Chapter 2

LITERATURE REVIEW

1.1 Introduction

The accurate forecasting of rainfall, also known as precipitation, is a significant challenge due to the intricate nature of meteorological events. The precise representation of precipitation in models plays a vital role in multiple industries, including flood and drought assessment, environmental planning, and water resource management. It assumes a prominent role in the reduction of extreme rainfall occurrences. This study formulates many artificial neural network models by manipulating the number of inputs, number of hidden layers, learning algorithm, and transfer function. A comprehensive literature review has been conducted to fulfil the research purpose.

Rainfall-runoff modeling using soft computing techniqueand artificial neural networks (ANNs) has been a subject of extensive research in recent years. The use of AI and ANNs in hydrological modeling has been shown to improve the accuracy and efficiency of rainfall-runoff models. In this literature review, the current state of research on the use of AI and ANNs in rainfall-runoff modeling has been discussed.

1.1.1 Literature review based on prediction of Rainfall

French et al. (1992) devised a neural network model with the purpose of predicting the spatial and temporal patterns of rainfall intensity. The training process involved the utilisation of back propagation, a widely employed technique in neural networks. In this approach, the input and output rainfall fields were presented to the neural network as a sequence of learning sets. Following the completion of the training process, the neural network was employed to predict the strength of rainfall fields with a lead time of 1 hour, utilising solely the present field as input. The generation of rainfall fields was accomplished by the utilisation of a space-time mathematical rainfall simulation model. Subsequently, the anticipated fields were juxtaposed with the fields produced by the model, which were known to be flawless. The findings demonstrated that a neural network exhibited the ability to acquire knowledge pertaining to the intricate correlation characterising the temporal and spatial progression of precipitation, as observed in a sophisticated rainfall simulation model.

The study conducted by Hsu et al., (1995) introduced a novel methodology called linear least squares simplex, which aimed to determine the structure and parameters of three-layer feed forward artificial neural network (ANN) models. The authors also showcased the potential of these models in accurately mimicking the nonlinear hydrologic behaviour of watersheds. The nonlinear artificial neural network (ANN) model approach demonstrated superior performance in capturing the rainfall-runoff relationship of the medium-size Leaf River Basin near Collins, Mississippi compared to the linear autoregressive moving average with exogenous inputs (ARMAX) time series approach. This is attributed to the fact that the ANN approach does not rely on models with physically realistic components and parameters. However, it is important to note that the ANN approach should not be considered a replacement for conceptual watershed modelling, as it serves a distinct purpose. The utilisation of neural networks presents a feasible and efficient alternative to the ARMAX time series methodology in the creation of input-output simulation and forecasting models.

In their research, Sudheer et al., (2002) introduced a novel methodology for the development of the network architecture within a rainfall-runoff model based on artificial neural networks. The researchers employed statistical techniques, specifically cross-auto and partial-auto-correlation analysis, to determine an optimal input vector that accurately reflected the basin process. Additionally, they utilised a typical training approach for model development. The approach employed in this study was verified by the utilisation of data collected from a river basin located in India. The findings of the study exhibited considerable promise, suggesting that it has the potential to substantially decrease the resources and computational duration necessary for constructing an artificial neural network (ANN) model. The algorithm under consideration has the potential to effectively achieve network compression, hence circumventing the need for a protracted and iterative trial and error process. The successful application of the methodology outlined has the potential to facilitate the automation of some procedures involved in model development.

According to research by, Wilby et al., (2003) conducted research on the identification of conceptual model rainfall-runoff processes within an artificial neural network. The study focused on the Test River basin located in southern England. Neural network methods were devised to address the daily discharge series generated by a hypothetical rainfall-runoff model, utilising observed daily precipitation totals and evaporation rates. The neural outputs linked to each hidden node were derived from the output node, following the removal of all other hidden nodes. These outputs were subsequently compared with the state variables and internal fluxes of the conceptual model. The results of the correlation study indicate that the hidden nodes in the neural network are associated with the primary processes inside the conceptual model.

In a study conducted by Ramirez et al., (2005) utilised the artificial neural network methodology to anticipate rainfall patterns specifically for the Sao Paulo region. The aim of the study was to produce precise and location-specific quantitative predictions of daily precipitation. The experiment was conducted at six different sites throughout the state of Sao Paulo, Brazil, during both the summer and winter seasons during the years 1997 to 2002. The investigation was conducted via a feedforward neural network and the robust propagation learning technique. The input data for the trained networks consisted of meteorological variables derived from the ETA model, including potential temperature, vertical component of the wind, specific humidity, air temperature, precipitable water, relative vorticity, and moisture divergence flux. These variables were utilised to generate a rainfall forecast for the subsequent time step. The comparative analysis involved evaluating the predictions generated by a multiple linear regression model and an artificial neural network (ANN). The findings of the study indicated that the artificial neural network (ANN) predictions outperformed those generated by the linear regression model.

In their research, Ghalhary et al., (2009) conducted a study on the prediction of seasonal rainfall in the Khorasan province of Iran, employing artificial neural network techniques. A total of 37 rainfall sites were chosen for analysis, covering a time span of 33 years from 1970 to 2002. The utilisation of the digital elevation model facilitated the computation of the mean precipitation in the designated geographical region. The artificial neural network model underwent training across the time span of 1970-2002, and its predictive capabilities were utilised to forecast rainfall patterns specifically for the period of 1993-2002. The findings indicate that the utilisation of the artificial neural network approach yielded significant success in the prediction of rainfall. Specifically, it was seen that this method achieved a high level of accuracy, accurately predicting rainfall in over 70% of the years examined. The model's root mean square error was determined to be 41 mm; a value deemed insignificant on an annual basis. It was anticipated that augmenting the quantity of statistical data over multiple years would enhance the accuracy of the model.

Gholizadeh and Darand (2009) studied an investigation on the application of artificial neural networks for the purpose of precipitation forecasting. A series of perceptron neural networks were trained using real monthly precipitation data obtained from the Tehran station over a span of 53 years. The anticipated quantities of precipitation during the upcoming month within the subsequent year. The artificial neural network (ANN) models demonstrated a strong correspondence with the observed data and exhibited a high degree of efficacy in forecasting monthly rainfall precipitation. The integration of neural networks with genetic algorithms yielded superior outcomes.

The study conducted by, Khalili et al., (2011) conducted an evaluation of the accuracy of daily rainfall predictions for the Mashhad synoptic station. This evaluation was carried out by employing artificial neural networks as a forecasting method. The researchers utilised daily precipitation data from the month of March, which is characterised by high humidity, as well as data from the months of May and December, which exhibit medium levels of humidity. The data spanned from 1986 to 2010 and was collected from a synoptic station. A novel methodology was introduced to analyse the precipitation patterns at a synoptic station by the implementation of an artificial neural network (ANN) model. The proposed strategy involved the utilisation of a three-layer feed-forward perceptron network, coupled with the back propagation algorithm. By employing the artificial neural network (ANN) model as an opaque model, the researchers were able to discern the latent dynamics of precipitation by leveraging historical data pertaining to the system. The methodology utilises the gradient descent technique for the purpose of training the network. The performance statistical analysis of the models obtained revealed that the selected model for daily forecasting demonstrated favorable results. Specifically, for the month of March, the correlation coefficient (R), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) were determined to be 0.89, 0.14 mm, and 1.15 mm, respectively. Similarly, for the month of May, the corresponding values were 0.85, 0.14 mm, and 1.16 mm, while for December, the values were 0.86, 0.15 mm, and 1.17 mm. These findings indicate the effectiveness of the proposed models.

The study conducted by Abbot and Marohasy (2012) elucidated the use of artificial neural networks for the purpose of rainfall prediction in the region of Queensland, Australia. The evaluation involved utilising a prototype stand-alone, dynamic, recurrent, time-delay, artificial neural network to input established climatic indices, monthly historical rainfall data, and atmospheric temperatures with the purpose of assessing the applicability of artificial intelligence in monthly and seasonal rainfall forecasting. The study involved the comparison of monthly rainfall forecasts made three months in advance for the period between 1993 and 2009 with the corresponding observed rainfall data. This comparison was conducted using time-series plots, Root Mean Squared Error (RMSE), and Pearson correlation coefficients. The results of comparing the RMSE values between the forecasts produced by the Australian

Bureau of Meteorology's Predictive Ocean Atmosphere Model for Australia (POAMA)-1.5 General Circulation Model (GCM) and the prototype model revealed that the prototype model had a reduced RMSE for 16 out of the 17 sites examined. The initial prototype design was regarded as preliminary, with the possibility for substantial enhancements, such as incorporating data from General Circulation Models (GCMs) and conducting experiments with alternative input qualities.

Wu et al., (2013) they employed several soft computing approaches for rainfall prediction. They utilized moving average (MA) and singular spectrum analysis (SSA) techniques for preprocessing. The modular model was composed of local support vectors regression (SVR) models or/and local artificial neural networks (ANN) models. Modular models involved preprocessing the training data into three crisp subsets (low, medium and high levels) according to the magnitudes of the training data, and finally two SVRs were performed in the medium and high-level subsets. And ANN or SVR was involved in training and predicting the low-level subset. The ANN-MA was displayed considerable accuracy in rainfall forecasts compared with the benchmark. Although, neural network should include backpropagation to increase the speed of the training process.

Mekanik et al., (2013) examined the use of Artificial Neural Network (ANN) and Multiple Regression (MR) analysis for the purpose of predicting long-term seasonal spring rainfall in Victoria, Australia. The researchers considered lagged El Nino Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) as viable predictors in their investigation. The MR models that did not exceed the thresholds of statistical significance were chosen for further predictions of spring rainfall. The artificial neural network (ANN) was created in the configuration of a multilayer perceptron employing the Levenberg-Marquardt (L-M) technique. The statistical evaluation of both the MR and ANN models involved the calculation of performance metrics such as Mean Square Error (MSE), Mean Absolute Error (MAE), Pearson correlation (r), and Willmott index of agreement (d). The multiple regression (MR) models exhibited limited generalisation capability for the eastern region of Victoria, as indicated by correlation coefficients ranging from -0.99 to -0.90. In contrast, artificial neural network (ANN) models demonstrated superior generalisation ability for both the central and western regions of Victoria, with correlation coefficients ranging from 0.68 to 0.85 and 0.58 to 0.97, respectively. The testing sets exhibited lower error rates in the case of artificial neural network (ANN) models as compared to multiple regression models.

The study conducted by Ganti (2014) involved an assessment of the accuracy of monthly monsoon rainfall predictions through the utilisation of artificial neural networks. Artificial neural networks (ANNs) have been offered as effective tools for the purpose of modelling and forecasting. The artificial neural network architecture employed in this study is a feed-forward multi-layer perceptron, which was trained using the widely-used back-propagation algorithm. Linear and non-linear regression models have been explored as alternative strategies for modelling monthly monsoon rainfall, with the aim of conducting comparative analyses. The dataset utilised in this research include monthly precipitation and monthly mean of the day maximum temperature within the north central area of India. A total of four regression models and two artificial neural network (ANN) models were built. The evaluation of the performance of different models involved the utilisation of a diverse range of conventional statistical parameters and scatter plots. The findings of the study indicated that the utilisation of Artificial Neural Networks (ANNs) for the purpose of predicting monsoon rainfalls yielded promising outcomes. The accurate monsoon rainfall projections have played a crucial role in properly managing India's economy and agricultural activities.

Gupta et al., (2014) study, they developed a rainfall time series prediction model using an artificial neural network (ANN), specifically a Multilayer Perceptron (MLP) network with backpropagation for training. Input parameters included discharge and rainfall data. Data pre-processing was conducted, and the model's sensitivity was investigated. The dataset was divided into training, validation, and test sets. The training set computed gradients to adjust network weights, the validation set monitored error during training, and the test set compared various models. Multiple neural network topologies were explored, assessing performance with mean absolute error (MAE), mean squared error (MSE), and correlation coefficient (CC). Results favored the MLP network for superior predictive accuracy over conventional models.

In a study conducted by Dubey (2015), various artificial neural networks were developed to forecast rainfall patterns in Pondicherry, an Indian coastal region. The artificial neural network (ANN) models were developed utilising three distinct training algorithms: the feed-forward back propagation strategy, the layer recurrent algorithm, and the feed-forward distributed time delay algorithm. The neuronal count for all models was maintained at 20. A total of twelve artificial neural network (ANN) models were trained using a dataset consisting of 800 samples from the region, which were collected over a span of one hundred years. Following the completion of the training, validation, and testing phases, the neural networks were evaluated using a dataset consisting of 200 samples for each phase. The mean squared error (MSE) was calculated for each model, and the feed-forward distributed time delay algorithm achieved the highest accuracy with an MSE score of 0.0083.

According to research by, Purnomo et al., (2017) devised an artificial neural network model with the purpose of predicting monthly rainfall rates. Two neural network models were suggested in order to anticipate the monthly rainfall rate. The evaluation of the suggested model was conducted using the monthly rainfall data in Ampel, Boyolali, spanning from 2001 to 2013. The experimental findings indicate that the first model exhibited significantly higher accuracy compared to the second model. The first model exhibits an average accuracy slightly over 98%, whereas the second model demonstrates an accuracy of roughly 75%. Moreover, both models exhibit improved performance in conditions of reduced rainfall variability.

Srivastava et al., (2017) studied artificial neural network (ANN) models to predict daily rainfall over a 4-month period in the Kumarganj area of Faizabad, Uttar Pradesh, India. The training data for the models consisted of a 4-month period from June 1st to September 30th, spanning the years 1986 to 2004. The models were then validated using data from the years 2005 to 2012. Sensitivity analysis was conducted to identify the primary parameter for rainfall prediction. The models were developed using a dataset consisting of twenty-seven years of daily data for vapour pressure, relative humidity, maximum temperature, and mean temperature as input parameters. The current day's rainfall was used as the output parameter for the models. The models were trained and tested using the sigmoid activation function. The models' performances were assessed through both qualitative means, such as visual observation, and quantitative means. The findings of the sensitivity analysis revealed that, in addition to rainfall, the vapour pressure emerged as the most significant input parameter in the context of rainfall forecasting for the specific research area.

The literature cited above highlights the importance of data normalisation in order to obtain trustworthy findings when utilising Artificial Neural Networks for rainfall forecasting. The process of selecting input variables is a critical aspect in artificial neural network (ANN) modelling. The existence of a universally applicable rule is absent, as the determination of an appropriate approach is contingent upon the specific problem at hand. Hence, the trial-and-error approach was employed to select input variables for artificial neural network (ANN) models. Moreover, a majority of researchers have determined that the Levenberg-Marquardt (LM) method and the logsigmoid transfer function are more effective in achieving successful execution of artificial neural networks (ANNs).

1.1.2 Literature review based on Rainfall-Runoff

The link between rainfall and runoff is of utmost significance for any given watershed. The nature of this relationship is contingent upon several parameters, including the attributes of precipitation, surface runoff, temperature, and infiltration. The elements discussed have a significant influence on the volume of runoff. The presence of a reliable link between rainfall and runoff allows for a more substantial increase in the available time for the local authority to develop appropriate decision-making strategies. Runoff is typically produced by precipitation events, specifically rainstorms, which are influenced by the attributes of the rainfall, including its frequency and volume. The correlation between rainfall and runoff is often employed in hydrological research and analysis. The utilisation of the rainfall-runoff model holds significant importance in the management and strategic planning of water resources (French et al., 1992).

The estimation of direct rainfall-runoff is conducted with efficiency; however, it is not feasible for the majority of locations within the needed timeframe. The runoff refers to the process of draining precipitation from a certain surface area, known as a catchment, into a canal. It represents the total amount of water discharged from the catchment within a specified timeframe. Prior to conducting an examination on runoff, some prerequisites must be met, including the presence of precipitation, evapotranspiration, initial loss, infiltration, and detection storage. Once these conditions are satisfied, any surplus precipitation will traverse the land surfaces via smaller channels known as overland flow, ultimately contributing to the formation of surface storage. This study examines the relationship between rainfall and runoff by analysing observed rainfall data in conjunction with observed stream flow data, as well as simulated runoff data for a certain year. In a similar manner, the output generated by the Artificial Neural Network (ANN) model will be utilised as a meteorological input in the Soil and Water Assessment Tool (SWAT) model, and afterwards executed for a specific time frame to conduct the simulation. Once again, the relationship between rainfall and runoff was established through the use of anticipated rainfall and simulated runoff. Extensive literature review has been conducted to fulfil the research purpose.

The Soil and Water Assessment Tool (SWAT) model is a commonly used semi-distributed hydrological model. It partitions a catchment into multiple sub basins, which are further divided into hydrologic response units (HRUs). These HRUs are defined by their specific combination of land cover, soil type, and slope class within each sub-basin. The integration of the SWAT model with the Geographic Information System (GIS) exhibits significant promise in effectively addressing spatial flood control methods. Furthermore, the

SWAT model is extensively utilised for the purpose of simulating runoff and water quality in response to altered circumstances. The SWAT model is a dynamic simulation tool that has a restricted scope in terms of its ability to accurately represent instantaneous hydrologic responses. In order to enhance the capacity of the SWAT model to replicate hydrological processes at sub-daily or sub-hourly intervals, the necessary adjustments were made largely to the model algorithms. These modifications were aimed at enabling the SWAT model to function at a more refined temporal resolution, while maintaining a continuous modelling loop. The current study incorporates the utilisation of the Soil and Water Assessment Tool (SWAT) to simulate stream flow, employing a range of meteorological parameters and gauge data as inputs. A comprehensive literature review has been conducted to achieve the research objective.

Spruill et al., (2000) conducted an evaluation of the Soil and Water Assessment Tool (SWAT) and examined the sensitivity of factors during the modelling of daily stream flows in a small watershed located in central Kentucky. The study spanned a duration of two years. The model was calibrated using streamflow data from 1996, while streamflow data from 1995 were utilised for evaluation purposes. The model demonstrated satisfactory prediction of the daily streamflow patterns within the specified time frame, despite the relatively low Nash-Sutcliffe R^2 values of 0.04 and 0.19 for the years 1995 and 1996, respectively. The exclusion of daily peak flow values from the months of August to December resulted in an enhancement of the daily R^2 to 0.15, a value comparable to the daily R^2 observed in 1996. The Nash-Sutcliffe coefficient of determination (R^2) for monthly total flows in the year 1995 was 0.58, while for the year 1996 it was 0.89. These figures align well with those reported in existing literature. A sensitivity analysis/calibration process was devised in order to assess the parameters believed to have an impact on the accuracy of stream discharge forecasts. The factors included in this study were the following: drainage area, slope length, channel length, saturated hydraulic conductivity, and available water capacity. The optimisation process aimed to minimise the average absolute divergence between observed and simulated stream flows, resulting in the identification of optimal values or ranges for each parameter. The findings of the study suggest that the SWAT model shown efficacy in accurately characterising the monthly runoff patterns observed in small watersheds located in central Kentucky.

The hydrologic evaluation of the lower Mekong River basin in Thailand was conducted by Rossi et al. (2009) using the Soil and Water Assessment Tool model. The study area encompassed a basin measuring 629,520 square kilometers. The SWAT model, specifically the

2003 version, was utilised to simulate the runoff inside the Mekong River basin. This basin has been partitioned into eight subareas, encompassing the regions upstream of Kratie, the vicinity of Tonle Sap, and certain portions of Vietnam. The parameters of the SWAT model were calibrated and validated for the gauged stream flows in the tributaries of the Mekong River over the time periods of 1985-1992 and 1993-2000, respectively. The statistical analysis conducted to assess the calibration and validation of the model revealed that the Nash-Sutcliffe efficiency (NSE) values, both on a monthly and daily basis, consistently fall within the range of 0.8 to 1.0 for all the primary monitoring stations. The application of the SWAT model to the Mekong River basin has been beneficial in establishing a hydrologic baseline for this expansive drainage area. The simulation conducted by the LMRB revealed the potential efficacy of the model as a tool for assessing water amount in the basin. The primary obstacle encountered in the process of modelling this particular watershed pertained to the substantial demands in terms of both time and computational resources.

Jain et al., (2010) studied the SWAT model in conjunction with ArcView to simulate the runoff and sediment production within a Himalayan watershed in India. The elevation of the watershed ranged from 600 to 3200 metres above mean sea level (MSL). The model underwent calibration for the time period spanning from 1993 to 1994, and subsequently underwent validation using observed data on runoff and sediment output for the years 1995, 1996, and 1997. The coefficient of determination (R^2) for the estimation of daily and monthly sediment yield during the calibration period was calculated to be 0.33 and 0.38, respectively. The coefficient of determination (R^2) for the daily and monthly sediment yield values from 1995 to 1997 was found to be 0.26 and 0.47, respectively.

Qui et al., (2012) conducted an evaluation of the Soil and Water Assessment Tool (SWAT) and examined its applicability for simulating runoff and sediment load in the Zhifanggou watershed, which is situated in a hilly-gullied area of China. The study utilised daily runoff and sediment event data spanning the period from 1998 to 2008. Specifically, the data from 1998 to 2003 were employed for calibration purposes, while the data from 2004 to 2008 were utilised for validation. The evaluation statistics pertaining to the daily runoff simulation indicated that the model outcomes were deemed satisfactory, albeit with a tendency to underestimate runoff during high-flow occurrences. The SWAT model had satisfactory performance in simulating sediment load, accurately capturing the overall trend. However, it consistently underestimated sediment load during both the calibration and validation periods. The discrepancy between the observed and simulated data is likely attributable to the inherent constraints of the SCS-CN and MUSLE methodologies employed in the model. The findings of the study suggest that adjustments to the components of the SWAT model are necessary in order to incorporate rainfall intensity and duration as factors that influence the model's accuracy in simulating peak flow and sediment load during periods of heavy rainfall.

In their work, Maharjan et al., (2013) investigated the prediction of runoff flow from a 0.8 ha field-sized agricultural watershed in South Korea. They employed the Soil and Water Assessment Tool (SWAT) sub-daily for this purpose. A series of SWAT sub-daily simulations were conducted, encompassing a total of 18 distinct rainfall events. These events were divided evenly, with 9 utilised for calibration purposes and the remaining 9 reserved for validation. The calibration method yielded a coefficient of determination (R^2) value of 0.88 and a Nash-Sutcliffe Efficiency (ENS) value of 0.88, indicating a strong overall trend and degree of matching simulated flow with measured data for the rainfall events in 2007-2008. The validation results indicated that the R^2 and ENS values were 0.9 and 0.84, respectively. The R^2 and ENS values obtained from the simulation results using daily rainfall data were found to be 0.79 and -0.01, respectively. These values were observed to exceed the permissible limits for the model simulation.

In their work, Patil et al., (2014) investigated the application of the SWAT hydrological model for runoff modelling of the Bhima River. The researchers employed automated calibration techniques, as well as conducted sensitivity and uncertainty analyses. Furthermore, they compared the results obtained from their approach with those obtained using the built-in auto-calibration tool of the SWAT model for parameter optimisation. The findings of the study demonstrated that the approach exhibited satisfactory performance and consistency in the exploration of a predetermined set of ideal parameters. The calibration and verification results demonstrated a high level of concordance between the simulated and observed data. The evaluation of model performance encompassed the utilisation of several statistical factors, such as the Nash-Sutcliffe coefficient and the normalised objective function. The coefficient of determination (R^2) during the Calibration phase was found to be 0.89, whereas in the Validation phase it was seen to be 0.74. The Nash-Sutcliffe Efficiency (NSE) values for calibration and validation were 0.81 and 0.77, respectively. The research shown that the SWAT model, when appropriately validated, may be utilised proficiently for evaluating management scenarios in watersheds. The utilisation of the SWAT model, in conjunction with Geographic Information System (GIS) technology, has demonstrated its efficacy as a versatile and dependable instrument for facilitating water-related decision-making processes.

The study conducted by Ahsan et al., (2015) involved the assessment of runoff in the Upper Betwa basin by the application of the SWAT Model. The primary aim of this study was to assess the efficacy of the Soil and Water Assessment Tool (SWAT) and determine its suitability as a simulator for runoff processes in the Berasia, Bhopal, Raisen, and Vidisha catchment areas within the upper Betwa basin. The hydrological and meteorological data were obtained from the Indian water site. The land use map of the studied area was obtained from the National Bureau of Soil Survey and Land Use Planning in Nagpur. Monthly surface runoff data for the monsoon months spanning from 1993 to 2002 were acquired for the regions of Berasia, Bhopal, Raisen, and Vidisha. The model underwent calibration and validation processes specifically for the monsoon seasons spanning from 1993 to 1999 and from 2000 to 2002, respectively. The evaluation of the model's performance involved the utilisation of statistical and graphical techniques to assess the model's ability to accurately simulate the runoff of the upper Betwa basin. The calibration period yielded coefficient of determination R^2 values of 0.97, 0.96, 0.94, and 0.98 for the locations of Berasia, Bhopal, Raisen, and Vidisha, respectively. The relative error values were calculated as 6.68, 8.00, 10.17, and 15.97, respectively. The Nash Sutcliffe model efficiency values derived for the locations of Berasia, Bhopal, Raisen, and Vidisha were 0.98, 0.97, 0.99, and 0.93, respectively. The validation period yielded R^2 values of 0.98, 0.97, 0.95, and 0.76 for the regions of Berasia, Bhopal, Raisen, and Vidisha, respectively. The relative errors for the relevant values were 6.77, 10.61, 7.91, and 10.56. The Nash-Sutcliffe model efficiency values found for monthly observed and simulated runoff in the locations of Berasia, Bhopal, Raisen, and Vidisha were 0.99, 0.99, 0.95, and 0.99, respectively. The calibration and validation findings demonstrated that the model accurately projected the total surface runoff at Berasia, Bhopal, Raisen, and Vidisha within the Upper Betwa basin. The model that underwent calibration and validation was employed to conduct assessments of water quantity and quality in the basin, encompassing both long-term analyses and evaluations of storm events.

According to research by, Sharma et al., (2015) conducted an evaluation of daily rainfall-runoff simulations in the Narmada River basin. They utilised the Soil and Water Assessment Tool (SWAT) model in the ArcGIS configuration for their analysis. The study utilised daily recorded data on rainfall and temperature spanning a period of 8 years. The model was then calibrated and validated using data from four gauge and discharge stations located at different points within the basin. This study involved doing model runs using three distinct combinations, each with varying time periods for calibration and validation. The study revealed that the performance of the model is satisfactory for gauge stations located upstream. However, the predicted runoff for downstream gauge stations exhibited a lack of agreement with the observed data. The study revealed that the ability to validate is contingent upon the specific rainfall parameters observed throughout the calibration years. The model's performance was superior in the first two instances of calibration and validation compared to the final combination. Experiments were conducted to investigate the influence of the number of hydrological response units (HRUs) employed in the simulations. This study emphasises the importance of utilising a significant number of years for the calibration of the SWAT model in order to enhance its accuracy in estimating rainfall-runoff within the entirety of the Narmada basin.

Pereira et al., (2016) conducted a hydrological simulation in the Pomba river basin, which spans an area of 8,600 square kilometres. The study focused on a region with a typical tropical climate and soil characteristics. The SWAT model was employed for the purpose of calibration and validation tests. The model underwent a process of calibration by trial and error from January 1996 to December 1999, followed by a validation phase from January 2000 to December 2004. The model was employed to simulate the maximum, average, and minimum annual daily stream flow using the paired t-test. This simulation contributes to the management of water resources in the region. However, it is important to note that the model still requires improvement, particularly in terms of accurately representing rainfall in order to provide more accurate estimates of extreme values.

Kangsabanik and Murmu (2017) conducted an investigation on the application of the Soil and Water Assessment Tool (SWAT) model for the purpose of rainfall-runoff modelling within the Ajay River basin. The delineation of the catchment area was accomplished by employing a Digital Elevation Model (DEM), which facilitated the division of said region into a total of 19 sub-basins. The landuse map was prepared using the IRS-P6 LISS-iii image, while the soil map was derived from the HWSD (Harmonised World Soil Database) Raster world soil map. The sub basins were subsequently partitioned into 223 hydrological response units (HRUs). The SWAT simulation was conducted using a dataset spanning 30 years of daily rainfall data and daily maximum and lowest temperature data. The purpose of the simulation was to determine the runoff corresponding to the rainfall on a daily, monthly, and yearly basis. The calculated coefficient of correlation (R) between rainfall and the related runoff for a specific period was determined to be 0.9419.

Khafaji et al., (2017) conducted an assessment of the efficacy of the SWAT model in simulating long-term runoff within the Al-Adhaim watershed in Iraq. The model underwent calibration using daily-measured flow data spanning from January 1, 1983, to September 31, 1984. Subsequently, validation of the model was conducted using data from January 1, 1985, to September 31, 1985. The statistical evaluation of the model's performance involved the utilisation of the coefficient of determination (R^2), Nash-Sutcliffe efficiency (NS), and standardisation root mean square error (RSR). The statistical parameters used to calibrate the runoff findings demonstrate an acceptable level of agreement between the daily values obtained from measurements and those generated by the model. The calibration (validation) process yielded R^2 values of 0.76, 0.75, and 0.5 (0.71, 0.69, and 0.55) for R^2, NS, and RSR, respectively. Subsequently, the calibrated model was employed to evaluate the relationship between runoff and changes in land cover/land use over an extended period from 1986 to 2013. Consequently, the model was executed on two separate occasions. The first run used a single land cover/land use (LC/LU) dataset from 1992 (referred to as Case 1). The second run, referred to as Case 2, incorporated two LC/LU datasets from 1992 and 2001. The computed statistical value was used to compare the two examples, and the performance of the SWAT model was examined for a long-term period. The findings indicate that there was no observed impact of land cover/land use change on the simulation of runoff in the studied region. The correlation coefficients for the runoff of case 1 and case 2 were found to be 0.99, 0.98, and 0.02 for R^2, NS, and RSR, respectively. Hence, the SWAT model can be employed to identify land cover and land use (LC/LU) alterations across an extended simulation period within the Al-Adhaim watershed.

According to research by, Sowmiya and Arul (2017) employed the Soil and Water Assessment Tool (SWAT) model to conduct a simulation of runoff inside an agricultural watershed. This work utilised QSWAT, a QGIS interface for the Soil and Water Assessment Tool (SWAT), to conduct sub-basin modelling and estimate runoff and related parameters. The data was collected over a span of 15 years, specifically from 2000 to 2014. The model was then calibrated and validated using the SUFi2 calibration technique, utilising the observed data from these 15 years. The model's performance was assessed using two metrics: the Nash-Sutcliffe Efficiency (NSE) and the Coefficient of Determination (R^2). The model performance coefficient yielded moderately average outcomes for both the calibration and validation phases.

According to research by, Mistry and Joshi (2018) employed the Soil and Water Assessment Tool (SWAT) model to conduct rainfall runoff modelling of the Shakkar River.

The delineation of the catchment area was accomplished by the utilisation of a Digital Elevation Model (DEM), which facilitated the division of said region into a total of 23 sub-basins. In order to create a land use map, the LANDSAT pictures were acquired from the earth explorer database, while the soil map was obtained from the National Bureau of Soil Survey (NBSS). The sub basins were subsequently partitioned into 223 Hydrological Response Units (HRUs). A SWAT simulation was conducted on a daily basis using 21 years of daily rainfall data and daily maximum and minimum temperature data in order to determine the corresponding runoff. The calculated coefficient of correlation (r) between rainfall and the associated runoff for a specific period was determined to be 0.8019.

Based on a comprehensive analysis of existing scholarly works, it has been acknowledged that the Soil and Water Assessment Tool (SWAT) model holds a prominent position among hydrological models that are widely employed to tackle hydrologic and environmental challenges. The inclusion of soil and land use/land cover information is of paramount significance in the creation of Hydrologic Response Units (HRUs), a critical component for stream modelling, as highlighted by numerous researchers. One of the primary challenges faced by SWAT users pertains to the accessibility of dependable data, as such data may either be inaccessible without cost or subject to restrictions on public access in certain countries. Therefore, it is imperative to prioritise future research endeavours that focus on the identification and development of dependable input data for SWAT models.

The study conducted by Harun et al., (2002) focused on investigating the application of an artificial neural network model to analyse the relationship between rainfall and runoff in the Sungai Lui watershed in Malaysia. The study introduced an artificial neural network (ANN) model as a means of predicting daily runoff. The model utilised rainfall data as input nodes. The method of selecting input nodes using a factor of 10 and 5 was implemented. Additionally, a comparison was made between the Artificial Neural Network (ANN) and the Hydrologic Engineering Center's Hydrologic Modelling System (HEC-HMS) model. The study revealed that the artificial neural network (ANN) models shown strong generalisation capabilities in predicting the rainfall-runoff relationship, surpassing the performance of the HEC-HMS model.

According to research by, Aytek et al., (2008) conducted a comparison between rainfall-runoff calculations utilising Gaussian process (GP) models and two alternative artificial neural network (ANN)-based models. The outcomes derived from GP models exhibit comparable levels of success to those achieved through the utilisation of ANN approaches.

These findings affirm the efficacy of the GP approach in offering a valuable tool for addressing distinct hydrological challenges, such as rainfall-runoff estimation. According to research by, Kalteh (2008) employed artificial neural networks (ANNs) to conduct rainfall-runoff modelling with the aim of enhancing modelling capabilities and gaining insights into the underlying processes. The findings suggest that artificial neural networks (ANNs) show promise as effective tools for accurately modelling intricate processes. Additionally, ANNs offer valuable insights derived from the learnt associations, which aid in the modeller's comprehension of the investigated process. Furthermore, ANNs contribute to the evaluation of the model's performance.

Welderufael et al., (2008) conducted a quantitative analysis of the rainfall-runoff correlations specifically on the Dera Calcic Fluvic Regosol ecotope located in Ethiopia. Rainfall-runoff measurements were conducted in the years 2003 and 2004 on plots measuring 2 m × 2 m. These plots were equipped with a runoff measuring device and were duplicated three times for each treatment. The study consisted of two distinct treatment groups, namely conventional tillage (CT) and no-till (NT). Measurements of rainfall-runoff were conducted throughout the rainy seasons of 2003 and 2004, specifically focusing on 25 instances of precipitation above 9 mm. There was no statistically significant difference observed in the runoff between the two treatments. The runoff (R) observed during the two rain seasons, represented as a ratio of the rainfall recorded during the measurement period (P), denoted as R/P, yielded values of 0.46 and 0.39 for the NT and CT treatments, respectively.

Sepaskhah and Fard (2010) investigated the correlation between rainfall and runoff by considering the soil's physical qualities. The purpose of their research was to provide valuable insights for the design of micro catchment water collecting systems. A subroutine has been developed to estimate the daily and average annual runoff within a computer model specifically designed for micro catchments. The subroutine made the assumption that all abstractions originate from infiltration. A methodology was then devised to calculate ponding duration and infiltration by utilising data from a recorded rain gauge and the physical parameters of the soil. This methodology was based on the GreenAmpt infiltration equation. The subroutine that was built yielded a daily micro catchment runoff coefficient of 0.0737 in the designated study region, which closely aligned with the observed value of 0.080. The subroutine that was built calculated that a daily threshold rainfall of 6.5 mm is required to induce daily runoff, while the actual value of rainfall was 4.6 mm. The subroutine that was built yielded a micro catchment average annual runoff coefficient of 0.0894 in the designated study region, which closely aligned with the measured value of 0.0875. The subroutine developed in this study was used to predict the yearly threshold rainfall required to generate annual runoff. The estimated value obtained was 158.8 mm, while the measured value was 106.5 mm. The model utilised the predicted correlation between annual runoff and rainfall to estimate the micro catchment area. The study revealed that the implemented subroutine successfully established the correlation between daily and annual runoff and rainfall. This correlation is crucial for including into the model used for designing the micro catchment area.

Raji et al., (2011) conducted an investigation to analyse the characteristics of rainfall-runoff in a specific area measuring 1.23 hectares. This area was chosen due to its notable occurrence of runoff throughout the rainy season. Measurements were conducted to determine the runoff from seven storms. A rectangular notch was utilised for this purpose. Additionally, a relationship between the discharge and the matching head for the notch located at the downstream end of the research area was computed. The calculated R^2 value was 0.98, indicating a strong correlation between the variables. Additionally, the runoff coefficient for the research area was determined to be 0.12, suggesting a relatively low proportion of rainfall that contributes to surface runoff in the area. The unit hydrograph for the examined area was generated by considering numerous storm hydrographs, specifically focusing on the storms Pl and P7. A linear relationship between rainfall and runoff was determined, with the equation $Y = 0.2X - 0.85$. The coefficient of determination (R^2) for this relationship was calculated to be 0.98. The findings of the study indicate that the implementation of rainwater recharge structures, designed using rainfall-runoff studies, resulted in a notable improvement in the water table level within the study region. The utilisation of the generated rainfall-runoff relation and representative unit hydrograph is advantageous for the design of rainwater harvesting and recharging systems in the investigated region.

Lakhote et al., (2014) conducted an analysis of rainfall runoff specifically for the Vena catchment in the Hinganghat District. The study assessed the rainfall runoff characteristics of an agricultural watershed area by analysing data from eight rain gauge stations. The study involves the examination of observed runoff in relation to three models: the SCS model, the modified SCS model, and the Mockus model.

Tandon and Nimbalkar (2014) conducted a study on the links between rainfall and runoff, employing the curve number approach. The study examined the five approaches for establishing the relationship between rainfall and runoff in the Morbe dam catchment watershed. Additionally, it was recommended to consider the curve number technique or the

Inglis method by modifying the used factor. The rainfall and runoff data spanning from the year 1958 to 2011 were employed for a comprehensive analysis. This study examined the relationship between rainfall and stream flow on gauged, small watersheds within the Morbe dam watershed project. Various approaches were employed to estimate runoff in order to assess the reaction of stream flow to precipitation. Calibrated curve numbers inherently possess considerable errors, necessitating the application of statistical methods to establish the satisfactory agreement between estimated runoff and observed data. The estimation of runoff was necessary to ascertain and predict its impacts.

In a study conducted by Bhagat (2016), a rainfall-runoff co-relationship was constructed for the lower Mahi basin in India. The rainfall data pertaining to the significant rain gauging sites was taken into account for a duration of ten years. The SCS-CN method was employed to compute the runoff depth. By employing the given parameters, a correlation between rainfall and runoff was established through the use of a regression equation. The study revealed that a significant correlation was seen between rainfall and runoff, leading to favourable outcomes. The quality of the data from which the regressions were derived had an impact on them to a certain degree. The evaluation of runoff data was deemed satisfactory, given it does not incorporate the infiltration occurring inside the catchment area. The correlation coefficient (r) for the basin approached unity, indicating a strong positive relationship. This high correlation suggests that the catchment exhibited a heightened responsiveness to the precipitation it received.

Parmar et al., (2016) conducted a study to create a rainfall-runoff connection in order to estimate runoff in the semi-arid catchment of Ozat River, which is situated in the Junagadh district of Gujarat, India. The estimation of the weighted curve number (73.03) for the research area was conducted by utilising land use/land cover maps produced from remote sensing and GIS techniques. The correlation coefficients of 0.657 and 0.860 were used to establish the correlations between daily rainfall and daily observed and calculated runoff, respectively.

Singh and Purty (2016) conducted an estimation of the link between rainfall and runoff in the East Singhbhum district of Jharkhand, India. The present study involved an analysis of the interrelationship between climatic parameters in the East Singhbhum district. Rainfall data from the years 2001 to 2013 was compiled on an annual basis to extract runoff data. Additionally, temperature data from the years 1991 to 2002 was collected to examine the correlation between temperature and runoff in the study area. Correlation analysis was employed to investigate the long-term trend of the hydrological time series, encompassing temperature, rainfall, and runoff. The impact of rainfall and runoff patterns on human activities is significant, making them crucial climatic parameters in the context of storm water management. The researchers employed several statistical analysis techniques to examine the correlation between meteorological factors, including rainfall, infiltration rate, and temperature, and the phenomenon of runoff. The findings of the study indicated a significant association between runoff and infiltration, as well as a substantial correlation between rainfall and runoff. However, the correlation between temperature and runoff was shown to be minimal. Hence, the significance of climatic factors, namely temperature, precipitation, and runoff, within the examined region was substantial.

According to research by, Kumari et al., (2019) conducted an estimation of the rainfall-runoff relationship in the east Singhbhum District of Jharkhand. Additionally, they examined the link between runoff and factors such as infiltration capacity and temperature. The purpose of this study was to examine the relationship between yearly precipitation and yearly water discharge over the time period spanning from 1901 to 2018. The results of the graphical analysis revealed a significant positive correlation between the yearly rainfall and runoff values. The Pearson correlation coefficient (R) exhibited a value close to +1, indicating a strong positive correlation between the variables, suggesting that they tend to move in the same direction. Additionally, this indicates that the two variables under consideration exhibit a perfect positive correlation, indicating a significant association between them. The strength of the linear association increases as the value of r approaches +1. Positive values in this context signify a positive correlation between rainfall and runoff, indicating that when rainfall levels rise, runoff levels likewise increase. The correlation exhibited a strong positive relationship as the value of r approached +1. The analysis of the rainfall-runoff relationship in the east Singhbhum district revealed a strong correlation between these two variables, with a coefficient of determination (R-squared) value of 0.953.

Tazyeen et al., (2019) investigated the correlation between rainfall and runoff in the Hulimavu watershed. The expansion of urban activity in the vicinity of the lake has resulted in the encroachment of its area, leading to a reduction in its surcharge storage capacity. Consequently, this has caused downstream areas to get inundated. The study region details were obtained from the SOI topographic map, while data was received from the Indian Meteorological Department (IMD), Karnataka State Remote Sensing Applications Centre (KSRSAC), and National Bureau of Soil Survey and Land Use Planning (NBSSLUP) for the research. The estimation of runoff was conducted using the Soil Conservation Service Curve

Number (SCS-CN) approach. The rainfall runoff connection was established in order to understand the interconnectedness between rainfall and runoff, thereby providing a foundation for conducting more hydrological investigations.

In a study conducted by Zeberie (2019), a model was created to analyse the relationship between rainfall and runoff in the BigAkaki watershed, located in the upper Awash basin of Ethiopia. The SCS curve number was employed for the purpose of estimating runoff from the surface of the basin. Additionally, the Soil and Water Assessment Tool (SWAT) was utilised to outline the basin and examine several factors such as the slope of the watershed, soil characteristics, and land uses. Furthermore, data pertaining to the classification of land use and land cover (LULC) was collected through the implementation of ground control stations, interviews, and field observation. Furthermore, the process of model calibration was conducted between the years 1991 and 1998, while model validation took place from 1999 to 2004. These activities were carried out specifically for the purpose of assessing the accuracy and reliability of the monthly flow measurements at the Akaki measuring station. The drainage area of the Big-Akaki watershed measures 971,849 square kilometres. The simulation was conducted by partitioning the watershed into 33 sub-basins and allocating a hydrological response unit (HRU) to each sub-basin, following the criteria outlined in the definition of multiple HRU. The SWAT model for the Big-Akaki watershed was calibrated and validated, and subsequently used to simulate the period from 1991 to 2016 on a monthly basis. This simulation aimed to examine the correlation between rainfall and runoff within the watershed. The findings of the study suggest that the Soil and Water Assessment Tool (SWAT) has satisfactory performance in simulating runoff, as evidenced by the evaluation of three key objectives: Nash-Sutcliffe Efficiency (NSE), coefficient of determination (R^2), and ratio of the root mean square error to the standard deviation of observed data (RSR). During the calibration period, the NSE, R^2, and RSR values for surface runoff were determined to be 0.81, 0.82, and 0.44, respectively. Similarly, during the validation period, same values were found to be 0.77, 0.77, and 0.48, respectively. The Big-Akaki basin exhibits an annual average precipitation of 1183.56 mm, while the surface runoff measures 227.63 mm. The findings of the study demonstrated a positive correlation between precipitation levels and the amount of water runoff observed on the Earth's surface.

In a study by Hu et al., (2020), a hybrid AI-based model combining ANNs and support vector regression (SVR) was proposed to predict streamflow in a large watershed. The results showed that the hybrid model outperformed other traditional models, including a single ANN model and a rainfall-runoff model based on the Soil and Water Assessment Tool (SWAT).

Another study by Wu et al., (2019) applied a deep neural network (DNN) model to predict streamflow in the Yellow River Basin in China. The DNN model was trained using long short-term memory (LSTM) neural networks, which are designed for time-series forecasting. The results showed that the DNN model outperformed other traditional models, including the SWAT model and an autoregressive integrated moving average (ARIMA) model.

In a study by Chen et al., (2021), a convolutional neural network (CNN) model was proposed to predict streamflow in a small watershed. The CNN model was trained using time-series data of rainfall and streamflow, and the results showed that the model outperformed traditional models, including the SWAT model and an ANN model. Genetic algorithms (GAs) have also been applied in conjunction with ANNs to optimize model parameters and improve the accuracy of rainfall-runoff models.

In a study by Zhang et al., (2020), a hybrid GA-ANN model was proposed to predict streamflow in a mountainous watershed. The results showed that the hybrid model outperformed other traditional models, including the SWAT model and an ANN model.

V. Kumar et al., (2020) introduced the use of the artificial neural network (ANN) in estimating the runoff of a river is popular among hydrologists and scientists for a long time. The classical gradient descent algorithm (GD) is the most commonly used algorithm for training the ANN runoff models so far. The achievement of the GD algorithm, however, is affected by changes to get stuck at the local minimum. In the paper, one of the popular evolutionary optimization algorithms, known as particle swarm optimization (PSO), has been explored to train the ANN rainfall-runoff model. However, the accurateness of the models using evolutionary algorithms is low.

Asadi et al., (2019) Proposed Inputs from the hydrogeomorphic and biophysical period series, including the Normalized Vegetation Difference Index (NDVI) and Connectivity Index were evaluated in addition to climatic and hydrological inputs. They used selected inputs for the creation of Artificial Neural Networks (ANNs) in the catchment of the Haughton River and the catchment of the Calliope River, Queensland. Results show that IC is incorporated as a hydrogeomorphic parameter and NDVI. Besides, the availability of remote sensing land use datasets (i.e., NDVI) and high-resolution feature maps to reflect geomorphologic catchment features may also be a limiting factor.

Roshni, et al., (2019) proposed the hydrological implementation of the Artificial Neural Network (ANN) and the Emotional Neural Network (ENN) for the simulation of rainfall-runoff in some order, Bihar as the area experiences flooding due to heavy rainfall. ENN is a revised version of ANN since it contains neural parameters that improve the process of network learning. The selection of inputs for the rainfall-runoff model is a key task. The paper makes use of a cross-correlation study to recognize possible predictors. Principal Component Analysis (PCA) was then conducted for the collection of data sets showing key patterns on the chosen data sets. The data sets collected after PCA were then used in the creation of the models ENN and ANN. Moreover, tremendous work is needed to estimate the reliability and performance of the models.

Üneş et al., (2019) implementation of the Rainfall and Runoff Partnership is very important for the effective use of water supplies and flood prevention. Nowadays various methods of artificial intelligence techniques are used to assess the relationship between rainfall and runoff. The present research utilizes Artificial Neural Networks (ANN). Popular approaches like Multiple Linear Regression (MLR) are also used. The data collected from the US have been integrated into the report. The Multi-Linear Regression and Feed-Forward Back-Propagation Artificial Neural Network models used 731 daily rainfalls, runoff, and temperature data to produce input data. However, lack of evaluation of watershed characteristics create issues in deciding the relationship between rainfall-runoff results.

Chadalawada et al., (2020) proposed a novel model building algorithm using the full potential of flexible modeling frameworks by searching for model space and using machine learning-based model configurations. The proposed algorithm for machine learning is based upon an evolutionary approach to computation using genetic (GP) programming. Up until now, state-of-the-art GP modeling implementations have used the algorithm as a short-term forecasting tool that produces an expected future time series very similar to neural network implementations. However, there are no sufficient observations were provided in the catchment area.

Minglei et al., (2020) developed an explored version of LSTM for streamflow prediction using historical data for forecasting for a particular period. The Kelantan River in the northeast region of the Malaysia Peninsula was selected for their case study. The feasibility of applying the developed LSTM model to streamflow prediction was verified, and the performance of the developed LSTM model was compared with the classic backpropagation neural network model. In the experimental process of applying the LSTM model to the prediction of streamflow, the influence of the training set size on the performance of the developed LSTM model was tested. Also, the effect of the time span of the prediction data on the performance of the developed LSTM model was tested. The bionic intelligent algorithm should be introduced into the LSTM model for better prediction.

Chen et al., (2019) they conducted uncertainty analysis on the hybrid double feed forward neural network (HDFNN) model for generating the sediment load prediction interval (PI). They used the LUBE technique, in which the lower and upper bounds are created directly as outputs of neural network-based models. Partitioning analysis demonstrates that the HDFNN model consistently performs well in creating PI for low, medium, and high loads. However, the LUBE approach does not solve highly nonlinear, discontinuous, and non-differentiable optimization problems.

Amir Mosavi et al., (2018) they analysed the state of the art of Machine learning models in flood prediction. They investigated to provide an extensive overview on the various ML algorithms used in the field through a qualitative analysis of robustness, accuracy, effectiveness, and speed. They found the effective strategies are hybridization, data decomposition, algorithm ensemble, and model optimization. However, hybridization method expensive costing up to five times the value of the normal process. Data decomposition method, ensemble algorithm, and model optimization technics suffer if not provided with the normal requirements.

Taormina et al., (2015) they tested the suitability of the LUBE approach in producing PIs at different confidence levels (CL) for the 6 h ahead streamflow discharges of the Susquehanna and Nehalem Rivers, US. The Particle Swarm Optimization (PSO) in LUBE applications, variants of this algorithm had employed for CWC minimization and they found to vary substantially depending on the chosen PSO paradigm. The algorithm has low convergence rate in the iterative process so that advanced algorithm should be utilized for prediction.

Chau et al., (2013) studied in neural network backpropagation should include for increase the speed of the training process. From the aforementioned issues, it is essential to develop a new framework for the rainfall-runoff model for enhanced prediction.

Based on the aforementioned reviews, it can be determined that several methodologies have been employed for the purpose of modelling rainfall-runoff relationships. The validation of the rainfall-runoff connection can be achieved by the assessment of many statistical measures, including the correlation coefficient (r), coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), index of agreement (IA), and volumetric error (VE). A correlation coefficient approaching unity is indicative of a strong relationship, suggesting that the catchment is very responsive to runoff.

Table 1.1 Summary of past literature works

S.NO	RESEARCHER'S	Primary objective	Major findings
1	Willems, 2000	classifications of hydrological models	Empirical models, conceptual or grey box models, and physically-based or white box models
2	Meresa et al. 2017	hydrological models	depending on the input data, hydrological parameter and structure of the model
3	Meresa and Gatachew (2018)	Comparison of hydrological model	compared three conceptual hydrological models for climate change impact study, and found that accuracy of the modeled flow is mainly depends on the model structure and number of model parameters
4	Yaghoubi and Massah (2014)	Comparison of hydrological model	In an investigation within the Azam Harat river catchment in Iran, three hydrological models—HBV, IHACRES, and HEC-HMS—were assessed and compared. Among these models, the HBV model demonstrated superior performance by providing a more accurate representation of river flow in terms of both mean and variability. Conversely, the HEC-HMS model displayed the weakest performance, as evidenced by higher

			root mean square values, indicating less accurate predictions of river flow characteristics.
5	Asati and Rathore (2012)	Develop a model	developed three models—a time series autoregressive model, an artificial neural network (ANN), and a multiple linear regression (MLR) model—to address the intricate behavior of a catchment characterized by a non-linear connection between rainfall and runoff. These models were compared without taking into account the inherent characteristics of the process.
6	Dastorani et al. (2009)	Compare of model	The study involved a comparison between artificial neural networks (ANNs) and several other data-driven models to reconstruct observed flow data. The findings indicated that ANNs outperformed the other models, including the conventional ratio and correlation methods, demonstrating their superiority in this context.
7	Yang et al., 2009;	Understanding of hydrological phenomena using of ANN and SVM	The advancement of artificial intelligence methods like Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs), among others, has brought about a substantial transformation in the forecasting of hydrological events.

8	Bafitlhile and Li, 2019	Details of computer-based model	Various data-driven models, such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Random Forest (RF), Fuzzy Rule-Based Systems, Model Trees (MT), Long Short-Term Memory (LSTM), Extreme Learning Machines (ELM), and others, are employed for prediction and classification purposes.
9	Gizaw and Gan, 2016	Comparison between ANN and SVM	For limited discharge data set available, Support Vector Machine (SVM) method based on machine learning techniques perform better or as par as artificial neural network models
10	Zhou et al., 2019	Comparison between RF and SVM	random forest (RF) method is another machine learning, decision tree-based method requiring less user defined parameters than SVM and often performing better in flood forecasting
11	Bhattacharya and Solomatine, 2005	M5P tree model	The M5P model tree is a machine learning technique that has seen increasing use in the field of water resources, particularly for predicting streamflow in recent times. It functions as a tree-based model, similar to linear regression, where the parameter space is divided into subspaces. Within each of these subspaces, a localized linear regression model is constructed,

			making it a versatile and adaptable tool for streamflow prediction.
12	(Tiwari et al., 2021).	compare between MP5 and ANN	M5P method also performed better in short-time flood forecasting than Artificial Neural Network (ANN) method
13	Taghi Sattari et al., 2013	Uses of MP5 MODEL	M5P has been used both in low flow forecasting as well as daily flow forecasting with less computation time and reasonable accuracy
14	Morid et al. (2010)	used the ANN and SWAT models	Their research revealed that, during the summer, fall, and winter seasons, Artificial Neural Networks (ANN) exhibited superior performance compared to the SWAT model when it came to predicting fundamental flow rates. However, in the spring season, characterized by frequent rainfall events and sudden peak flow rates, the SWAT model demonstrated better modeling capabilities.

It has been reported by many researchers that hydrological models are mainly depending on the input data, hydrological parameter and structure of the model (Meresa et al. 2016; Meresa et al. 2017; Meresa and Gatachew 2018). Particularly, studies on river modeling in ungauged catchment using the climate and physiographic characteristics are possible only if detailed information about topography, land use, soil, vegetation, and climate are depending on available data (Gunter and Bloschl, 2005; Wale et al., 2009; Adib et al., 2010; He et al., 2011). Runoff response estimation from ungauged river catchments is currently a topical issue in hydrology and water resources management (Gunter and Bloschl 2005; Wale et al., 2009; Adib et al., 2010; He et al., 2011) and in developing countries for hydraulic infrastructure construction.

Nowadays, there are several studies performed the rainfall and runoff process simulation using empirical, data driven, hydrological model and statistical models comparisons. Meresa and Gatachew (2018) compared three conceptual hydrological models for climate change impact study, and found that accuracy of the modeled flow is mainly depends on the model structure and number of model parameters. Yaghoubi and Massah (2014) compared three models of HBV, IHACRES and HEC-HMS in Azam Harat river catchment in Iran. Among these models HVB model performed better in proved resinable river flow in mean and variability whereas HEC-HMS exhibited worst performance in root mean square value. Asati and Rathore (2012) developed an autoregressive model, ANN and MLR for a complex catchment behavior which is non-linear relationship between rainfall and runoff, which is compared without incorporating the nature of process. Dastorani et al., (2009) compared artificial neural network with various data driven models for rebuilding the observed flow data and they concluded the ANN were dominant in comparison to other models (the normal ratio and correlation methods).

Moreover, in recent decades, the development of artificial intelligence techniques, such as Artificial Neural Networks (ANN), Support Vector Machine (SVM) and more, have provided a significant evolution in the predictors of hydrological phenomena (Yang et al., 2009; Kisi et al., 2009; Kocabasa et al., 2009; Kisi and Cigizoglu, 2007). Mathematically, the SVM is used for both classification and regression algorithms, which are formulated through the principles of statistical learning theory by Vapnik (1995). Due to the wide capability of the SWAT and SVM model regarding water and soil research studies, many studies have been performed all over the world by these models separately (Shepherd et al., 1999; Spruill et al., 2000; Saleh and Du, 2004; Birhanu et al., 2007; Gassman et al., 2007).

These advantages and increased computing power of computers have propelled the research and use of data driven models for discharge prediction. These methods are particularly useful in rivers basin for predicting floods, and drought sequence, where scarce data records and resource limitations hinder the use of physical based models. The data driven models used for prediction and classification are Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Fussy rule-based system and Model Trees (MT), Long Short-Term Memory (LSTM), Extreme Learning Machines (ELM) and so on (Bafitlhile and Li, 2019; Adnan et al., 2021). For limited discharge data set available, Support Vector Machine (SVM) method based on machine learning techniques perform better or as par as artificial neural network models (Gizaw and Gan, 2016; Liong and Sivapragasam, 2002). In classification and regression problems, SVM is widely used, developed by Cortes and Vapnik, 1995. SVM theoretically reduces the expected error of a learning model, and minimizes the problem of over fitting. The details of SVM are well described in the literature and can be found elsewhere (Cortes and Vapnik, 1995; Tongal and Booij, 2018; Wang et al., 2006; Yu et al., 2017; Yoon et al., 2011). Similarly, random forest (RF) method is another machine learning, decision tree-based method requiring less user defined parameters than SVM and often performing better in flood forecasting than the latter (Yu et al., 2017; Zhou et al., 2019). RF was used in flood hazard prediction in Dongxiang River (Sadler et al., 2018) and for drought prediction in Australia (Deo and Şahin, 2015). More details about RF can be found in many published literatures (Breiman, 2001).

In general, it seems HEC-HMS and RF are the most widely applied to predict discharge in river catchments. That is why in this research work, the comparison was done using these two hydrological and data driven model. Due to the reason that there are no previous studies in M.H. Halli station that are focused on discharge forecasting. This research work provides innovative research approach and robust solutions in discharge estimation. The few available meteorological data often present significant gaps. This makes the research work very innovative and original in terms of study area, methodology and framework approach.

Generally, river discharge models are designed to gain a better understanding of the hydrologic characteristics of a catchment and to generate a synthetic hydrologic data for river flow facility design like flood protection, water resources planning, mitigation of contamination, or for flood early warning and forecasting.

Overall, the literature review suggests that the use of AI and ANNs in rainfall-runoff modeling has the potential to improve the accuracy and efficiency of these models. The application of AI and ANNs in conjunction with other optimization techniques, such as GAs, has also shown promising results in improving the performance of rainfall-runoff models. However, more research is needed to further validate the use of these techniques in real-world applications and to address the challenges associated with their implementation, such as overfitting and the selection of appropriate input variables and model parameters.

This chapter gives a detailed overview of the available model. The next Section details the evolution in the rainfall runoff model. Section 2.2 details the literature gaps. Section 2.3 provides the summary of the chapter. The next section details the literature gap.

1.2 Summary

This chapter serves as an in-depth exploration of the various rainfall runoff models that are currently available. From this extensive pool of models, four distinct ones have been identified for closer examination and inclusion in this research work. These selected models are HEC-HMS, Random Forest, SWAT, and the M5P Model. The choice to focus on these particular models reflects a deliberate and considered selection process, taking into account their relevance and potential applicability to the research objectives.

The subsequent chapter delves into the research methodology in finer detail. It outlines the systematic approach and framework that will be employed to investigate and analyze the performance and effectiveness of these chosen models. This chapter serves as a crucial bridge between the theoretical understanding of the models and the practical implementation of the research, providing a roadmap for the entire investigative process. It will offer insights into how these models will be employed, the data sources utilized, and the specific criteria for evaluating their performance, ensuring transparency and rigor in the research process.