensing applications: A meta-analysis and review," ISPRS J. Photogrammetry Remote Sens., 2019.
20. L. F. J. Y. G. a. Q. D. J. Fan et al., "From brain science to artificial intelligence,," Engineering, 2020.
21. Ke Q. et al., "Urban pluvial fooding prediction by machine learning approaches: A case study of Shenzhen city, China," Adv. Water Resour., p. 145, 2020.
22. I. M. A. G.-T. S. &. E. Mekanik F. et al., "A. Multiple regression and Artifcial Neural Network for long-term rainfall forecasting using large scale climate modes," 2013.

23. C. C.-T. &. Q. L. Wang W et al., "A comparison of performance of several artifcial intelligence methods for forecasting monthly discharge time series," Hydrology, 2009.
24. K. O. Erdal H. İ et al., "Advancing monthly streamfow prediction accuracy of CART models using ensemble learning paradigms," Hydrology, 2013.
25. Z.-M. M. A. D. M. L. H. &. M. Bozorg-Haddad O. et al., "A self-tuning ANN model for simulation and forecasting of surface fows," Water Resour. Manag., 2016.
26. Khan M. & Coulibaly P. et al., "Application of support vector machine in lake water level prediction.," Hydrology, p. 11, 2006.
27. Wen X. et al., "Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions," Water Resour. Manag., p. 29, 2015.
28. A. T. &. B. Y. Bilali A., "Groundwater quality forecasting using machine learning algorithms for irrigation purposes.," Agric. Water Manag., 2020.
29. B. P. e. a. M. D. &. K. P. Parida, "Forecasting runof coefcients using ANN for water resources management: The case study of Notwane catchment in Eastern Botswana," Phys. Chem. Earth, 2006.
30. T. Awchi, "River discharges forecasting in northern Iraq using diferent ANN techniques.," Water Resour. Manag., p. 28, 2014.
31. M. e. a. M. S. &. S. S. Shrestha, "Forecasting water demand under climate change using artifcial neural network: A case study of Kathmandu Valley, Nepal.," Water Supply, p. 20, 2020.
32. N. D. &. M. V. Golzar F. et al., "Forecasting wastewater temperature based on artifcial neural network (ANN) technique and monte carlo sensitivity analysis.," Sustainability, p. 12, 2020.
33. C. N. I. K. &. K. G. Trichakis I. et al., "Artifcial neural network (ANN) based modeling for karstic groundwater level simulation.," Water Resour. Manag., p. 25, 2011.
34. Liong S.-Y. et al., "Genetic programming: A new paradigm in rainfall runof modeling.," Water Resour. Assoc., p. 38, 2002.
35. K. O. Aytek A. et al., "A genetic programming approach to suspended sediment modeling," Hydrology, 2008.
36. B.-H. O. &. M. M. A. Fallah-Mehdipour E. et al, "Prediction and simulation of monthly groundwater levels by genetic programming," Hydro-Environ, p. 7, 2013.
37. B.-H. O. &. M. M. Fallah-Mehdipour E. et al., "Genetic programming in groundwater modeling.," Hydrol. Eng., p. 19, 2014.
38. M. S. J. M. Z. M. &. E. A. Babaee M.et al., " Artifcial intelligence approach to estimating rice yield.," Irrig. Drain., p. 70, 2021.
39. J. S. J. L. e. a. Lee, "Speckle filtering of synthetic aperture radar images: A review.," Remote Sensing Reviews, 1994.
40. K. M. M. A. &. M. F. J. e. a. Uddin, "Operational Flood Mapping Using Multi-Temporal Sentinel-1 SAR Images: A Case Study from Bangladesh," Remote Sensing, 2019.
41. P. S. e. a. Ramon Torres, "GMES Sentinel-1 mission," Remote Sensing of Environment, vol. 120, 2012.
42. S. Z. e. a. Zhiheng Chen, "Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine," Remote Sensing of Environment, 2022.
43. Moriasi Daniel N et al, "MODEL EVALUATION GUIDELINES FOR SYSTEMATIC QUANTIFICATION OF ACCURACY IN WATERSHED SIMULATIONS," American Society of Agricultural & Biological Engineers, vol. 50, p. 16, 2007.
44. Peng C Xia et al., "Knowledge Graphs: Opportunities and Challenges," Artificial Intelligence Review, vol. 56, p. 36, 2023.
45. R. K. G. e. a. K. Gupta, "Daily, Weekly and Monthly Rain Forecasting using Deep Learning," 3rd International Conference for Innovation in Technology, 2024.
46. A. O. P. e. a. Mosavi, "Flood Prediction Using Machine Learning Models," Water, 2018.
47. R. C. D. e. a. Zaher Mundher Yaseen, "Hybrid Data Intelligent Models and Applications for Water Level Prediction," Handbook of Research on Predictive Modeling and Optimization Methods in Science and Engineering, 2018.
48. A. F. e. a. Mahdi Abbasi, "A hybrid of Random Forest and Deep Auto-Encoder with support vector regression methods for accuracy improvement and uncertainty reduction of long-term streamflow prediction,," Journal of Hydrology, vol. 597, 2021.
49. D. W. X. X. e. a. Alexander Y. Sun, "Monthly streamflow forecasting using Gaussian Process Regression," Journal of Hydrology, vol. 511, 2014.
50. H. W. e. a. Wei Zhang, "Reservoir inflow predicting model based on machine learning algorithm via multi-model fusion: A case study of Jinshuitan river basin," IET Cyber-Systems and Robotics, vol. 3, 2021.
51. P. D. e. a. Hanisah Musor, "Reservoir Inflow Forecasting of the Bhumibol Dam Using XGBOOST Algorithm," International Associaation of Hydro Environment Engineering & Research, 2023.
52. A. K. e. a. Shabir Jan, "Use of machine learning techniques in predicting inflow in Tarbela reservoir of Upper," Journal of Agrometeorology, vol. 26, 2024.
53. W. Z. N. A. A. e. a. Wee, "A review of models for water level forecasting based on machine learning," Earth Sci Inform, vol. 14, 2021.

54. P.-a. Z. e. a. Jieyu Li a, "Intelligent identification of effective reservoirs based on the random forest classification model," Journal of Hydrology, vol. 591, 2020.
55. S. D. e. a. Yogesh Bhattarai, "Leveraging machine learning and open-source spatial datasets to enhance flood susceptibility mapping in transboundary river basin," International Journal of Digital Earth, 2024.
56. K. S. H. &. A. H.-S. M. Mahdiani et al., "Modeling viscosity of crude oil using k-nearest neighbor algorithm," Advances in Geo-Energy Research, 2020.
57. M. K. &. A. M. A. Ahmadi et al., "Robust methods for identifying optimal reservoir operation strategies using deterministic and stochastic formulations," Water Resources Management, 2010.
58. W. Zhao et al., "Approaches of combining machine learning with nmr-based pore structure characterization for reservoir evaluation," Sustainability, 2024.
59. H. W. Y. L. J. J.-l. W. L. &. X. A. W. Zhang et al., "eservoir inflow predicting model based on machine learning algorithm via multi-model fusion: a case study of jinshuitan river basin," Let Cyber-Systems and Robotics.
60. S. H. K. S. K. &. J. H. T. Kim, "hThe use of large-scale climate indices in monthly reservoir inflow forecasting and its application on time series and artificial intelligence models," Water, 2019.