Enhancing Runoff Estimation in the Pra River Basin, Ghana, using Remote Sensing and LSTM-based Deep Learning Modeling

Nicholas Junior Appiah

Supervised by:

Prof. Dr. Edzer Pebesma (ifgi, WWU)

Co-Supervised by:

Prof. Dr. Marco Painho (NOVA IMS, UNL)

Dr. Edward Ackom (external, University of Ghana)

In partial fulfillment of the requirements for the Degree of Master of Science in Geospatial Technologies

Westfälische Wilhelms-Universität Münster (WWU), Germany
Universidade Nova de Lisboa (UNL), Portugal
Universitat Jaume I (UJI), Spain

Münster, Germany

February, 2024

Declaration Of Academic Originality

I hereby confirm that the entirety of this thesis is my own creation, except where duly attributed. All sources utilized, including text, images, and data, have been appropriately referenced. Any assistance received during the research and writing process has been acknowledged. Furthermore, I assert that this thesis has not been presented for any other academic qualification elsewhere.

APPIAH, Nicholas Junior

Münster, Germany

February 28, 2024

Acknowledgment

I extend my heartfelt gratitude to the coordinators of the Erasmus Mundus Joint Master Geospatial Technologies program for establishing an exceptional academic platform that fosters intellectual growth and provides a supportive international environment for students.

I am deeply grateful to my supervisors, Prof. Dr. Edzer Pebesma, Prof. Dr. Marco Painho, and Dr. Edward Ackom, for their steadfast support and invaluable guidance throughout my research journey. Special thanks are extended to Dr. Ackom for his contribution to the thesis topic, idea, and data. I would like to express my gratitude to Dr. Collins Owusu and George Kwabena Fordjour for their invaluable assistance in implementing the ideas presented in this thesis.

My heartfelt thanks go to my mother, whose unwavering support has been a constant source of motivation, and to my grandparents for their enduring inspiration. Special appreciation to my late father, Nicholas Appiah, whose love for education and constant encouragement to pursue higher degrees has brought me to this point.

Lastly, I wish to acknowledge the dedication of all the esteemed professors in the master's program for their tireless efforts in impacting knowledge and shaping the academic experience of students.

Abstract

Estimating runoff is vital for managing water resources, especially in areas where water is essential for agriculture, industry, and households. Traditionally, this involved using data from rain gauge stations and complex mathematical equations based on physical principles. However, this study explores advanced modeling techniques integrating satellite observations, particularly in regions lacking historical hydrological data. By employing in-situ precipitation data and various remote sensing products such as CHIRPS precipitation, MOD09GA_006 NDVI, MOD11A1 V6.1 LST, and MOD16A2GF.061 ET from 2001 to 2020, a Fully Connected Long Short-Term Memory (FC-LSTM) model was developed to predict monthly runoffs in the basin. Compared to the Soil and Water Assessment Tool (SWAT) model, the LSTM model outperformed, achieving higher accuracy with an R^2 of 0.94 and Nash-Sutcliffe Efficiency (NSE) of 0.93, while SWAT obtained R^2 and NSE values of 0.78 and 0.75, respectively. Additionally, integrating remote sensing variables enhanced model performance, underscoring the potential of remote sensing in hydrological studies. These findings advance our understanding of runoff estimation methods and underscore the importance of LSTM models and remote sensing technology in hydrology.

Keywords: runoff estimation, LSTM, remote sensing, Pra River Basin

Table of Content

List of figures

List of tables

Chapter 1

Introduction

1.1 Background

Effective water resource management is a crucial element in fostering sustainable development, particularly in regions heavily dependent on agriculture and grappling with challenges related to scarce water availability, drought, and flooding (Agarwal & Global Water Partnership, 2000). In this complex scenario, successful water resource management goes beyond mere allocation, necessitating a thorough understanding of hydrological systems and their intricate dynamics.

The modeling of runoff holds paramount significance in the realm of hydrological systems and water resource management, given its substantial implications for water availability, ecological equilibrium, and human activities (Beven, 2011). The accurate forecasting of runoff, particularly in the medium and long term, plays a pivotal role in the comprehensive development, effective utilization, scientific management, and optimization of water resources. The heightened occurrence of extreme floods in recent years, attributed to climate change, has led to substantial human suffering and considerable economic losses worldwide. Consequently, the accurate prediction of the timing and magnitude of peak flow before a flood event becomes an imperative undertaking.

Rainfall–runoff modeling, boasts a rich history within hydrological sciences, with initial attempts to forecast discharge based on precipitation events dating back 170 years (Beven, 2001; Mulvaney, 1850). Currently, hydrological models fall into two main categories: physical models and empirical models (C. Li et al., 2019).

Physical-based models, alternatively termed process-based models, encompass tools like the Hydrologic Engineering Center - Hydrologic Modeling System (HEC-HMS) and the Soil and Water Assessment Tool (SWAT). These models operate on a comprehensive understanding of the underlying physical processes governing hydrological phenomena, simulating water movement through equations rooted in the principles of physics, such as the conservation of mass and energy. However, the application of physical models faces constraints due to high expertise requirements, data acquisition challenges, and significant computational costs (Kratzert et al., 2018; Young & Liu, 2015). Therefore, physically based models are still rarely used in operational rainfall–runoff forecasting. In contrast, empirical models, categorized as statistical models, forego explicit representation of physical processes. Instead, they derive patterns and relationships from historical data to formulate predictions. Empirical models, including regression and neural networks, establish direct connections among hydrologic variables by extracting pertinent information from input data without delving into the intricacies of physical mechanisms (Liu et al., 2021; Young et al., 2017) These models are characterized by their simplicity, lower computational costs, flexibility, and accuracy.

In the 1990s, the initial applications of artificial neural networks (ANNs) for predicting rainfall-runoff, as demonstrated by Daniell (1991) and Halff et al. (1993), showcased their efficacy in capturing nonlinear relationships. Subsequent to these pioneering studies, a multitude of research endeavors have employed ANNs for modeling runoff processes. Despite ANNs exhibiting superior performance compared to process-based models, a notable drawback lies in the significant loss of sequential order information within inputs, attributed to the independence of neurons within the same layer. To address this limitation, recurrent neural networks (RNNs) have emerged as a specialized neural network architecture explicitly designed to comprehend temporal dynamics by processing input data in its sequential order (Rumelhart et al., 1986). This adaptation has proven beneficial in the domain of rainfall-runoff modeling, with RNNs demonstrating accuracy in both single- and multiple-step-ahead forecasting scenarios. (Nagesh Kumar, Raju, et al., 2004; R. Zhao et al., 2019). Hochreiter S. and J. Schmidhuber (1997) proposed the Long Short-Term Memory (LSTM) network to overcome the vanishing gradient problem of the traditional RNN in storing long term dependencies. Today, the LSTM network has great potential and plays an important role in hydrological processes and modeling applications.

In ungauged catchments or basins, remote sensing technology offers continuous observations of variables such as precipitation and temperature over extensive spatial extents and extended time periods. This capability is vital for simulating runoff in areas where traditional measurements are unavailable (Huang et al., 2020). Remote sensing data have been widely applied in combination with hydrological models for runoff estimation, flood estimation and water resource evaluation tasks. However, few studies have combined remote sensing data with LSTM–based models in simulating runoffs in a watershed or catchment area.

1.2 Problem Statement

The Pra River Basin (PRB), second only to the Volta basin in Ghana, holds immense importance due to its substantial contributions to crucial socio-economic activities, including agriculture, tourism, and mining. This basin is a primary freshwater source for four out of Ghana's sixteen administrative regions, supporting a population exceeding 4.2 million people (Awotwi et al., 2017a). The upper reaches of the PRB act as the main source of runoff for the entire basin, playing a pivotal role in supplying water for both domestic and commercial purposes, including irrigation activities and household needs in the middle and lower regions.

Given the indispensable role of the PRB in sustaining various sectors of Ghana's economy and the livelihoods of millions, it becomes crucial to comprehend the dynamics or changes in runoff within the basin. Accurate estimation and prediction of streamflow or runoff at the output discharge station become imperative for effective water resource management. Previous research endeavors have aimed to model hydrological dynamics in the basin, utilizing various methodologies and models with specific aims.

Several process-based models, such as the PolFlow model used for long-term annual Pra river discharge estimation (G. Owusu et al., 2017), have been applied to simulate runoff and hydrological dynamics. Awotwi et al. (2017) explored the variability in runoff and quantified the contributions of precipitation and anthropogenic activities on runoff at the Lower Pra River Basin using the Mann–Kendall statistical test. Additionally, Awotwi et al. (2019a) deployed the Soil and Water Assessment Tool (SWAT) to simulate and assess impacts of Land Use and Land Cover (LULC) changes on Water Balance Components (WBC) in the PRB.

However, despite these efforts, the application of LSTM, a powerful deep learning model, in modeling hydrological dynamics, particularly runoff, within the PRB remains notably absent. Additionally, there have been no instances of employing deep learning models to simulate runoff in the PRB, and the integration of remote sensing data products with LSTM models for runoff estimation in this region has not been explored.

This study therefore seeks to bridge the gaps by integrating remote sensing and neural networks to model the hydrological and runoff dynamics in the Pra Basin. The research aspires to explore and compare the performance of this robust approach with traditional models such as SWAT applied in the basin. By leveraging the capabilities of remote sensing data and the predictive power of neural networks specifically the long-term dependencies capability of LSTM, this research aims to enhance the precision and efficiency of runoff estimation in the PRB, contributing valuable insights for sustainable water resource management in the region. With this focus on deep learning and remote sensing integration for accurate predictions, the following research questions can be stated:

1) How does the incorporation of remote sensing data, including NDVI, land surface temperature, and evapotranspiration, impact the precision and accuracy of LSTM-based runoff predictions?

2) How does the overall performance of the LSTM model compare with SWAT model in estimating monthly runoff discharge in the Pra Basin, Ghana?

1.3 Objectives

The goal of this research is to accurately estimate runoff in the Pra Basin in Ghana, with the explicit objective of integrating in situ data (observed precipitation) and remote sensing products; normalized difference vegetation index (NDVI), land surface temperature (LST), evapotranspiration (ET) with LSTM to optimize monthly runoffs estimates routes.
To achieve this goal, the following intermediate objectives are defined:

- Develop and implement the LSTM model for estimating monthly runoff discharge.
- Utilize established process-based models (SWAT) for monthly runoff discharge estimation in the same study area.

- Assess overall performance of the LSTM model against process-based models based on accuracy metrics such as Root Mean Squared Error (RMSE), Nash-Sutcliffe Efficiency (NSE).
- Analyze model sensitivity: determining the relative contributions of NDVI, LST, and ET to the LSTM model's performance.

1.4 Significance and Relevance

The integration of remote sensing technology plays a pivotal role in advancing hydrological science, particularly in the context of runoff estimation. Remote sensing provides a unique vantage point, enabling a comprehensive understanding of the hydrological dynamics of watersheds or basin over long time frames and large spatial coverages (Xue et al., 2022). The incorporation of NDVI, LST, and ET facilitates a nuanced examination of environmental factors influencing runoff in the Pra Basin.

While most well-known applications of deep learning are in the field of computer vision (Farabet et al., 2013; Krizhevsky et al., 2012; Tompson et al., 2014), the utilization of Long Short-Term Memory (LSTM) in this study underscores its remarkable abilities in capturing intricate temporal patterns, a critical aspect in hydrological modeling. LSTM's adaptability to changing conditions and its capacity to discern complex relationships make it a potent tool for precise runoff estimation in the Pra Basin.

Drawing inspiration from successful applications in various jurisdictions and regions, the approach of incorporating NDVI, LST, and ET with LSTM has been proven effective in enhancing hydrological modeling precision and optimizing water resource management strategies (Xue et al., 2022). By adopting and implementing this proven approach in the unique context of the Pra Basin, the study pioneers a transformative narrative. It brings forth immense benefits empirical approach for hydrological modeling and water resource management in the region, offering tailored insights and solutions. This approach, backed by evidence of its efficacy elsewhere, stands poised to contribute significantly to the advancement of hydrological sciences and the sustainable management of water resources in the Pra Basin.

1.5 Thesis Organization

The theoretical background of the study including the interdisciplinary concepts of hydrological sciences, remote sensing and deep learning as well as the objectives and justification of the project are introduced in this first chapter.

A thorough literature review and use case examples further explains the concepts of hydrological modeling and deep learning neural networks (with focus on LSTM) in the case of runoff estimation in the second chapter.

Chapter three details the overall project methodology. Data acquisition (in-situ observations and remote sensing products), data preprocessing and preparation as well as selection of model parameters are discussed. LSTM model development and implementation is also explained in this chapter. Sensitivity analysis is explored to determine the contribution or influence of each of the remote sensing products on the overall performance of the LSTM model.

A demonstration of the results including graphs, charts and tables (performance of LSTM and SWAT) discussed in chapter four is followed by a discussion of their implications in the fifth chapter.

Finally, a summary of the research contribution and possible recommendations for future works are discussed in the concluding chapter.

Chapter 2

Literature Review

In the pursuit of understanding and advancing hydrological modeling, this chapter initiates a comprehensive exploration of the existing body of knowledge. A synthesis of literature is presented, delving into the hydrological processes, the historical evolution of hydrological models, ranging from physical-based approaches to the contemporary era of empirical methodologies. The roles played by artificial neural networks (ANNs), recurrent neural networks (RNNs), and, in particular, Long Short-Term Memory (LSTM) networks in shaping the landscape of hydrological (runoff) modeling are unveiled throughout the chapter. Furthermore, the pivotal integration of remote sensing technology is scrutinized, examining its contributions to observed variables and its applications in hydrological sciences specifically runoff modeling. The threads of past research endeavors, gaps, challenges, and the theoretical foundations that paved the way for the unique integration of LSTM networks and remote sensing in modeling runoff in watershed or river basins are unraveled.

2.1 Hydrological Process

Hydrological processes refer to the various processes involved in the movement and distribution of water within the Earth's hydrological cycle. These processes include precipitation, runoff generation, evaporation, infiltration, and the behavior of water in different hydrological phases (Zhou et al., 2023). Hydrological processes within a catchment perform an important role in the functioning of the ecosystem, by incorporating the complex processes (physical, chemical and biological) that sustain life (Ekundayo, 2020). A catchment is an area of land that is drained by a river and its tributaries. In studying hydrological processes, the most common spatial unit of consideration is the catchment or drainage basin.

The catchment at various scales within a defined area forms a landscape element that integrates all aspects of the hydrologic processes, which can be studied, quantified, and acted upon (Wagener et al., 2004). A catchment may range in size from a matter of hectares to millions of square kilometers, and all catchments are in reality made up of set of nested subcatchments. River basin is the area of land from which water flows towards a river and then in that river to the sea.

2.2 Components of Hydrological Processes

Within the realm of hydrological processes, distinct components govern the intricate journey of water. From the initial stage of precipitation to the final discharge as runoff, these components, including interception, infiltration, evapotranspiration, and groundwater flow, collectively contribute to the hydrological cycle.

2.3 Hydrological Modelling (Runoff Modeling)

Moradkhani Hamid and Sorooshian, (2008) define a model as a simplified depiction of a real-world system. Models are mainly used for predicting system behaviour and understanding various hydrological processes (Devia et al., 2015). Hydrological modelling according to Islam, (2011) entails formulation of mathematical models to illustrate the interaction between the hydrologic processes as well as represent them (evapotranspiration, interception, infiltration, sub-surface flow, surface flow, precipitation, and snowmelt). Hydrological models are constructed to simulate the genuine relationships among geo-climato-hydrological variables, forecasting their future interactions (Majumder Mrinmoyand Barua, 2010).

According to Devia et al. (2015), a runoff model is characterized as a collection of mathematical equations designed to calculate runoff based on a range of parameters describing watershed attributes. Essential inputs for all models include rainfall data and drainage area. Additionally, watershed features such as soil properties, vegetation cover, watershed topography, soil moisture content, and characteristics of the groundwater aquifer are taken into account. Hydrological models are recognized as crucial tools in contemporary water and environmental resource management.

Runoff models are classified based on model input and parameters and the extent of physical principles applied in the model. These models can be classified based on the type of independent variable (time, space, or both) and can be single or multievent, as well as lumped or distributed (Devia et al., 2015; Majumder Mrinmoyand Barua, 2010; Moradkhani Hamidand Sorooshian, 2008; Sitterson et al., 2018)

Devia et al., (2015) and Sitterson et al., (2018) explain that with lumped models the entire river basin is taken as a single unit hence the outputs (runoffs) are generated without considering the spatial processes and spatial variability within the catchment. A distributed model has the capability to generate spatially distributed predictions by subdividing the entire catchment into small units, typically square cells or a triangulated irregular network. This allows for spatial variation in parameters, inputs, and outputs (Khakbaz et al., 2012).

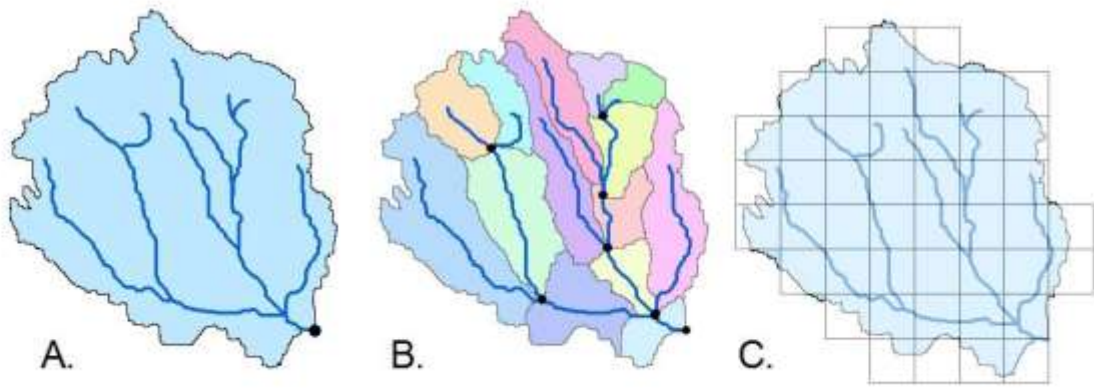


Fig 1. Visualization of the spatial structure in runoff models. A: Lumped model, B: Semi-Distributed model by sub-catchment, C: Distributed model by grid cell. Runoff is calculated for each sub-catchment at the pour point represented by the black dots in Figure 1B. Distributed models calculate runoff for each grid cell, while lumped models calculate one runoff value for the entire catchment at the river outlet point represented by the black dot in Figure 1A. Images from (Sitterson et al., 2018)

Khakbaz et al. (2012) investigate the benefits of employing a semi-distributed modeling framework in hydrologic models. The research introduces an enhanced semi-distributed version of the NWS SACramento Soil Moisture Accounting (SAC-SMA) model, explicitly addressing distinctions between fast and slow response runoff components. Findings reveal that calibration strategies, whether lumped or distributed, enhance simulations at the outlet and yield meaningful streamflow predictions at internal points. Moreover, utilizing a semi-distributed structure calibrated at both interior points and the outlet further enhances model performance at the outlet, even with limited historical data available.

Lavenne et al. (2016) highlight the difficulties faced by the semi-distributed model in identifying stable optimal parameter sets, particularly for the parameters responsible for the closure of the water balance, such as intercatchment groundwater flow parameter. In the study (de Lavenne et al., 2016), emphasis is laid on the importance of considering parameter identifiability in the calibration strategy and evaluation approach of a semi-distributed model, especially when exploring future scenarios in a changing world.

Forecasting or modeling runoff, especially in the medium and long term, plays a crucial role in the comprehensive development, effective utilization, scientific management, and optimization of water resources. Substantial research and application studies have been conducted over the past decades, and these methods can be classified into two categories: empirical models and physically-driven models.

2.3.1 Physically-based models

Physical (process-based) models, also referred to as mechanistic models, are simplified mathematical representations of actual phenomena that incorporate concepts from physical processes (Jehanzaib et al., 2022). These models typically construct a simplified watershed system and express internal behavior by solving mathematical equations that best represent current knowledge of hydrological processes (Rezaeianzadeh et al., 2013a). Finite definite equations are used to represent the hydrological processes of water movement in these process-based models (Devia et al., 2015). Physical models utilize mass, momentum, and energy conservation equations to represent various hydrological processes (Jaiswal et al., 2020). Capable of considering the spatial variability of land use, slope, soil, and climate, physical models, distributed in nature, require extensive topographical, soil, land use, and climatic data. This data-intensive approach provides a framework to explore changes in the hydrological cycle due to human interferences and climate change, crucial for water resources management (Y. Chen et al., 2016; Devia et al., 2015; Yang et al., 2008).

Hydrologic modeling system HEC-HMS (Feldman, 2000) developed by the US Army Corps of Engineers, Hydrologic Engineering Center is one of the well-known physical based models. The Soil and Water Assessment Tool (SWAT) model (Santhi et al., 2001) is a complex physically based model and was designed to test and forecast the water and sediment circulation and agriculture production with chemicals in ungauged basins. Système Hydrologique Europèen (SHE) is also another well-known distributed process-based model (Abbott et al., 1986).

2.3.2 Empirical models

Empirical models, are observational models that solely rely on current data without considering the characteristics and processes of the hydrological system (Pandi et al., 2021). Empirical models establish a direct correlation between hydrologic variables, extracting their relationships from historical measured data through algorithms in the domains of statistics, computational intelligence, machine learning, and data mining (Abrahart et al., 2004; Dawson & Wilby, 2001; Tabari et al., 2010). These models are commonly used in water resources forecasting, especially when relating variables such as rainfall, runoff, and antecedent moisture content (Alaba et al., 2010). They perform good modelling results in ungauged watersheds owing to a lack of particular knowledge about the watershed (Jackson et al., 2011). Empirical models are selected for a variety of reasons, including ease of implementation, quicker calculation speeds, and cost efficiency (Amatya et al., 2022).

Beven (2011) and Granata et al., (2016) elaborate that the majority of empirical models are considered black-box models, indicating limited or no understanding of the internal processes generating runoff results. If the aim is to quantify the flow volume, the target function should center on either the volume of discharge or the shape of the hydrograph (Varvani et al., 2019). In most empirical models, the governing equation for calculating runoff is expressed as a function of inputs, and expressed as

$$Runoff = F\ (X, Y) \quad \text{(Sitterson et al., 2018)} \qquad (1)$$

where X and Y are precipitation and historical runoff.

Madsen (2000) noted that the constraint of empirical models is that their parameters cannot be directly derived from the watershed; consequently, calibration is required.

Artificial neural networks (ANN) and regression analysis are widely used empirical approaches applied in runoff modeling studies.

2.4 Deep Learning (Neural Networks) for Runoff Estimation

Artificial neural networks (ANNs), serving as models for hydrological systems, replicate the functionality of human brain cells by learning patterns from extensive training data (Garbrecht, 2006; Haykin & Network, 2004; Rezaeianzadeh et al., 2013b). The structure of ANN is composed of three layers: (i) input layer, (ii) hidden layer, and (iii) output layer. In engineering applications, the Feedforward Backpropagation (FFBP) is widely utilized among artificial neural networks for making non-linear general estimates (Hornik et al., 1989). ANNs, categorized as empirical models, have gained prominence in rainfall-runoff modeling, finding widespread applications in recent decades (Mukerji et al., 2009; Nourani & Kalantari, 2010; Rezaeianzadeh et al., 2014; Shamseldin, 1997).

Rezaeianzadeh et al. (2014) presents artificial neural networks (ANN), adaptive neuro-fuzzy inference systems (ANFIS), multiple linear regression (MLR), and multiple nonlinear regression (MNLR) for predicting maximum daily flow at the Khosrow Shirin watershed in Iran. The study further evaluates the performance of these models using different input precipitation data sets and concludes that the MNLR model is superior to the other models, with an R^2 value of 0.81 and an RMSE of 0.145 m3 s^{-1}. Furthermore, the authors indicates that the inclusion of antecedent flow with one- and two-day time lags significantly improves flow prediction.

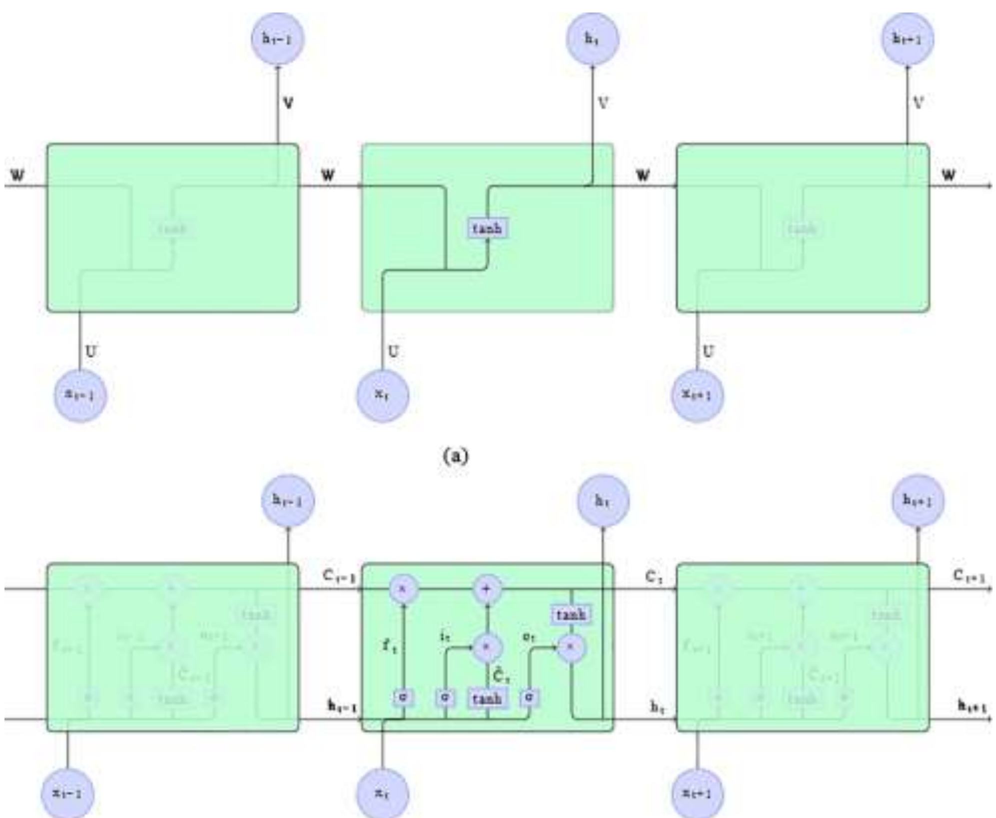
In a study by Gholami & Sahour (2022), ANN is applied to simulate the rainfall-runoff process using data from field sampling plots in conjunction with rainfall and hydrometric data. The modeling process in the study utilizes a multi-layer perceptron network with inputs including rainfall time series, initial loss, soil antecedent moisture condition (A.M.C), and the time to peak of the basin, and the output being the runoff time series. The authors' modeling processes demonstrated positive outcomes, with high R-squared values and low mean square errors, signifying the capability of artificial neural networks (ANNs) for simulating runoff time series and flood hydrographs in basins lacking active hydrometric stations. This highlights the usefulness of ANNs as a valuable tool for hydrological modeling in data-scarce regions (Gholami & Sahour, 2022).

Kişi (2008) explored the application of different ANN techniques, namely feed forward neural networks (FFNN), generalized regression neural networks (GRNN), and radial basis ANN (RBF), for monthly streamflow estimation and forecasting. The performance of these ANN techniques was evaluated for one-month ahead streamflow forecasting using monthly flow data from two stations in the Eastern Black Sea region of Turkey in the first part of the study. It was identified that the GRNN technique outperforms the other ANN techniques in monthly flow forecasting. The study further investigated the impact of periodicity on the forecasting performance of the models. The author used the Levenberg-Marquardt (LM) optimization technique for adjusting the weights in the FFNN simulations, which is considered more powerful than the conventional gradient descent technique. The RBF and FFNN techniques were found to be better than the GRNN technique in monthly river flow estimation using nearby river data during the second part of the study. The potential of artificial neural networks (ANN) as a promising alternative for rainfall-runoff modeling, streamflow prediction, and reservoir inflow forecasting in the hydrological forecasting context was highlighted in the study (Kişi, 2008).

ANNs have notable limitations including challenges associated with over-fitting, negotiating local minima, vanishing gradient point, overseeing learning rate processes, coping with computation time and cost, and demanding manual interventions like training. In recent rainfall–runoff modeling (runoff estimation), advanced forms of artificial neural networks such as deep neural networks (DNN), convolutional neural networks (CNN), long short-term memory (LSTM), and recurrent neural networks (RNN) (Han et al., 2021; Kratzert et al., 2018; Roy et al., 2021; Xue et al., 2022; Yin et al., 2021; Yokoo et al., 2022a) have gained prominence.

2.5 Recurrent Neural Networks for Runoff Estimation

Recurrent Neural Networks introduced by (Rumelhart et al., 1986), represent a distinct variant of neural networks characterized by interconnected neurons forming a directed cycle. This architectural design incorporates internal self-looped cells, enabling dynamic temporal behavior. RNNs adopt chain-like structures comprising repeating modules, facilitating the retention of past information. Conventional RNNs take into consideration the sequential details of time series, allowing them to retain and comprehend past information, thereby possessing the capability to effectively handle varied and prolonged sequences. This inherent ability to remember prior data contributes to the success of RNNs in tasks related to learning sequences and temporal dynamics (Bhattacharjee & Tollner, 2016; Pan et al., 2007).



(a)

(b)

Fig. 2. (a) Architectural representation of the recurrent neural network. The self-connected hidden units allow information to be passed from one step to the next. (b) Graphical representation of LSTM's memory block Image sourced from (J. Zhang et al., 2018).

Carriere et al. (1996) conducted one of the first studies utilizing RNN for runoff modeling. The study presented a Virtual Runoff Hydrograph System (VROHS) trained using data from 45 lab experiment sets, with 29 sets used for training and 16 sets for testing. The authors further highlight the use of a recurrent back-propagation neural network in the VROHS, which contributes to its ability to generate accurate predictions. (Hsu et al., 1997) also conducted a comparative analysis between a traditional ANN and an RNN. Although the conventional ANN demonstrated similar performance to the RNN, the study underscored the importance of the number of delayed inputs, a critical hyperparameter for the ANN's efficacy. However, the architecture of the RNN rendered the search for this number obsolete.

RNNs were found to perform better than feed forward networks in forecasting monthly river flows, as indicated in a comparative study by (Nagesh Kumar, Srinivasa Raju, et al., 2004). The selected RNN models showed improved accuracy in both single step ahead and multiple steps ahead forecasting of river flows. The architecture size and training time required for RNNs were found to be less compared to feed forward networks. The authors recommended the use of recurrent neural networks as a tool for river flow forecasting, based on their superior performance and efficiency.

2.6 Long Short-Term Memory for Runoff Modeling

The typical configuration of a standard RNN involves simple looping units, commonly utilizing a single layer with hyperbolic tangent (tanh) activation, as observed in various studies (Gelly & Gauvain, 2018; Mou et al., 2017; Poornima & Pushpalatha, 2019; Shewalkar et al., 2019). To address the vanishing gradient issue and facilitate the learning of longer-term dependencies in sequential data, merits of nonlinear predictive capability and faster convergence, the LSTM architecture (improved variant of RNN) was introduced by Hochreiter & Schmidhuber, (1997).

The architecture of the LSTM network encompasses an input layer, memory cell(s), and an output layer. Key to its functionality are the three gates within each memory cell: a forget gate (f_t), an input gate (i_t), and an output gate (o_t). At each time-step t, these gates are influenced by both the input x_t and the output h_{t-1} from the memory cells of the previous time-step $t-1$ (Hochreiter & Schmidhuber, 1997; Hu et al., 2018).

f_t: specifies the content removed from the cell state.

i_t: specifies the content added to the cell state.

o_t: specifies the content from the cell state is used.

The sequential update formulas are

Input node

$$g(t)=tanh(W_{gx}x(t)+W_{gh}h(t-1)+b_g) \quad (2)$$

Input gate

$$i(t)=\sigma(W_{ix}x(t)+W_{ih}h(t-1)+b_i) \quad (3)$$

Forget gate

$$f(t)=\sigma(W_{fx}x(t)+W_{fh}h(t-1)+b_f) \quad (4)$$

Output gate

$$o(t)=\sigma(W_{ox}x(t)+W_{oh}h(t-1)+b_o) \quad (5)$$

Cell state

$$s(t)=g(t) \,.\, i(t)+s(t-1) \,.\, o(t) \quad (6)$$

Hidden gate

$$h(t)=tanh(s(t) \,.\, o(t) \quad (7)$$

Output layer

$$y(t)=(W_{hy}h(t)+b_y) \quad (8)$$

where σ is the activation function (sigmoid),

x_t is the input vector for time step t,

W_s is the weight, b_s is the bias,

y_t is the output at time step t,

h is the hidden state,

s is the memory cell state

Despite being widely employed in computer vision (Farabet et al., 2013; Krizhevsky et al., 2012) domains, such as speech recognition (bukhari et al., 2017; Greff et al., 2016; Hinton et al., 2012) and natural language processing (NLP) (Sutskever Google et al., 2014.), LSTM has also shown promise in hydrological modeling and by extension rainfall-runoff modeling, as indicated by previous studies (Boulmaiz et al., 2020; W. Li et al., 2021; Sahoo et al., 2019; Yokoo et al., 2022b; Yuan et al., 2018).

W. Li et al. (2021) developed an empirical rainfall runoff (RR) model using a sequence-to-sequence LSTM network for high temporal resolution modeling of RR. The LSTM network was capable of learning long-term dependencies between input and output, allowing for modeling the RR at a 15-minute time resolution. The model was tested using 10 years of precipitation data from 153 rainfall gages and river channel discharge data, with over 5.3 million data points. In the study, the LSTM model showed efficient prediction of river discharge and achieved good model performance when compared to a process-driven model, Gridded Surface Subsurface Hydrologic Analysis (GSSHA). The overall performance of the developed empirical rainfall runoff (RR) model using the LSTM network was evaluated using the test Nash-Sutcliffe model efficiency (from 0.666 to 0.942).

Yuan et al. (2018) employed a hybrid model called LSTM–ALO, which combines LSTM network and the ant lion optimizer, for monthly runoff forecasting. The LSTM parameters were calibrated using the ant lion optimizer to improve the prediction performance of the LSTM model. The study utilized historical monthly runoff data from the Astor River Basin in Pakistan for the case study. The simulation results from the Astor River Basin demonstrated that the LSTM–ALO model has higher accuracy compared to three other models; Back Propagation (BP), Radial Basis Neural Network (RBNN) and least squares support vector machine (LSSVM) for monthly runoff forecasting.

Hu et al. (2018) also compared the performance of ANN and LSTM models in terms of rainfall-runoff simulation and prediction in the Fen River basin in North China. The study used data from 14 rainfall stations and one hydrologic station in the catchment, covering a period from 1971 to 2013. Both models were found to be suitable for rainfall-runoff modeling, but LSTM models outperform ANN models in terms of simulation performance and stability. The performance of the models was assessed using statistical error measures such as the coefficient of determination (R2), root mean square error (RMSE), Nash-Sutcliffe Efficiency (NSE), mean absolute error (MAE), error of time to peak discharge and error of peak discharge. The LSTM model achieved high values of R^2 and Nash-Sutcliffe Efficiency (NSE), indicating its superior performance in simulating the rainfall-runoff process.

In their recent work, Man et al. (2023) developed an enhanced LSTM model for daily runoff prediction in China's upper Huai River Basin, focusing on accurately forecasting peak runoff crucial for flood defense. The model introduces unique loss functions, including peak error tanh (PET) and peak error swish (PES), emphasizing peak runoff predictions. Utilizing a feature extractor with three LSTM networks, the model captures temporal features, enhancing its capacity to handle the complexity of runoff processes. In validation, the enhanced LSTM model outperforms widely used process-based hydrological (Australian Water Balance Model, Sacramento, SimHyd, and Tank Model) and empirical models (artificial neural network, support vector regression, and gated recurrent units), with the PES loss function exhibiting superior performance in predicting extreme runoff, achieving a mean NSE for floods of 0.873 (Man et al., 2023).

Boulmaiz et al. (2020) also compared the performance of LSTM and traditional FFNN models in rainfall-runoff modeling, utilizing varying training data sizes from 3 to 15 years. Results indicated that the deep LSTM consistently outperformed the FFNN across different training set sizes. Notably, the LSTM demonstrated satisfactory performance with a 9-year training data length, further improving with 12 years of data. The study highlighted the deep LSTM's competence even with a small 3-year data size, surpassing the FFNN, which required 9 years for comparable outcomes.

J. Li et al. (2022) employed LSTM to forecast inland runoff in arid regions, utilizing a model with 5 layers, Relu activation, Mean Absolute Error (MAE) as the loss function, and Adam as the optimizer (learning rate 0.01). Trained and tested on Yarkant River runoff data in Northwest China, the LSTM model showed better prediction accuracy compared to models such as convolutional neural network (CNN), Decision Tree Regressor (DTR), and Random Forest (RF), with MAE and RMSE values of 3.633 and 7.337.

2.7 Remote Sensing applications (products) in Runoff modeling

Remote sensing products have become invaluable tools in the field of hydrological modeling, providing essential data for understanding and predicting runoff dynamics (Jiang & Wang, 2019; Moradkhani, 2008; Thakur et al., 2017). Many studies have highlighted the use of satellite-based remote sensing as an alternative means to observe hydrological variables such as remotely sensed precipitation, land surface temperature, evapotranspiration, soil moisture, and snow water equivalents for effective hydrological modeling and by extension runoff predictions (X. Chen et al., 2017; Hartanto et al., 2017; Kalma et al., 2008; Laiolo et al., 2016; Y. Li et al., 2018).

Sirisena et al. (2020) adopted two calibration approaches: single variable calibration with measured station streamflow and remotely sensed Global Land Evaporation Model (GLEAM ET) evapotranspiration separately, and multi-variable calibration with both variables together to calibrate SWAT models in the Chindwin Basin, Myanmar. The performances of each calibrated model were evaluated using the NSE metric. The simulation results showed that the GLEAM ET data together with the streamflow data can be used for model calibration as the performance was reasonable with an NSE > 0.85.

Owusu et al. (2019) evaluated the accuracy of three satellite rainfall products (TMPA 3B42RT, TMPA 3B42, and CMORPH) in the Pra basin of Ghana using point-to-pixel method to compare the satellite rainfall data with gauged rainfall measurements. Performance evaluation methods such as correlation coefficient (r), bias, and percent bias (%Bias) were used to analyze the data at daily, monthly, annual, and seasonal timescales. The TMPA 3B42 and 3B4RT products performed better than the CMORPH concluding that remotely sensed rainfall data can be used to supplement gauged rainfall measurements in the basin and estimate rainfall in ungauged basins with similar characteristics.

Furthermore, Xue et al. (2022) investigated the use of multiple remote sensing parameters (NDVI, Precipitation, LST and ET provided by MOD13A2, TRMM3B43, MOD11A2 and MOD16A2 respectively) along with hydrological station data (precipitation and runoff or discharge) in estimating runoff Upper Reaches of the Heihe River Basin in China using an LSTM model. The study period was covered from 2001 to 2016, which may not have captured long-term variations or trends in the runoff. However, the study showed that the incorporation of remote sensing data improved the quality of runoff estimation with limited data availability compared to rain gauge observations, with an increase in R^2 from 0.91 to 0.94 and an improvement in NSE from 0.89 to 0.93.

Investigating the impact of land cover changes on watershed hydrological response, Nourani et al. (2017) utilized a conceptual rainfall-runoff model. The researchers harnessed Landsat-derived NDVI to establish the relationship between the land cover characteristics of the watershed and the model parameter. NDVI values played a crucial role in considering the effects of land cover changes on the response of rainfall-runoff.

In their research, Park et al. (2014) assessed the effectiveness of MODIS NDVI and LST in explaining forest soil moisture using the SWAT model. Calibration and verification of the SWAT model were conducted with daily streamflow data from three gauging stations and daily measured soil moisture data from three locations within the watershed. The study showed a higher correlation between SWAT soil moisture and MODIS LST during the forest leaf growing period (March-June). The correlation between SWAT soil moisture and MODIS NDVI was also higher during the leaf falling period (September-December), emphasizing the practical application of MODIS NDVI and LST as valuable indicators of forest soil moisture in hydrological modeling.

While not focused on rainfall-runoff, Thapa et al. (2020) utilized various machine learning techniques, including LSTM, nonlinear autoregressive exogenous model (NARX), Gaussian process regression (GPR), and support vector regression (SVR), along with satellite data such as 3B42RT TRMM precipitation and MOD10A2 snow cover datasets. The study, conducted in the Himalayan basin, Nepal, aimed to model snow discharge. Through a Gamma test, the authors determined the optimal input combination for the models. The results, indicating higher R^2, NSE, RMSE, and modified Kling–Gupta efficiency (KGE') values, underscored the effectiveness of LSTM and the use of remote sensing datasets (MODIS and TRMM) in snowmelt and, by extension, hydrological modeling.

In summary, this chapter delved into pivotal aspects of hydrological science and runoff modeling, covering various hydrological processes, diverse model types and approaches. The incorporation of remote sensing alongside neural networks for runoff modeling was extensively explored, drawing insights from various research findings. These insights not only shed light on the capabilities of different methodologies but also serve as inspiration and motivation for the chosen modeling approach in this study. The synthesis of these elements not only deepens our understanding of hydrological processes but also underscores the transformative potential of advanced technologies in achieving more robust rainfall runoff predictions.

Chapter 3

Methodology

This chapter transitions from the theoretical underpinnings discussed in the literature review to the practical implementation of the robust methodology for rainfall runoff modeling in the study area. The chosen methodology (LSTM) is a carefully crafted framework designed to address the mechanical and computational complexities associated with utilizing the process-based approaches to model hydrological processes specifically runoff modeling, incorporating insights learned from the literature review. Through a systematic approach, this chapter outlines the data collection methods, model development procedures, and the integration of advanced technologies, such as remote sensing and neural networks, in the pursuit of accurate and efficient rainfall runoff predictions.

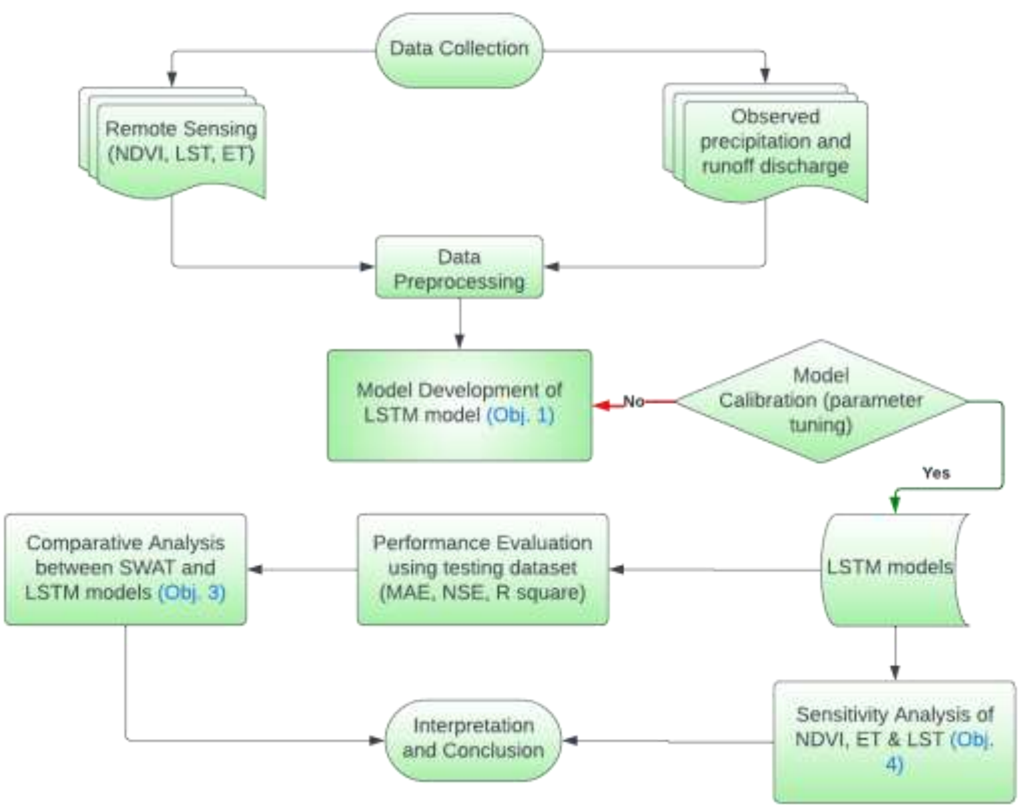


Fig 3. Methodology flow chart

3.1 Study Area

The Pra Basin, situated in Ghana's forest ecological zone, spans latitudes 5°N to 7° 30’N and longitudes 2° 30’W to 0° 30’W in south-central Ghana, with an approximate surface area of 23,330 km^2. Comprising three interconnected sub-basins; Offin, Birim and Pra, it ultimately discharges into the Gulf of Guinea, part of the Atlantic Ocean in the northeast (Awotwi et al., 2019b). The Pra River and its branches form the largest river basin within the southwestern region of Ghana (Commission, 2012)

The basin displays diverse topography with flat terrain in the south and scattered mountains in the central to northern areas, reaching peak elevations of 800 meters along the northern and eastern borders. Predominantly covered by moist semi-deciduous forest, the basin exhibits varied land cover, including agricultural land (60%), forest (30%), and a mix of grassland and human settlements (10%) (C. Owusu et al., 2019b). Significant features include isolated forest reserves and extensive commercial tree plantations located within the basin.

Two main seasons occur in the basin: rainy/wet season (major and minor occurring in April-July and September-November respectively) and the dry season (December - March). The annual rainfall ranges from 1300 mm to 1900 mm with a mean of 1500 mm and the mean surface temperatures range between 23 °C and 33 °C (Akrasi & Ansa-Asare, 2008; Awotwi et al., 2017b, 2019b; C. Owusu et al., 2019b).

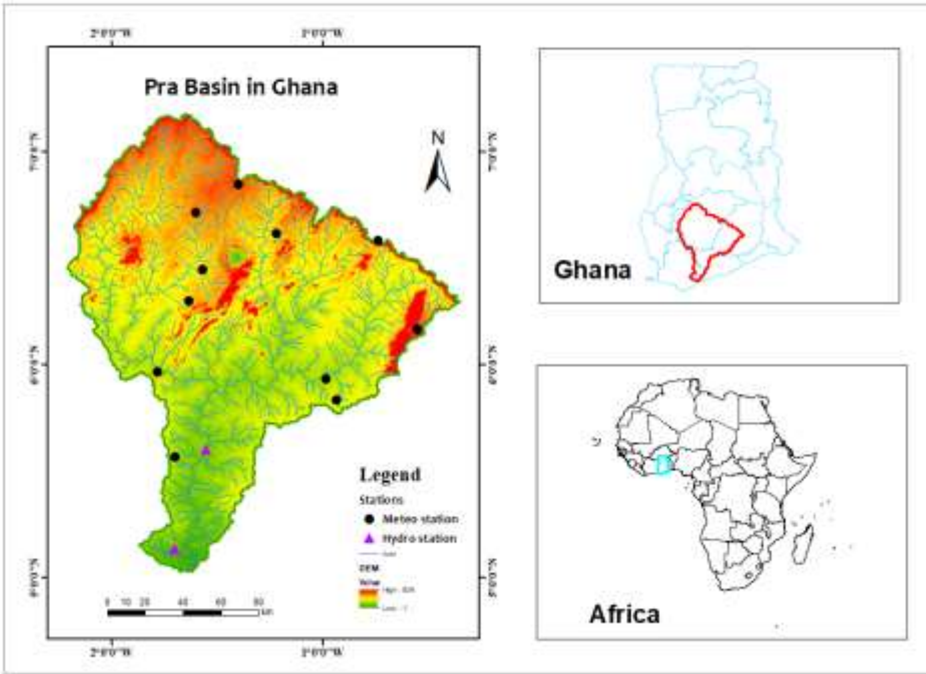


Fig 4. Spatial distribution of gauge station in the Pra basin and boundaries of Ghana and Africa

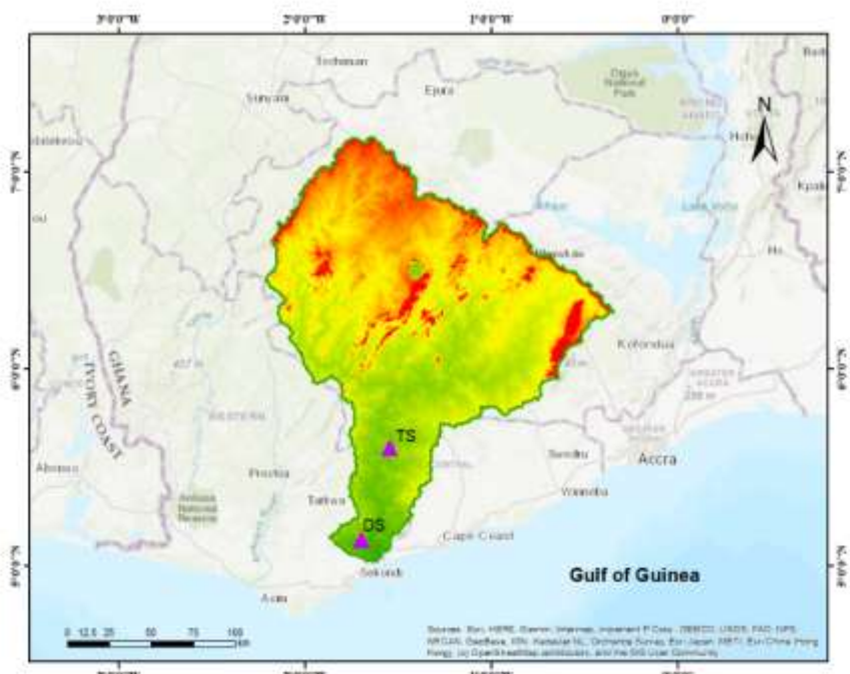


Fig 5. Location map of the study area

3.2 Data Acquisition

3.2.1 Field Observations

A network of 11 meteorological stations, spatially distributed provided the daily precipitation data, offering a localized perspective and understanding the temporal patterns of precipitation within the study period (2001-2020). Two hydrological stations; Twifo Praso station (TS) and Daboase station (DS) as shown in fig 4 also supplied daily runoff discharge data for the specified timeframe (2001-2020). This data was crucial for understanding the hydrological response or flow of water within the basin. All field data were obtained from the Ghana Meteorological Agency (GMA).

3.2.2 Satellite Data

NDVI

The NDVI data, essential for evaluating vegetation dynamics which is a pivotal factor influencing hydrological processes was acquired from the MODIS Terra Daily NDVI. This dataset, accessible through the Google Earth Engine platform, is derived from the near-infrared and red bands of the MODIS/006/MOD09GA surface reflectance composites, at a spatial resolution of 500m. The MOD09GA derived NDVI are extremely useful for monitoring vegetation dynamics as it has been shown in past studies including monitoring vegetation growing season, crop phenology, estimating phenological metrics and large-scale rice yield estimation, etc. (Rykin et al., 2019; Son et al., 2014; Testa et al., 2018; H. Zhao et al., 2009).

Land Surface Temperature

The Land Surface Temperature (LST) data was provided by the MOD11A1 V6.1 product. This product provides daily LST information in a 1200 x 1200 km grid, featuring distinct bands for day and night temperatures. To synthesize the data, an averaging process was applied to the day and night bands, resulting in a consolidated band that represents the mean LST for each time step throughout the study period. The LST data was resampled to 500m spatial resolution.

Evapotranspiration

The ET data were obtained from the 8-day MOD16AGF.016 product at a spatial resolution of 500m. This data is essential in understanding the water loss through evaporation of soil water and transpiration of water from plant leaves in the basin. Kim et al. (2012) affirms the reliability of the MODIS derived ET data in estimating actual ET values.

Precipitation (PT)

Remotely sensed precipitation was also obtained for this study to be used in together with the in-situ rainfall recorded for comparative analysis and modeling the runoff in the basin. The data was acquired from the global rainfall dataset, Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) with a spatial resolution of 5km.

All data used in the study is shown in table 1.

Table 1. Datasets used for building LSTM for runoff prediction

Hydrological Data	Data Source	Temporal Resolution	Spatial Resolution
Precipitation	GMA	Daily	
	CHIRPS	Daily	5km
Land surface temperature	MOD11A1 V6.1	Daily	1 km
NDVI	MOD09GA_006	Daily	500 m
Evapotranspiration	MOD16A2GF.061	8-Day	500m
Observed Discharge	GMA	Daily	

3.3 Data Preparation and Preprocessing

All the remotely sensed data were obtained and processed using the Google Earth Engine (GEE) platform, known for its efficient batch processing of large images through the online JavaScript application programming interface (API). The study area's vector data (boundary shapefile) was imported into GEE to extract all relevant remotely sensed data within the specified study period. The CHIRPS and LST images were resampled from 5km and 1km respectively to 500m spatial resolution. The 8-day MOD16A2 – ET data was synthesized to daily temporal resolution, and the reduceRegions function in GEE was utilized to compute spatial mean values over the study area for each image. Subsequently, the mean values of PT, NDVI, LST, and ET were exported as CSV files for further use in developing the LSTM model for runoff predictions. The rainfall measurements recorded by the 13 meteorological stations distributed in the basin were aggregated to a single value representing the mean rainfall in the basin for each daily time step within the study period (2001-2020).

The runoff discharge recorded at the outlet station (DS) or downstream (of the two hydrological stations (TS and DS) as shown in fig 4) was used to represent the runoff of the entire basin.

3.4 Feature Selection

The process of selecting input variables is a critical aspect of runoff estimation, exerting a substantial influence on the accuracy of estimation outcomes. Numerous variables contribute to the precipitation–runoff relationship, necessitating a thoughtful grouping of inputs for the LSTM model. Drawing on literature and existing knowledge, certain variables were designated as standalone inputs, while others were combined to assess their impact on estimation results. Input combinations were determined based on observed correlations among variables.

To explore the relationship between precipitation and NDVI, both variables were grouped together with the runoff variable as a unified input. Additionally, evapotranspiration was incorporated with NDVI as inputs, acknowledging the influence of vegetation cover on evaporation and transpiration. Standalone inputs included both observed and remotely sensed precipitation, given the pivotal role of precipitation as the primary input for runoff. The following input combinations were examined:

C1: P T(observed)

C2: PT(RS)

C3: PT (observed), NDVI

C4: PT (observed), NDVI, ET

C5: PT (observed), NDVI, ET, LST

C6: PT (RS), NDVI, ET, LST.

3.5 LSTM Model Development

3.5.1 Data Splitting

In order to effectively train, validate, and test the LSTM model within the study's timeframe (2001-2020), a systematic data splitting approach was employed. Following the widely adopted 70-30 principle, the dataset was partitioned into three distinct subsets: training, validation, and testing.

Training Data (70%): The training dataset encompassed the period from 2001 to 2014, constituting 70% of the overall data. This extensive timeframe provided the foundation for training the LSTM model, allowing it to learn the intricate patterns and relationships within the hydrological processes.

Validation Data (15%): To assess the model's performance and fine-tune its parameters, a dedicated validation dataset was carved out from the years 2015 to 2017, representing 15% of the total data. Both the training and validation datasets collaboratively contributed to the model development process.

Testing Data (15%): The remaining 15% of the data, spanning from 2018 to 2020, served as the independent testing dataset. This dataset played a crucial role in evaluating the model's predictive capabilities and generalization to unseen data, providing insights into its real-world applicability.

3.5.2 Data Transformation

To facilitate efficient training and enhance the model's learning process, a crucial step involved the transformation of all datasets (PT, NDVI, LST and ET as inputs and runoff as output variables) comprising training, validation, and testing sets. The transformation process was executed utilizing the MinMaxScaler function from the known scikit-learn library. The MinMaxScaler function fits and transformed data based on the computed minimum and maximum values. This MinMax Scaling process ensures that all datasets share a consistent scale between 0 and 1, preventing features with larger magnitudes from dominating the model training. The scaled datasets collectively contribute to an optimized and standardized learning experience for the LSTM model. MinMaxScaler is expressed as:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where:

- X_{scaled} is the scaled value.
- X is the original value of the variable.
- X_{min} is the minimum value of the variable in the dataset.
- X_{max} is the maximum value of the variable in the dataset.

3.5.3 Fully Connected LSTM model

Following the transformation of datasets through MinMax scaling, the subsequent step involved the training of the LSTM model. Ensuring proper formatting of the training, validation, and testing datasets was crucial to align with the requisite input shape for LSTM: $N \times T \times D$, where N signifies the number of observations, T denotes the time step (with a chosen window size of 7 for capturing short to mid-term dependencies), and *D* represents the number of features/variables.

The architectural composition of each LSTM model included an input layer utilizing the input combinations, an LSTM layer, a fully connected dense layer, and a single output layer predicting the target variable (runoff).

The LSTM model construction, executed using the TensorFlow-Keras framework, involved the configuration of various hyperparameters. Hyperparameter tuning was conducted using the random search approach. The learning rate, acting as a critical factor influencing the size of steps during optimization, was set to govern the convergence speed of the model. Adam optimizer was employed to adjust model weights based on computed gradients, ensuring optimal performance. Determining the number of hidden LSTM units and dense (fully connected layer) units was crucial for capturing intricate temporal patterns and enabling effective learning.

In addition, dropout (0.5), a regularization technique, was introduced to mitigate overfitting by randomly deactivating a fraction of neurons during training. The batch size (32), defining the number of samples processed before updating the model, impacted the computational efficiency and convergence speed. The activation function used was Rectified Linear Unit (ReLU), introducing non-linearity to the model and influencing the information flow.

During training, the validation dataset coupled with MSE as the validation metric guided the model's training by ensuring it not only learns from the training data but also generalizes well to diverse unseen data. Monitoring the validation metric, such as MSE, helps in fine-tuning the model's hyperparameters and ensuring it achieves good predictive performance on new data.

Multiple FC-LSTM models were constructed for each input combination, and the hyperparameter tuning process aimed to select the most optimal parameters for each model was done. Table 2 shows the hyper parameters for each model.

Table 2. Hyperparameters for LSTM models

Model	Learning rate	Dropout	No. of LSTM units	No. dense units
C1	0.0009829175534613746	True	80	80
C2	0.0008792587648188515	True	90	60
C3	0.000968299956121367	True	100	40
C4	0.0009231024572743306	True	60	20
C5	0.0008481676645756039	True	100	40
C6	0.0008792463061777378	True	70	80

3.6 Evaluation of Model Performance

The performance of the models was assessed or evaluated to determine the accuracy and reliability of the LSTM model in predicting runoff. Multiple metrics were employed for a comprehensive evaluation, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R-squared (R^2), and Nash-Sutcliffe Efficiency (NSE).

MSE is used to assess the average squared difference between the predicted and observed values quantifying the overall magnitude of errors. A lower MSE indicates a better fit of the model to the data. MSE is defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \qquad (2)$$

where:

- n is the number of observations,
- Y_i is the observed value,
- $\hat{Y}_i$ is the predicted value.

RMSE used to evaluate the accuracy of the model by measuring the average magnitude of the errors between predicted and observed values. A lower RMSE indicates that, on average, the model's predictions are closer to the observed values. Conversely, a higher RMSE suggests larger prediction errors. Root Mean Square Error is then calculated as:

$$RMSE = \sqrt{MSE} \tag{3}$$

The R^2 metric assesses how well the model captures the actual values, with values ranging from 0 to 1. A higher R^2 value, nearing 1, indicates a more effective model fit. R^2 is further calculated as:

$$R^2 = \frac{\sum_{i=1}^{n}[(Y_i - \bar{Y}_i)\,(Y_i - \hat{Y}_i)]^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2\ \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y}_i)^2} \tag{4}$$

where:

- n is the number of observations,
- Y_i is the observed value,
- $\hat{Y}_i$ is the predicted value.
- $\bar{Y}_i$ is the mean of the observed values (Y_i)

The NSE metric employed to assess the accuracy of hydrological model predictions, with NSE values ranging from negative infinity to 1. A higher NSE value indicates a better fitting effect of the model. NSE is expressed as:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \tag{5}$$

- where $\hat{Y}_i$ is the simulated/predicted runoff, Y_i represents the observed runoff and $\bar{Y}_i$ is the mean of the observed values (Y_i).

Chapter 4

Results and Discussion

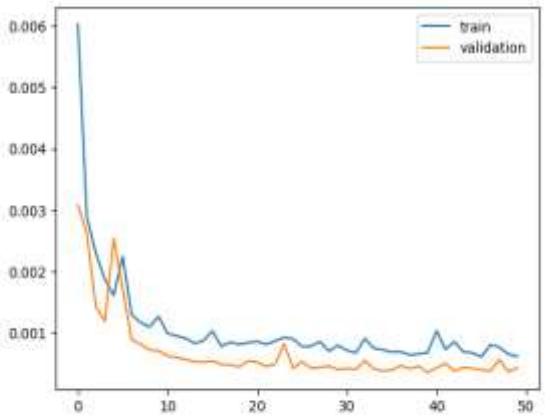
In this chapter, the results obtained from the implementation of the proposed methodology outlined in Chapter 3 is implemented. Through rigorous analysis and evaluation, this section delves into the outcomes of utilizing advanced technologies, including remote sensing datasets and LSTM modeling, to predict rainfall runoff in the study area. The findings offer valuable insights into the effectiveness and reliability of the developed models, shedding light on the intricate dynamics of hydrological processes within the river basin.

4.1 Model Training Results

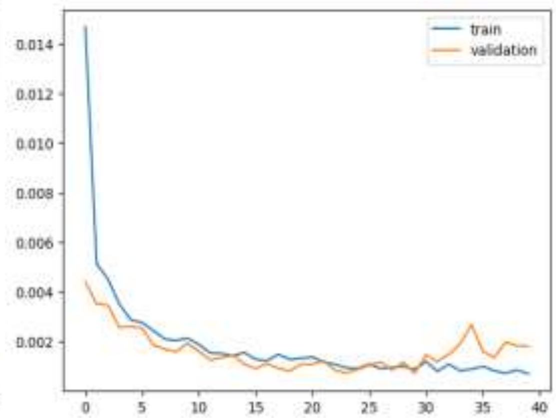
After training six distinct LSTM models with varied input combinations and predefined hyperparameters as detailed in Chapter 3, an iterative fitting process ensued over 200 epochs per model. The objective was to achieve convergence between validation and training loss functions (figure 6), aimed to refine models for optimal performance. The C5-model, incorporating observed precipitation and remote sensing data (NDVI, LST, and ET), exhibited the best convergence, with training and validation losses of 0.000595 and 0.000531, respectively. Complete details on minimal loss values are presented in Table 3.

Table 3. Minimal loss values of the six distinct LSTM models

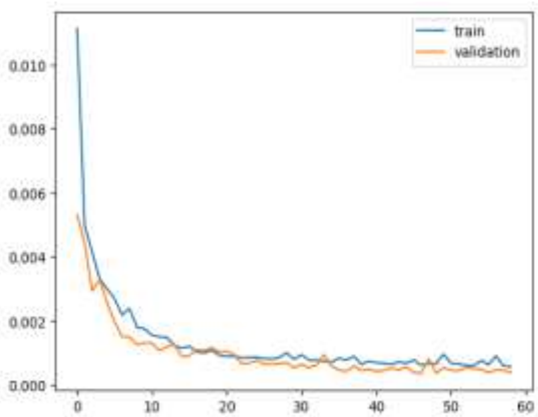
LSTM Model	Training Loss	Validation Loss
C1	0.000622	0.000432
C2	0.000653	0.000374
C3	0.000659	0.001719
C4	0.000595	0.000418
C5	0.000595	0.000531
C6	0.000704	0.000521



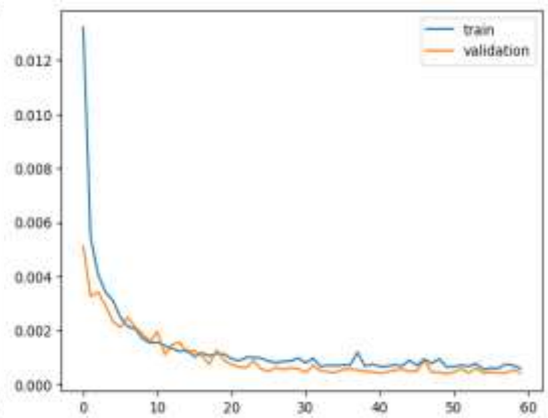
C1



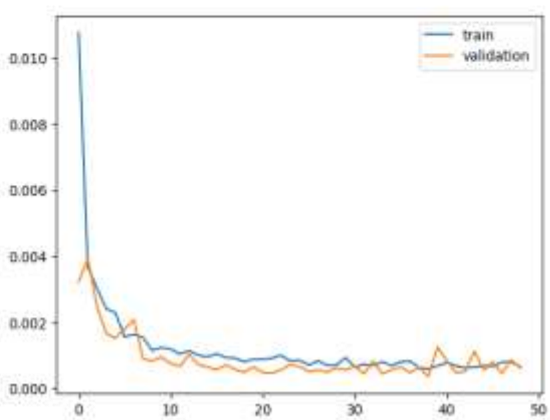
C2



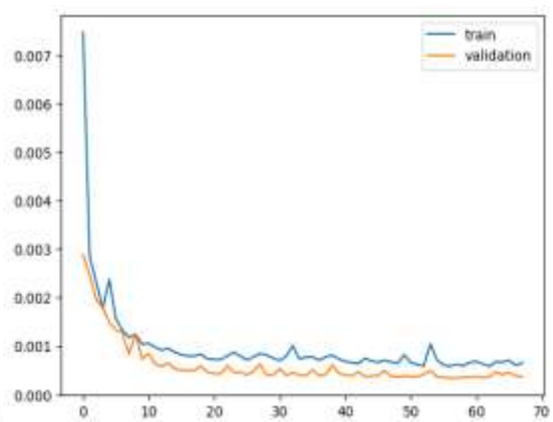
C3



C4



C5



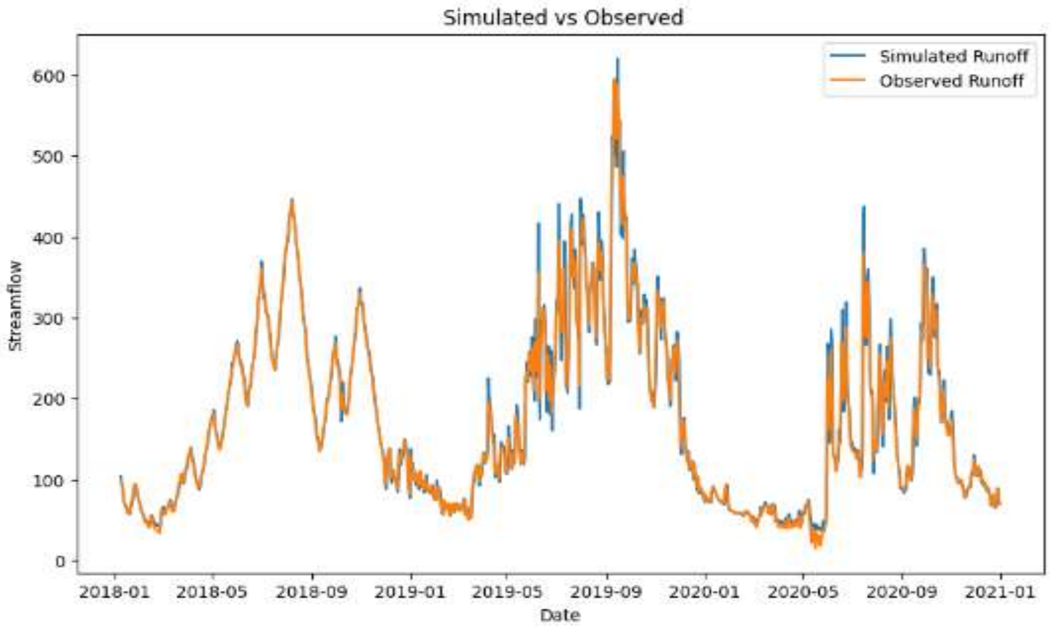
C6

Fig 6. Training and validation loss curves of LSTM models

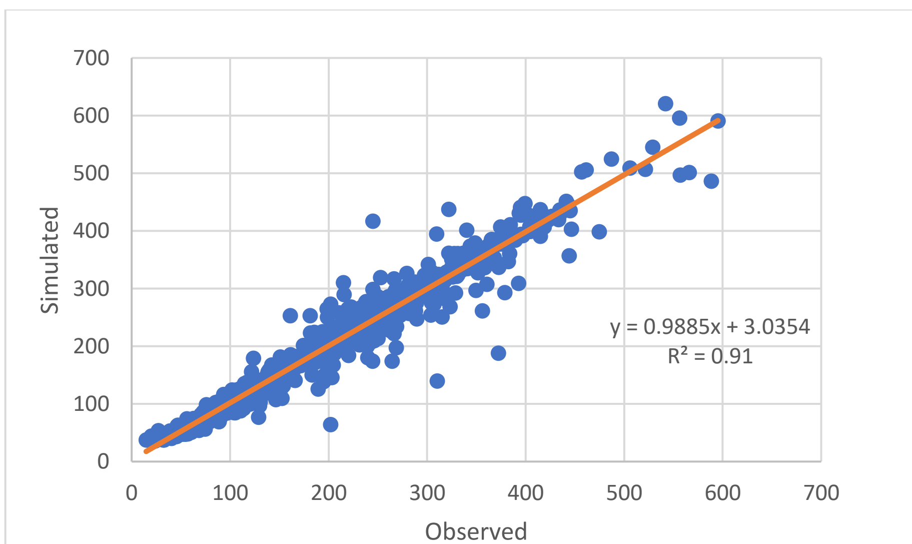
4.2 LSTM Runoff Prediction or Simulation

The LSTM models developed in chapter 3 were employed to predict the runoff across the entire basin for the period spanning 2018 to 2022, utilizing the testing dataset. Subsequently, the accuracy of these predictions or simulations was assessed using the testing data. Within the testing period (i.e. 2018-2020), the peak observed runoff discharge recorded at the Daboase hydro-station was on September, 2019 equal to 595.21mm/s with the lowest runoff observed on May, 2020 at 14.673mm/s. The peak observed runoff values for the testing period occurred roughly around the rainy or wet season within the basin.

The predictions made using the C1-LSTM model, based solely on observed precipitation data from rainfall gauge stations as input, resulted in simulated values that were generally higher than the observed flows. The hydrograph in Figure 7a illustrates the observed flow compared to the simulated flow for the 2018-2020 period within the basin. While the C1-LSTM model was able to capture the overall trend or pattern of the observed runoff, there were differences in the peak values. The simulated runoff value on the observed peak date was 599.61mm/s, closer to the observed peak runoff value. Evaluation of prediction accuracy using R^2, RMSE, and NSE revealed high R^2 and NSE values of 0.91 and 0.89, respectively, and a low RMSE value of 26.006.



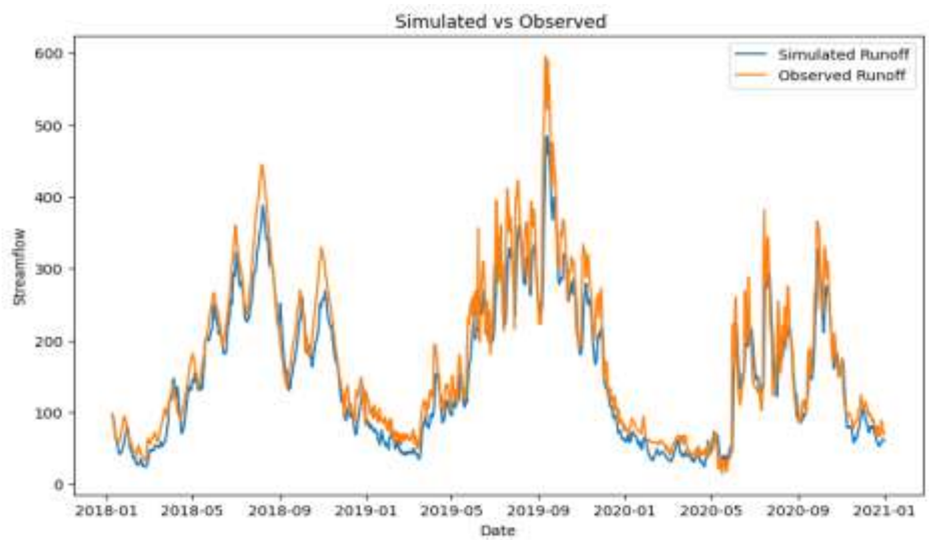
a)



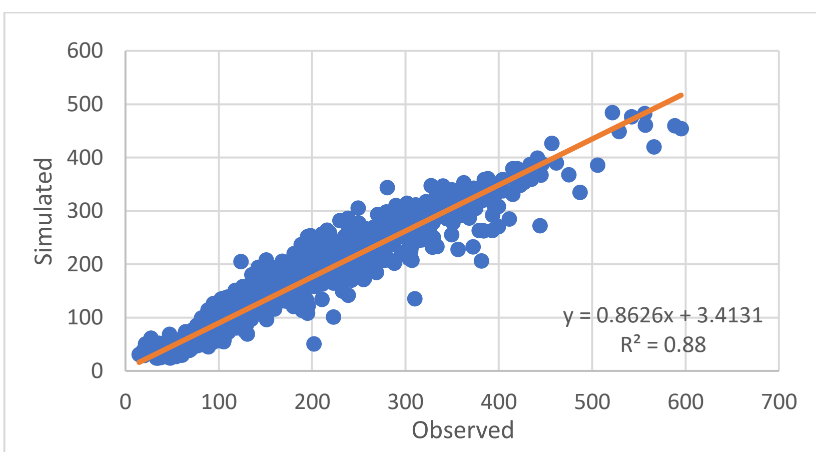
b)

Fig 7. a) Hydrograph from C1-LSTM model prediction b) Regression trendline of model simulations and observed.

Figure 8a depicts the hydrograph generated by the C2-LSTM model, which utilized remotely sensed precipitation data, alongside the observed runoff. While the simulations generally mirror the observed flow pattern, notable underestimations in the simulated runoff values are evident compared to the observed data. Additionally, the C2-LSTM model's predictions exhibit a decrease compared to those of the C1-LSTM model. Specifically, the R2 and NSE values decrease to 0.88 and 0.79, respectively, compared to the performance of the C1-LSTM model and an increase in the RMSE value to 27.109.



a)

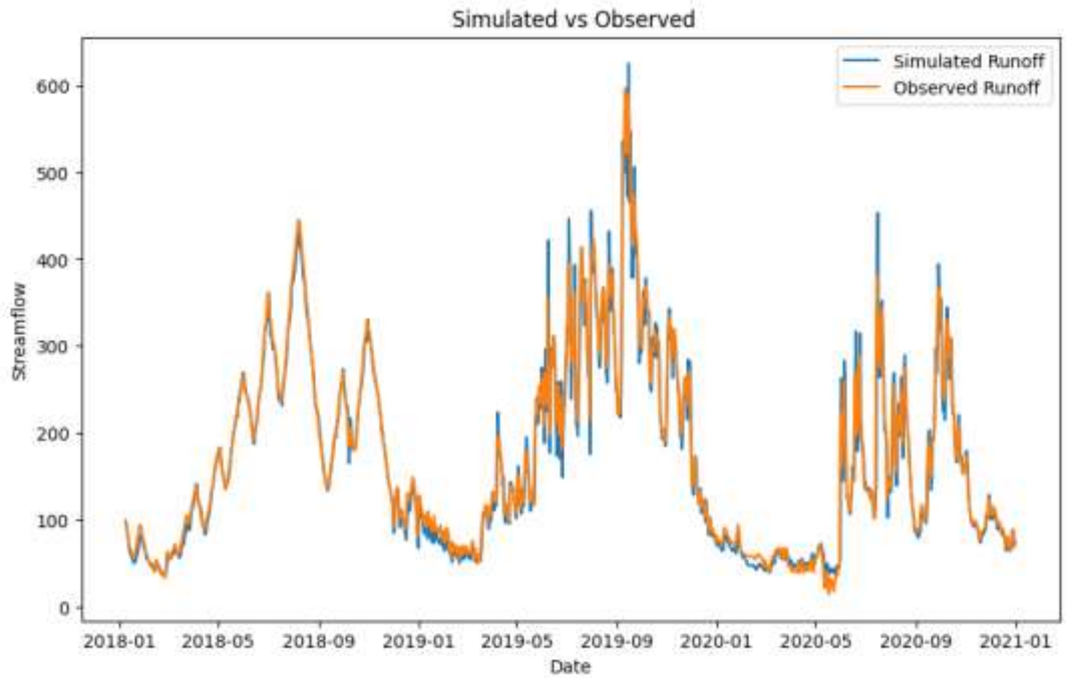


b)

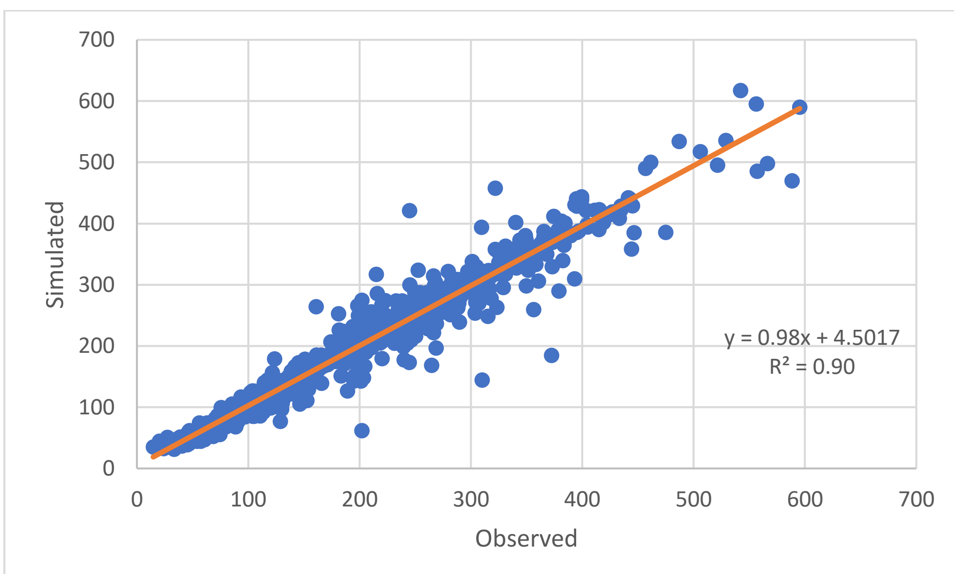
Fig 8. a) Hydrograph from C2-LSTM model prediction b) Regression trendline of model simulations

Figure 9a presents the hydrograph generated by the C3 model, incorporating observed precipitation and NDVI data. The inclusion of NDVI, reflecting vegetation dynamics and its contribution in the runoff process, enhances the model's performance compared to the C2 model. The simulated hydrograph depicts the general pattern of the observed flow for the year 2018. The C3 model demonstrates improved accuracy with higher R^2 and NSE values (0.90 and 0.86, respectively) and a reduced RMSE value of 26.525. However, the model still struggles to accurately simulate extreme runoff events during heavy rainfall periods (rainy seasons).

In Figure 10a, the gap between the simulated and observed hydrograph further decreases under C4 compared to the C3 input combination. The addition of ET to the input combination improves the model's performance and predictive ability. Although there is a marginal improvement in simulating peak runoff values, significant overestimations are still notable. Moreover, the model reproduces low runoff levels, particularly during the dry season. R^2 and NSE values show improvement under the C4 model, reaching 0.92 and 0.90 respectively, slightly outperforming C3 model. RMSE decreases under C4, with a value of 25.881.

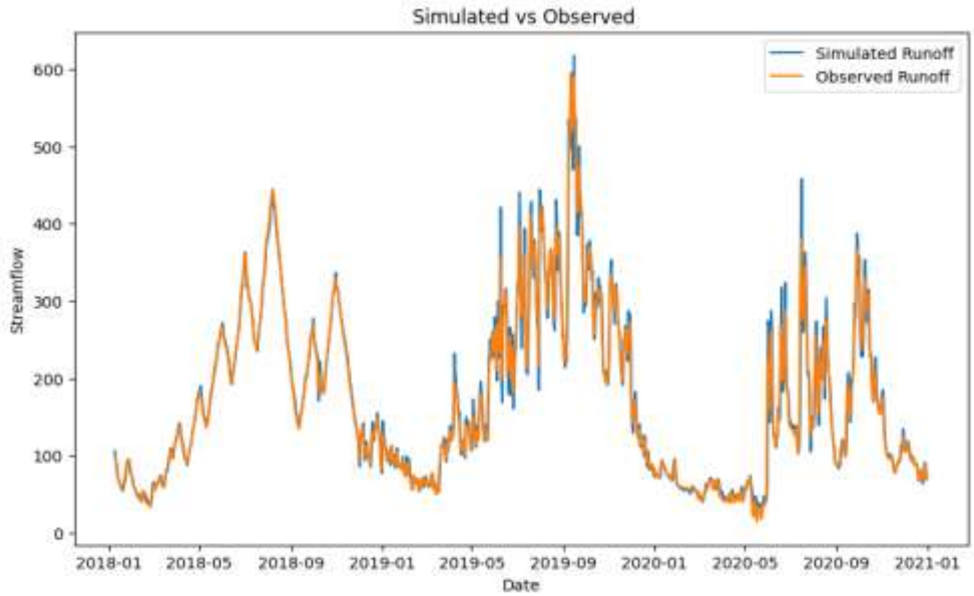


a)

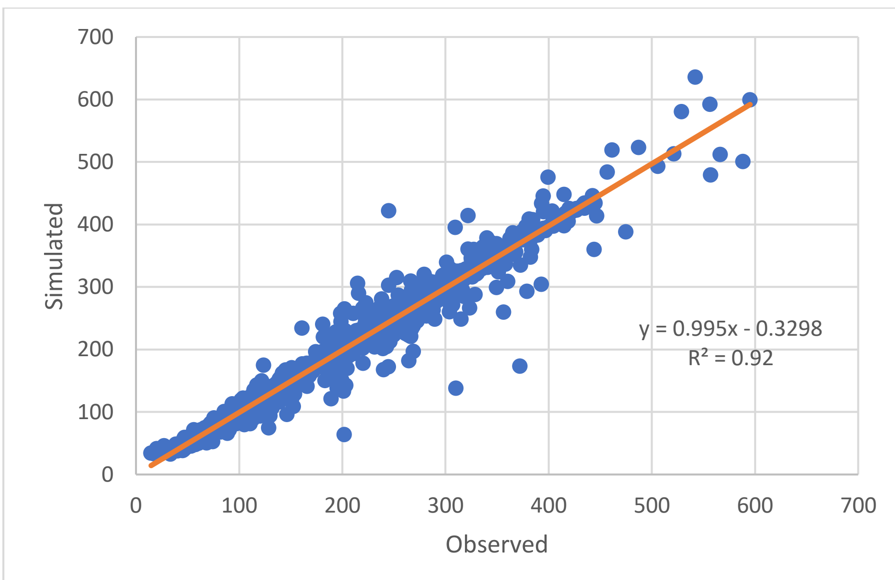


b)

Fig 9. a) Hydrograph from C3-LSTM model prediction b) Regression trendline of model simulations



a)



b)

Fig 10. a) Hydrograph from C4-LSTM model prediction b) Regression trendline of model simulations

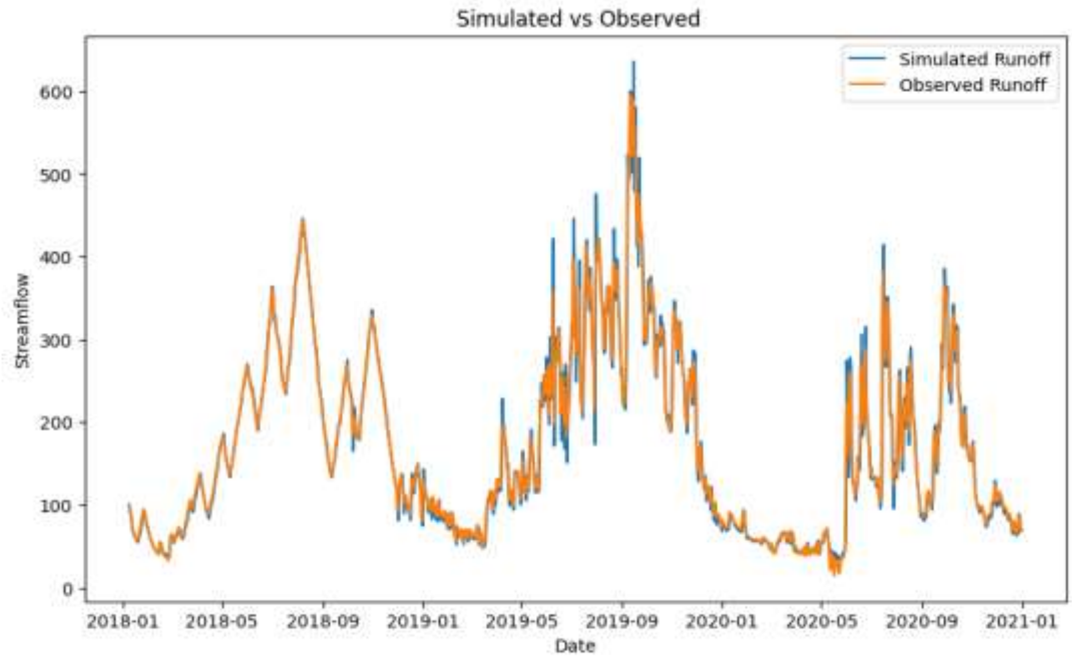
The C5 model, integrating observed precipitation, NDVI, LST, and ET as inputs demonstrated the best performance in reproducing the runoff patterns. The simulated hydrograph generated by the model in figure 11a closely mirrors the observed hydrograph and outperformed the other models in the evaluation indices, with higher R^2 and NSE values of 0.94 and 0.93, respectively. Moreover, C5 achieved the lowest RMSE at 25.860, confirming its precision. The regression trendline in figure 11b indicates a strong linear relationship between the observed and predicted runoff with an R^2 value of 0.9403 closer to 1 suggesting a better fit of the regression line to the data.

The C6 predictions significantly differ from those of C5, primarily due to the incorporation of remotely sensed precipitation, NDVI, LST, and ET as inputs. This notably enhances the model's performance compared to the predictions under C2, where only precipitation from satellite observation was utilized. However, there is a reduction in evaluation metrics, particularly with R^2 and NSE values, compared to the metrics obtained under the C5 simulations with noticeable underestimations depicted in the hydrograph (figure 12b).

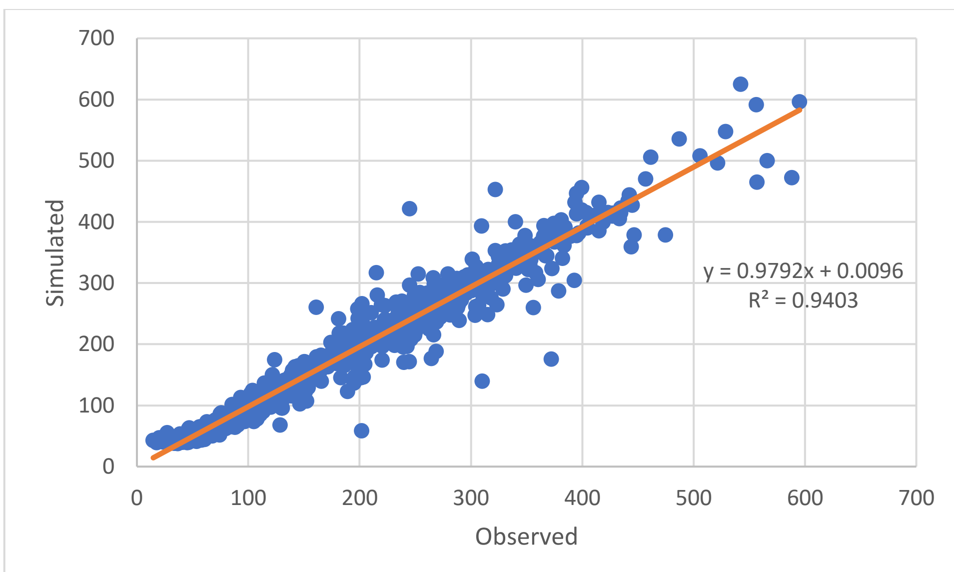
Table 4 provides a comprehensive overview of the performance of all six input combination models.

Table 4. Performance of the six models used to simulate runoff

Model	R2	NSE	RMSE
C1	0.91	0.89	26.006
C2	0.84	0.79	27.109
C3	0.90	0.86	26.525
C4	0.92	0.90	25.881
C5	0.94	0.93	25.860
C6	0.88	0.84	26.536

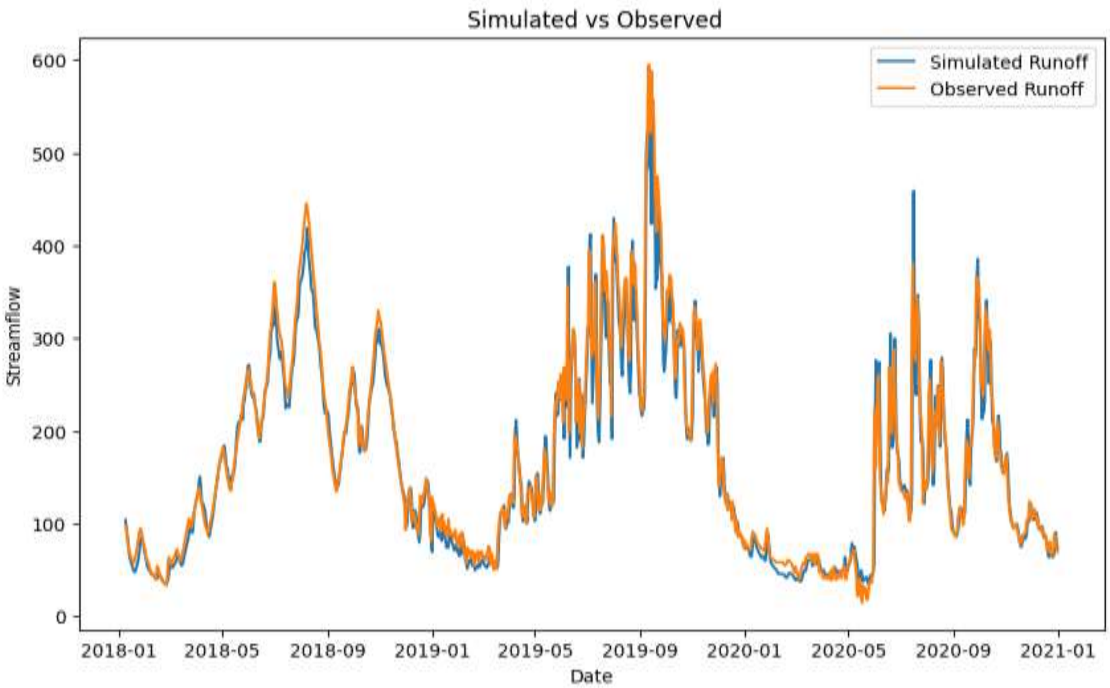


a)

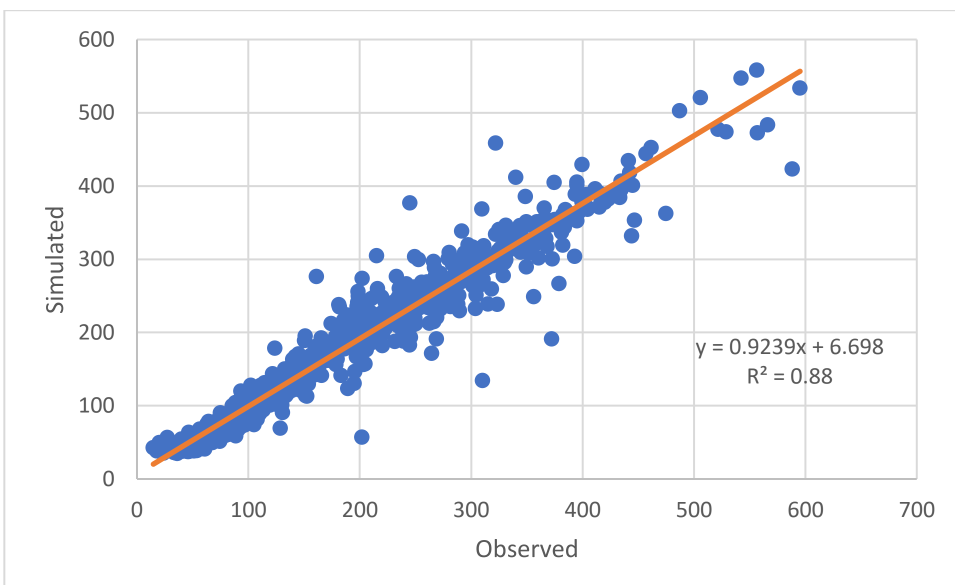


b)

Fig 11. a) Hydrograph from C5-LSTM model prediction b) Regression trendline of model simulations



a)

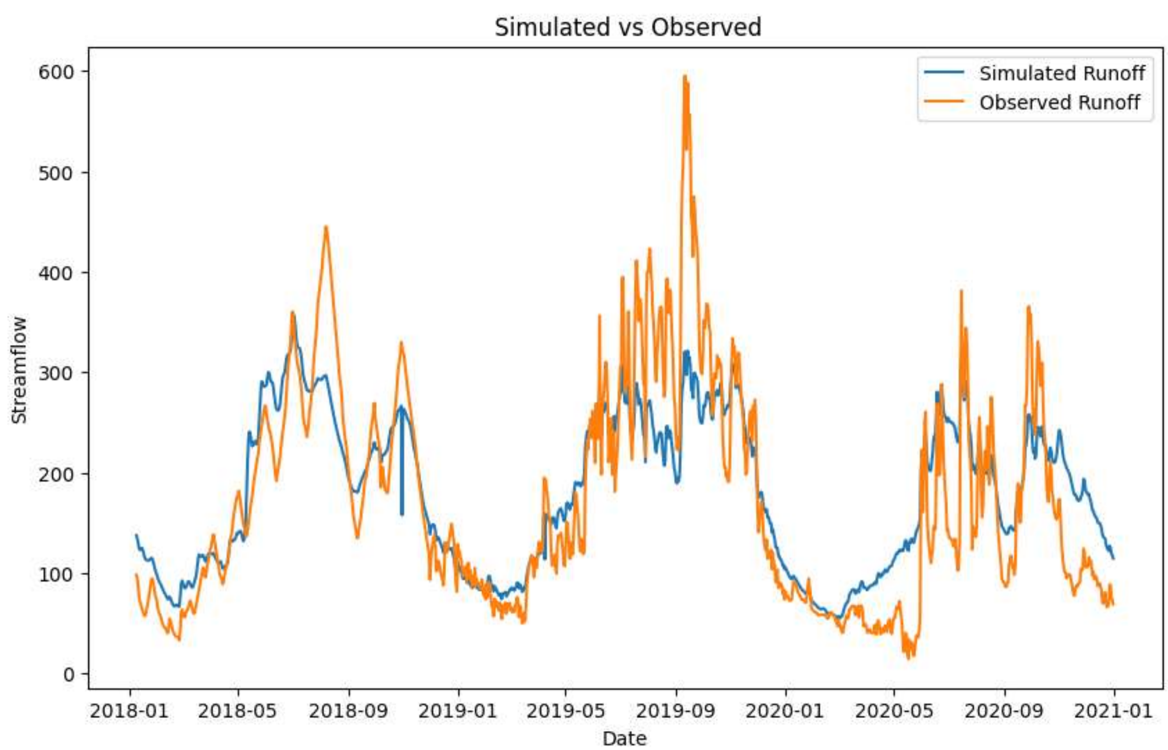


b)

Fig 12. a) Hydrograph from C6-LSTM model prediction b) Regression trendline of model simulations

4.3 SWAT Predictions

The SWAT model has been utilized in previous works to simulate the runoff process in the basin for the same timeframe as the LSTM predictions (2018-2020). SWAT, being a physical model, incorporates mechanical and physical processes within the hydrological system. The dataset used in the SWAT model included observed precipitation from rain gauge stations, land use and land cover (LULC) data, and digital elevation model (DEM) data. However, the prediction accuracy of the SWAT model, with R^2 and NSE values of 0.78 and 0.75 respectively, was lower than that of the best-performing LSTM model (C5). Figure 13a illustrates a notable difference between the simulated and observed hydrographs. The SWAT model struggled to reproduce both peak and low runoff values accurately. Moreover, in Figure 13b, many points deviated from the trendline, indicating discrepancies or inconsistencies between the simulated and observed runoff values.



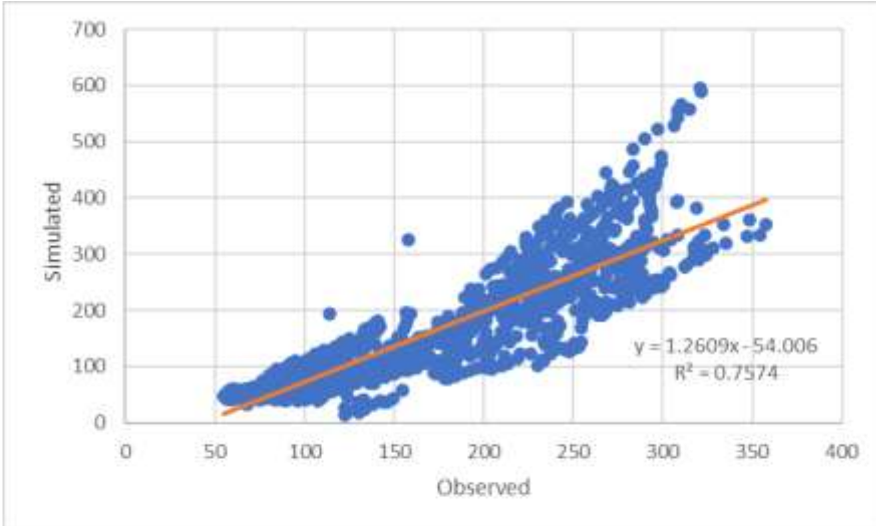


Fig 13. a) Hydrograph from SWAT model prediction b) Regression trendline of SWAT model simulations

4.4 Discussion

Over the years, LSTM architectures have demonstrated potential in hydrological modeling, especially in runoff estimation. In this study, a fully connected LSTM was applied to forecast runoff in Ghana's Pra River Basin, resulting in satisfactory to excellent outcomes. The model performed well when in-situ precipitation data was incorporated in the input combinations, indicating the significant role of precipitation as the primary driver in rainfall-runoff processes. This importance is reflected in the similar hydrograph patterns observed across models C1, C3, C4, and C5.

4.4.1 Runoff Trends

The regression plots displayed in figures 7b, 9b, 10b, and 11b mirror the patterns observed in their corresponding hydrographs in figures 7a, 9a, 10a, and 11a. These figures reveal that points exceeding 400 mm^3/s, indicating high runoff values, diverge from the straight line, indicating overestimation of peak values by the models, despite their overall satisfactory performance. In the Pra River Basin, peak runoff values typically occur during the wet seasons, namely between April - July and September - November. Uncertainties in rainfall events, largely influenced by climate change impacts such as delayed or excessive rains during the wet season, may contribute to errors in simulating peak runoff values. Despite these challenges, models C1, C3, C4, and C5 effectively predicted low runoff values (>100 mm^3/s) during the dry season between December - March.

The hydrographs produced by the C2 and C6 models, illustrated in figures 8a and 12a, exhibit a similar shape or trend. These models utilize satellite-derived precipitation data, resulting in lower prediction accuracies compared to models incorporating in-situ precipitation data. The decreased performance of C2 and C6, relative to C1, C3, C4, and C5, can be attributed to the utilization of remotely sensed precipitation data with a coarse resolution (500m). Despite this limitation, both C2 and C6 achieved satisfactory results in terms of evaluation metrics with R^2 value of 0.84 and 0.88, as outlined in Table 4, underscoring the potential of remote sensing data as an alternative source in the absence of field observations for runoff modeling in basins and watersheds.

4.4.2 Comparison between SWAT and LSTM model

It is important to emphasize that the accuracy achieved by the SWAT model was lower compared to the LSTM model used for runoff estimation in the basin. The SWAT's relatively low performance stems from the intricate interactions, feedback loops, and complexities within the hydrological cycle, as well as the diverse geophysical characteristics, which pose challenges in the development of the physical based-model. SWAT requires a variety of topographical and meteorological data in a watershed with sufficient historical hydrological data. Comparing the evaluation metrics in this study, SWAT yielded R^2 and NSE values of 0.78 and 0.75 respectively, while the LSTM model (specifically C5, the highest-performing LSTM model) achieved R^2 and NSE values of 0.94 and 0.93 respectively.

The superior performance of the LSTM model compared to SWAT aligns with findings from previous studies that have evaluated both models in simulating runoff. For instance, Zou et al. (2016) used an integrated SWAT model to simulate monthly runoff in the Heihe River Basin, achieving R^2 and NSE values of 0.83 and 0.82 respectively. In contrast, Xue et al. (2022) employed FC-LSTM for monthly runoff estimation in the same region, yielding higher R^2 (0.95) and NSE (0.94) values than the SWAT model. Similarly, C. Kim & Kim (2021) compared LSTM and SWAT performances for rainfall-runoff analysis, with SWAT obtaining R^2 and NSE values of 0.90 and 0.80 respectively, while LSTM achieved superior results with $R^2 = 0.98$ and NSE = 0.94. The R^2 and NSE values of 0.94 and 0.93 attained by LSTM in this study highlight the effectiveness of the fully connected LSTM architecture in capturing temporal dynamics and extracting pertinent information from diverse parameters, thereby enhancing runoff modeling accuracy.

4.4.3 Impacts of Remote Sensing Products on Runoff Estimation

In runoff estimation, the selection of input data significantly impacts the accuracy of the results (Bai et al., 2021; Ouma et al., 2021; Xue et al., 2022). Including more variables as inputs generally improves model performance, leading to more precise outcomes. With the exception of in-situ precipitation, the input variables were derived from satellite imagery using remote sensing techniques. Numerous prior studies have demonstrated the effectiveness of remote sensing data in runoff estimation (Pangali Sharma et al., 2020; Yeditha et al., 2022; Y. Zhang et al., 2009; Zhu et al., 2023).

In this research, ET served as a critical indicator of soil and plant water loss, influencing the hydrological process. Models incorporating both ET and in-situ precipitation (C4, C5) outperformed those with precipitation alone. This improvement can be attributed to ET's role in reflecting the available water content, which contributes to the runoff response in hydrological processes.

NDVI was utilized as a biophysical indicator, representing both the physical characteristics, such as vegetation cover, and the biological aspects, including the health and growth of vegetation within the basin. Its role is pivotal in influencing hydrological processes such as interception and evapotranspiration rates, thereby affecting the response of rainfall in the runoff discharge process. The integration of NDVI into the C3 model led to significant performance improvements.

LST was incorporated as a climatic factor due to its reflection of the Earth's surface temperature, influenced by various climatic elements such as solar radiation. Its integration into runoff modeling helps capture the climatic conditions affecting processes like evaporation, thus impacting the available water for runoff. The C5 model, which achieved the best performance, utilized both in-situ and satellite data (NDVI, LST, and ET), yielding the best results. This is attributed to its inclusion of more input variables compared to other models, except for C6, which had the same number of input variables (remotely sensed precipitation, NDVI, LST, and ET).

The successful outcomes of the models with multiple remote sensing input variables underscore the benefits of integrating new features into model inputs, improving estimation results and providing alternative methods for obtaining relevant hydrological data for runoff modeling.

Chapter 5

Conclusion and Recommendation

5.1 Conclusion

In this study, runoff modeling in the Pra River Basin, Ghana, was carried out using a fully connected long short-term memory modeling approach combined with remote sensing data. Literature indicates that LSTM, a form of recurrent neural network architecture, is a sophisticated deep learning algorithm highly effective in runoff modeling. It is apparent that studies utilizing LSTM tend to outperform physically-based models like SWAT, achieving significant accuracies. Moreover, the importance of remote sensing data in supplying crucial hydrological information for runoff modeling has been underscored in several studies.

A promising research opportunity therefore arose with the aim to harness the strength and capabilities of the LSTM network and remote sensing technology to enhance the accuracy of runoff estimations in the Pra River Basin. Previous studies in the area have not delved into this approach for simulating runoff, marking the novelty of this thesis. It highlights the significance of precise runoff modeling for effective water resource management and environmental planning in the region.

Overall, the research questions defined to achieve the aim of the thesis have been answered.

1) *How does the incorporation of remote sensing data, including NDVI, land surface temperature, and evapotranspiration, impact the precision and accuracy of LSTM-based runoff predictions?*

   The methodology involved the development of six LSTM models utilizing various input combinations derived from both in-situ and remote sensing data spanning the period from 2001 to 2020. Remote sensing datasets including CHIRPS-PT, MODIS-NDVI, MODIS-LST, and MODIS-ET were processed using Google Earth Engine. The fully connected LSTM exhibited the capability to capture temporal patterns in the input data, demonstrating strong learning capacity and accurately predicting runoff for the testing period (2018-2020). Models utilizing a single input (observed precipitation) and target (observed runoff) achieved R^2 and NSE values of 0.91 and 0.89, respectively, while those incorporating mixed data of field and remote sensing yielded R^2 and NSE values of 0.94 and 0.93, respectively. Among the six models developed, two utilized single inputs (observed and satellite-based precipitation), while the remaining four integrated both remote sensing and in-situ data. Integrated models consistently outperformed single-input models, with improved accuracy and performance attributed to the inclusion of additional variables (NDVI, LST, and ET).

2) *How does the overall performance of the LSTM model compare with Soil Water Assessment Tool physical-based model in estimating monthly runoff discharge in the Pra Basin, Ghana?*

The top-performing LSTM model incorporated in-situ precipitation, satellite-derived NDVI, LST, and ET as inputs, achieving R^2 and NSE values of 0.94 and 0.93, respectively. This model's performance was then compared to that of a SWAT model developed for the same basin during the testing period (2018-2020), which yielded prediction accuracy metrics of R^2 = 0.78 and NSE = 0.75. Clearly, the LSTM outperformed the SWAT model in simulating runoff within the basin, emphasizing the robustness and efficiency of deep learning, particularly LSTM, in runoff simulation.

5.2 Limitations and Future Applications

This study encountered limitations in several areas. Firstly, there were challenges regarding the availability of sufficient in-situ precipitation data, which affected the temporal range of the study period. This resulted in a limited number of data points for model development, with the study period set between 2001 and 2020. The absence of earlier precipitation records from meteorological stations necessitated preprocessing of the in-situ data to address missing values and ensure consistency. Additionally, remote sensing data had to be obtained within the same timeframe to maintain dataset coherence for LSTM modeling.

In future studies, addressing data sufficiency challenges and increasing the amount of input datasets might involve relying solely on remote sensing datasets. Remote sensing offers extensive historical data, and previous works including this study has shown that models utilizing such data can accurately simulate runoffs, presenting an alternative source of hydrological data for runoff modeling. Additionally, improving model accuracy could entail the inclusion of supplementary variables. Therefore, variables like the soil moisture index, which impacts infiltration and runoff in hydrological processes, could be explored and integrated into forthcoming research.

Additionally, the modeling process did not consider the various spatial resolutions of remote sensing products, as all data were resampled to a 500m resolution. While the simulated runoffs captured the general trends in the basin's runoff and achieved satisfactory results, future research could explore the sensitivity of models to different spatial resolutions by incorporating remote sensing images with varied resolutions.

Lastly, this study employed a lumped model approach to estimate basin runoff, treating the study area as a single unit. However, the basin comprises three sub-basins, raising concerns about spatial dependency. The runoff recorded at downstream points or within sub-basins may not fully represent basin-wide runoff. Additionally, temporal dependency exists, where the time taken for surface water to flow between points impacts recorded downstream flow. Future research should address these issues by considering spatiotemporal dependency and exploring distributed models to better simulate sub-basin runoffs, providing a more accurate representation of runoff in the Pra River Basin.

References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system. *Journal of Hydrology*, *87*(1), 61–77. https://doi.org/https://doi.org/10.1016/0022-1694(86)90115-0

Abrahart, R., Kneale, P. E., & See, L. M. (2004). *Neural networks for hydrological modeling*. CRC Press.

Akrasi, S. A., & Ansa-Asare, O. D. (2008). Assessing sediment and nutrient transport in the Pra Basin of Ghana. *West African Journal of Applied Ecology*, *13*(1), 45–54.

Alaba, B., Mateu, J., Pebesma, P. E., & Cabral, P. (2010). *HYDROLOGIC MODELING AND UNCERTAINTY ANALYSIS OF AN UNGAUGED WATERSHED USING MAPWINDOW-SWAT*.

Amatya, D. M., Walega, A., Callahan, T. J., Morrison, A., Vulava, V., Hitchcock, D. R., Williams, T. M., & Epps, T. (2022). Storm event analysis of four forested catchments on the Atlantic coastal plain using a modified SCS-CN rainfall-runoff model. *Journal of Hydrology*, *608*, 127772. https://doi.org/https://doi.org/10.1016/j.jhydrol.2022.127772

Awotwi, A., Anornu, G. K., Quaye-Ballard, J. A., Annor, T., Forkuo, E. K., Harris, E., Agyekum, J., & Terlabie, J. L. (2019a). Water balance responses to land-use/land-cover changes in the Pra River Basin of Ghana, 1986–2025. *CATENA*, *182*, 104129. https://doi.org/https://doi.org/10.1016/j.catena.2019.104129

Awotwi, A., Anornu, G. K., Quaye-Ballard, J. A., Annor, T., Forkuo, E. K., Harris, E., Agyekum, J., & Terlabie, J. L. (2019b). Water balance responses to land-use/land-cover changes in the Pra River Basin of Ghana, 1986–2025. *CATENA*, *182*, 104129. https://doi.org/https://doi.org/10.1016/j.catena.2019.104129

Awotwi, A., Anornu, G. K., Quaye-Ballard, J., Annor, T., & Forkuo, E. K. (2017a). Analysis of climate and anthropogenic impacts on runoff in the Lower Pra River Basin of Ghana. *Heliyon*, *3*(12), e00477. https://doi.org/https://doi.org/10.1016/j.heliyon.2017.e00477

Awotwi, A., Anornu, G. K., Quaye-Ballard, J., Annor, T., & Forkuo, E. K. (2017b). Analysis of climate and anthropogenic impacts on runoff in the Lower Pra River Basin of Ghana. *Heliyon*, *3*(12).

Bai, Y., Bezak, N., Zeng, B., Li, C., Sapač, K., & Zhang, J. (2021). Daily runoff forecasting using a cascade long short-term memory model that considers different variables. *Water Resources Management*, *35*, 1167–1181.

Beven, K. J. (2011). *Rainfall-runoff modelling: the primer*. John Wiley \& Sons.

Bhattacharjee, N. V, & Tollner, E. W. (2016). Improving management of windrow composting systems by modeling runoff water quality dynamics using recurrent neural network. *Ecological Modelling*, *339*, 68–76. https://doi.org/https://doi.org/10.1016/j.ecolmodel.2016.08.011

Boulmaiz, T., Guermoui, M., & Boutaghane, H. (2020). Impact of training data size on the LSTM performances for rainfall–runoff modeling. *Modeling Earth Systems and Environment*, *6*(4), 2153–2164. https://doi.org/10.1007/s40808-020-00830-w

bukhari, D., Wang, Y., & Wang, H. (2017). Multilingual Convolutional, Long Short-Term Memory, Deep Neural Networks for Low Resource Speech Recognition. *Procedia Computer Science*, *107*, 842–847. https://doi.org/https://doi.org/10.1016/j.procs.2017.03.179

Carriere, P. P. E., Mohaghegh, S., & Gaskari, R. (1996). Performance of a Virtual Runoff Hydrograph System. *Journal of Water Resources Planning and Management*, *122*, 421–427. https://api.semanticscholar.org/CorpusID:109143779

Chen, X., Long, D., Hong, Y., Zeng, C., & Yan, D. (2017). Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the U pper B rahmaputra R iver basin? *Water Resources Research*, *53*(3), 2431–2466.

Chen, Y., Li, J., & Xu, H. (2016). Improving flood forecasting capability of physically based distributed hydrological models by parameter optimization. *Hydrology and Earth System Sciences*, *20*(1), 375–392. https://doi.org/10.5194/hess-20-375-2016

Commission, W. R. (2012). Pra River Basin—integrated water resources management plan. *Water Resources Commission, Accra*.

Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in Physical Geography: Earth and Environment*, *25*(1), 80–108. https://doi.org/10.1177/030913330102500104

de Lavenne, A., Thirel, G., Andréassian, V., Perrin, C., & Ramos, M.-H. (2016). Spatial variability of the parameters of a semi-distributed hydrological model. *Proceedings of the International Association of Hydrological Sciences*, *373*, 87–94. https://doi.org/10.5194/piahs-373-87-2016

Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A Review on Hydrological Models. *Aquatic Procedia*, *4*, 1001–1007. https://doi.org/10.1016/j.aqpro.2015.02.126

Ekundayo, A. M. (2020). The Role of Hydrologic Processes in Catchments. In *International Journal of Innovative Science and Research Technology* (Vol. 5, Issue 8). www.ijisrt.com1216

Farabet, C., Couprie, C., Najman, L., & LeCun, Y. (2013). Learning Hierarchical Features for Scene Labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1915–1929. https://doi.org/10.1109/TPAMI.2012.231

Feldman, A. D. (2000). *Hydrologic modeling system HEC-HMS: technical reference manual*. US Army Corps of Engineers, Hydrologic Engineering Center.

Garbrecht, J. D. (2006). Comparison of three alternative ANN designs for monthly rainfall-runoff simulation. *Journal of Hydrologic Engineering*, *11*(5), 502 – 505. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:5(502)

Gelly, G., & Gauvain, J.-L. (2018). Optimization of RNN-Based Speech Activity Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *26*(3), 646–656. https://doi.org/10.1109/TASLP.2017.2769220

Gholami, V., & Sahour, H. (2022). Simulation of rainfall-runoff process using an artificial neural network (ANN) and field plots data. *Theoretical and Applied Climatology*, *147*(1), 87–98. https://doi.org/10.1007/s00704-021-03817-4

Granata, F., Gargano, R., & De Marinis, G. (2016). Support Vector Regression for Rainfall-Runoff Modeling in Urban Drainage: A Comparison with the EPA's Storm Water Management Model. *Water*, *8*(3). https://doi.org/10.3390/w8030069

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: a search space odyssey (2015). *ArXiv Preprint ArXiv:1503.04069*.

Han, H., Choi, C., Jung, J., & Kim, H. S. (2021). Deep learning with long short term memory based sequence-to-sequence model for rainfall-runoff simulation. *Water*, *13*(4), 437.

Hartanto, I. M., van der Kwast, J., Alexandridis, T. K., Almeida, W., Song, Y., van Andel, S. J., & Solomatine, D. P. (2017). Data assimilation of satellite-based actual evapotranspiration in a distributed hydrological model of a controlled water system. *International Journal of Applied Earth Observation and Geoinformation*, *57*, 123–135.

Haykin, S., & Network, N. (2004). A comprehensive foundation. *Neural Networks*, *2*(2004), 41.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., & Kingsbury, B. (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, *29*(6), 82–97. https://doi.org/10.1109/MSP.2012.2205597

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*, 1735–1780.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/https://doi.org/10.1016/0893-6080(89)90020-8

Hsu, K. lin, Gupta, H. V., & Soroosh Sorooshian. (1997). Application of a recurrent neural network to rainfall-runoff modeling. *Proceedings of the 1997 24th Annual Water Resources Planning and Management Conference ; Conference Date: 06-04-1997 Through 09-04-1997*, 68–73.

Hu, C., Wu, Q., Li, H., Jian, S., Li, N., & Lou, Z. (2018). Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation. *Water*, *10*(11). https://doi.org/10.3390/w10111543

Huang, Q., Qin, G., Zhang, Y., Tang, Q., Liu, C., Xia, J., Chiew, F. H. S., & Post, D. (2020). Using Remote Sensing Data-Based Hydrological Model Calibrations for Predicting Runoff in Ungauged or Poorly Gauged Catchments. *Water Resources Research*, *56*(8). https://doi.org/10.1029/2020WR028205

Jackson, B., Mcintyre, N., Pechlivanidis, I. G., Jackson, B. M., Mcintyre, N. R., & Wheater, H. S. (2011). Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and app CATCHMENT SCALE HYDROLOGICAL MODELLING: A REVIEW OF MODEL TYPES, CALIBRATION APPROACHES AND UNCERTAINTY ANALYSIS METHODS IN THE CONTEXT OF RECENT DEVELOPMENTS IN TECHNOLOGY AND APPLICATIONS. In *Article in Global NEST Journal* (Vol. 13, Issue 3). https://www.researchgate.net/publication/266341273

Jaiswal, R. K., Ali, S., & Bharti, B. (2020). Comparative evaluation of conceptual and physical rainfall–runoff models. *Applied Water Science*, *10*(1), 48. https://doi.org/10.1007/s13201-019-1122-6

Jehanzaib, M., Ajmal, M., Achite, M., & Kim, T.-W. (2022). Comprehensive Review: Advancements in Rainfall-Runoff Modelling for Flood Mitigation. *Climate*, *10*(10). https://doi.org/10.3390/cli10100147

Jiang, D., & Wang, K. (2019). The role of satellite-based remote sensing in improving simulated streamflow: A review. *Water*, *11*(8), 1615.

Kalma, J. D., McVicar, T. R., & McCabe, M. F. (2008). Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics*, *29*, 421–469.

Khakbaz, B., Imam, B., Hsu, K., & Sorooshian, S. (2012). From lumped to distributed via semi-distributed: Calibration strategies for semi-distributed hydrologic models. *Journal of Hydrology*, *418–419*, 61–77. https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.02.021

Kim, C., & Kim, C.-S. (2021). Comparison of the performance of a hydrologic model and a deep learning technique for rainfall- runoff analysis. *Tropical Cyclone Research and Review*, *10*(4), 215–222. https://doi.org/https://doi.org/10.1016/j.tcrr.2021.12.001

Kim, H. W., Hwang, K., Mu, Q., Lee, S. O., & Choi, M. (2012). Validation of MODIS 16 global terrestrial evapotranspiration products in various climates and land cover types in Asia. *KSCE Journal of Civil Engineering*, *16*, 229–238.

Kişi, Ö. (2008). River flow forecasting and estimation using different artificial neural network techniques. *Hydrology Research*, *39*(1), 27–40. https://doi.org/10.2166/nh.2008.026

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 25). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Rudari, R., Pulvirenti, L., Boni, G., Fascetti, F., & Pierdicca, N. (2016). Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model. *International Journal of Applied Earth Observation and Geoinformation*, *48*, 131–145.

Li, C., Zhu, L., He, Z., Gao, H., Yang, Y., Yao, D., & Qu, X. (2019). Runoff prediction method based on adaptive Elman Neural Network. *Water (Switzerland)*, *11*(6). https://doi.org/10.3390/w11061113

Li, J., Qian, K., Liu, Y., Yan, W., Yang, X., Luo, G., & Ma, X. (2022). LSTM-Based Model for Predicting Inland River Runoff in Arid Region: A Case Study on Yarkant River, Northwest China. *Water*, *14*(11). https://doi.org/10.3390/w14111745

Li, W., Kiaghadi, A., & Dawson, C. (2021). High temporal resolution rainfall–runoff modeling using long-short-term-memory (LSTM) networks. *Neural Computing and Applications*, *33*(4), 1261–1278. https://doi.org/10.1007/s00521-020-05010-6

Li, Y., Grimaldi, S., Pauwels, V. R. N., & Walker, J. P. (2018). Hydrologic model calibration using remotely sensed soil moisture and discharge measurements: The impact on predictions at gauged and ungauged locations. *Journal of Hydrology*, *557*, 897–909.

Liu, Y., Zhang, T., Kang, A., Li, J., & Lei, X. (2021). Research on runoff simulations using deep-learning methods. *Sustainability (Switzerland)*, *13*(3), 1–20. https://doi.org/10.3390/su13031336

Madsen, H. (2000). Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology*, *235*(3), 276–288. https://doi.org/https://doi.org/10.1016/S0022-1694(00)00279-1

Majumder Mrinmoy and Barua, A. and B. B. (2010). An Overview of Hydrologic Modeling. In M. Jana Bipal K. and Majumder (Ed.), *Impact of Climate Change on Natural Resource Management* (pp. 375–391). Springer Netherlands. https://doi.org/10.1007/978-90-481-3581-3_21

Man, Y., Yang, Q., Shao, J., Wang, G., Bai, L., & Xue, Y. (2023). Enhanced LSTM Model for Daily Runoff Prediction in the Upper Huai River Basin, China. *Engineering*, *24*, 229–238. https://doi.org/https://doi.org/10.1016/j.eng.2021.12.022

Moradkhani, H. (2008). Hydrologic Remote Sensing and Land Surface Data Assimilation. *Sensors*, *8*(5), 2986–3004. https://doi.org/10.3390/s8052986

Moradkhani Hamid and Sorooshian, S. (2008). General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis. In K.-L. and C. E. and T. B. and V. M. and V. G. Sorooshian Soroosh and Hsu (Ed.), *Hydrological Modelling and the Water Cycle: Coupling the Atmospheric and Hydrological Models* (pp. 1–24). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77843-1_1

Mou, L., Ghamisi, P., & Zhu, X. X. (2017). Deep Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(7), 3639–3655. https://doi.org/10.1109/TGRS.2016.2636241

Mukerji, A., Chatterjee, C., & Singh Raghuwanshi, N. (2009). Flood forecasting using ANN, neuro-fuzzy, and neuro-GA models. *Journal of Hydrologic Engineering*, *14*(6), 647–652. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000040

Nagesh Kumar, D., Raju, K. S., & Sathish, T. (2004). River Flow Forecasting using Recurrent Neural Networks. In *Water Resources Management* (Vol. 18).

Nagesh Kumar, D., Srinivasa Raju, K., & Sathish, T. (2004). River Flow Forecasting using Recurrent Neural Networks. *Water Resources Management*, *18*(2), 143–161. https://doi.org/10.1023/B:WARM.0000024727.94701.12

Nourani, V., Fard, A. F., Gupta, H. V, Goodrich, D. C., & Niazi, F. (2017). Hydrological model parameterization using NDVI values to account for the effects of land cover change on the rainfall–runoff response. *Hydrology Research*, *48*(6), 1455–1473.

Nourani, V., & Kalantari, O. (2010). Integrated Artificial Neural Network for Spatiotemporal Modeling of Rainfall–Runoff–Sediment Processes. *Environmental Engineering Science*, *27*(5), 411–422. https://doi.org/10.1089/ees.2009.0353

Ouma, Y. O., Cheruyot, R., & Wachera, A. N. (2021). Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: case study of Nzoia hydrologic basin. *Complex & Intelligent Systems*, 1–24.

Owusu, C., Adjei, K. A., & Odai, S. N. (2019a). Evaluation of Satellite Rainfall Estimates in the Pra Basin of Ghana. *Environmental Processes*, *6*(1), 175–190. https://doi.org/10.1007/s40710-018-0344-1

Owusu, C., Adjei, K. A., & Odai, S. N. (2019b). Evaluation of Satellite Rainfall Estimates in the Pra Basin of Ghana. *Environmental Processes*, *6*(1), 175–190. https://doi.org/10.1007/s40710-018-0344-1

Owusu, G., Owusu, A. B., Amankwaa, E. F., & Eshun, F. (2017). Analyses of freshwater stress with a couple ground and surface water model in the Pra Basin, Ghana. *Applied Water Science*, *7*(1), 137–153. https://doi.org/10.1007/s13201-015-0279-x

Pan, T., Wang, R., & Lai, J. (2007). A deterministic linearized recurrent neural network for recognizing the transition of rainfall–runoff processes. *Advances in Water Resources*, *30*(8), 1797–1814. https://doi.org/https://doi.org/10.1016/j.advwatres.2007.02.009

Pandi, D., Kothandaraman, S., & Kuppusamy, M. (2021). Hydrological models: a review. *International Journal of Hydrology Science and Technology*, *12*(3), 223–242. https://doi.org/10.1504/IJHST.2021.117540

Pangali Sharma, T. P., Zhang, J., Khanal, N. R., Prodhan, F. A., Paudel, B., Shi, L., & Nepal, N. (2020). Assimilation of snowmelt runoff model (SRM) using satellite remote sensing data in Budhi Gandaki River Basin, Nepal. *Remote Sensing*, *12*(12), 1951.

Park, J.-Y., Ahn, S.-R., Hwang, S.-J., Jang, C.-H., Park, G.-A., & Kim, S.-J. (2014). Evaluation of MODIS NDVI and LST for indicating soil moisture of forest areas based on SWAT modeling. *Paddy and Water Environment*, *12*(1), 77–88. https://doi.org/10.1007/s10333-014-0425-3

Poornima, S., & Pushpalatha, M. (2019). Prediction of Rainfall Using Intensified LSTM Based Recurrent Neural Network with Weighted Linear Units. *Atmosphere*, *10*(11). https://doi.org/10.3390/atmos10110668

Rezaeianzadeh, M., Stein, A., Tabari, H., Abghari, H., Jalalkamali, N., Hosseinipour, E. Z., & Singh, V. P. (2013a). Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. *International Journal of Environmental Science and Technology*, *10*(6), 1181–1192. https://doi.org/10.1007/s13762-013-0209-0

Rezaeianzadeh, M., Stein, A., Tabari, H., Abghari, H., Jalalkamali, N., Hosseinipour, E. Z., & Singh, V. P. (2013b). Assessment of a conceptual hydrological model and artificial neural networks for daily outflows forecasting. *International Journal of Environmental Science and Technology*, *10*(6), 1181–1192. https://doi.org/10.1007/s13762-013-0209-0

Rezaeianzadeh, M., Tabari, H., Arabi Yazdi, A., Isik, S., & Kalin, L. (2014). Flood flow forecasting using ANN, ANFIS and regression models. *Neural Computing and Applications*, *25*(1), 25–37. https://doi.org/10.1007/s00521-013-1443-6

Roy, B., Singh, M. P., Kaloop, M. R., Kumar, D., Hu, J.-W., Kumar, R., & Hwang, W.-S. (2021). Data-driven approach for rainfall-runoff modelling using equilibrium optimizer coupled extreme learning machine and deep neural network. *Applied Sciences*, *11*(13), 6238.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Rykin, I., Shagnieva, A., Panidi, E., & Tsepelev, V. (2019). Highly discrete mapping of the growing season time frames and time dynamics. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 357–361.

Sahoo, B. B., Jha, R., Singh, A., & Kumar, D. (2019). Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophysica*, *67*(5), 1471–1481. https://doi.org/10.1007/s11600-019-00330-1

Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R., & Hauck, L. M. (2001). VALIDATION OF THE SWAT MODEL ON A LARGE RWER BASIN WITH POINT AND NONPOINT SOURCES1. *JAWRA Journal of the American Water Resources Association*, *37*(5), 1169–1188. https://doi.org/https://doi.org/10.1111/j.1752-1688.2001.tb03630.x

Shamseldin, A. Y. (1997). Application of a neural network technique to rainfall-runoff modelling. *Journal of Hydrology*, *199*(3), 272–294. https://doi.org/https://doi.org/10.1016/S0022-1694(96)03330-6

Shewalkar, A., nyavanandi, D., & Ludwig, S. A. (2019). Performance Evaluation of Deep neural networks Applied to Speech Recognition: Rnn, LSTM and GRU. *Journal of Artificial Intelligence and Soft Computing Research*, *9*(4), 235–245. https://doi.org/10.2478/jaiscr-2019-0006

Sirisena, T. A. J. G., Maskey, S., & Ranasinghe, R. (2020). Hydrological model calibration with streamflow and remote sensing-based evapotranspiration data in a data poor basin. *Remote Sensing*, *12*(22), 3768.

Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Avant, B., Overview, A., & Muche, M. (2018). *An Overview of Rainfall-Runoff Model Types an Overview of Rainfall-Runoff Model Types An Overview of Rainfall-Runoff Model Types*. https://scholarsarchive.byu.edu/iemssconferencehttps://scholarsarchive.byu.edu/iemssconference/2018/Stream-C/41Thisoralpresentation

Son, N. T., Chen, C. F., Chen, C. R., Minh, V. Q., & Trung, N. H. (2014). A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agricultural and Forest Meteorology*, *197*, 52–64.

Sutskever Google, I., Vinyals Google, O., & Le Google, Q. V. (n.d.). *Sequence to Sequence Learning with Neural Networks*.

Tabari, H., Marofi, S., Abyaneh, H. Z., & Sharifi, M. R. (2010). Comparison of artificial neural network and combined models in estimating spatial distribution of snow depth and snow water equivalent in Samsami basin of Iran. *Neural Computing and Applications*, *19*(4), 625–635. https://doi.org/10.1007/s00521-009-0320-9

Testa, S., Soudani, K., Boschetti, L., & Borgogno Mondino, E. (2018). MODIS-derived EVI, NDVI and WDRVI time series to estimate phenological metrics in French deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, *64*, 132–144. https://doi.org/https://doi.org/10.1016/j.jag.2017.08.006

Thakur, J. K., Singh, S. K., & Ekanthalu, V. S. (2017). Integrating remote sensing, geographic information systems and global positioning system techniques with hydrological modeling. *Applied Water Science*, *7*(4), 1595–1608.

Thapa, S., Zhao, Z., Li, B., Lu, L., Fu, D., Shi, X., Tang, B., & Qi, H. (2020). Snowmelt-Driven Streamflow Prediction Using Machine Learning Techniques (LSTM, NARX, GPR, and SVR). *Water*, *12*(6). https://doi.org/10.3390/w12061734

Varvani, J., Khaleghi, M. R., & Gholami, V. (2019). Investigation of the Relationship between Sediment Graph and Hydrograph of Flood Events (Case Study: Gharachay River Tributaries, Arak, Iran). *Water Resources*, *46*(6), 883–893. https://doi.org/10.1134/S0097807819060204

Xue, H., Liu, J., Dong, G., Zhang, C., & Jia, D. (2022). Runoff Estimation in the Upper Reaches of the Heihe River Using an LSTM Model with Remote Sensing Data. *Remote Sensing*, *14*(10). https://doi.org/10.3390/rs14102488

Yang, J., Reichert, P., Abbaspour, K. C., Xia, J., & Yang, H. (2008). Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology*, *358*(1), 1–23. https://doi.org/https://doi.org/10.1016/j.jhydrol.2008.05.012

Yeditha, P. K., Rathinasamy, M., Neelamsetty, S. S., Bhattacharya, B., & Agarwal, A. (2022). Investigation of satellite precipitation product driven rainfall-runoff model using deep learning approaches in two different catchments of India. *Journal of Hydroinformatics*, *24*(1), 16–37.

Yin, H., Zhang, X., Wang, F., Zhang, Y., Xia, R., & Jin, J. (2021). Rainfall-runoff modeling using LSTM-based multi-state-vector sequence-to-sequence model. *Journal of Hydrology*, *598*, 126378. https://doi.org/https://doi.org/10.1016/j.jhydrol.2021.126378

Yokoo, K., Ishida, K., Ercan, A., Tu, T., Nagasato, T., Kiyama, M., & Amagasaki, M. (2022a). Capabilities of deep learning models on learning physical relationships: Case of rainfall-runoff modeling with LSTM. *Science of The Total Environment*, *802*, 149876.

Yokoo, K., Ishida, K., Ercan, A., Tu, T., Nagasato, T., Kiyama, M., & Amagasaki, M. (2022b). Capabilities of deep learning models on learning physical relationships: Case of rainfall-runoff modeling with LSTM. *Science of The Total Environment*, *802*, 149876. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.149876

Young, C.-C., & Liu, W.-C. (2015). Prediction and modelling of rainfall–runoff during typhoon events using a physically-based and artificial neural network hybrid model. *Hydrological Sciences Journal*, *60*(12), 2102–2116. https://doi.org/10.1080/02626667.2014.959446

Young, C.-C., Liu, W.-C., & Wu, M.-C. (2017). A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. *Applied Soft Computing*, *53*, 205–216. https://doi.org/https://doi.org/10.1016/j.asoc.2016.12.052

Yuan, X., Chen, C., Lei, X., Yuan, Y., & Muhammad Adnan, R. (2018). Monthly runoff forecasting based on LSTM–ALO model. *Stochastic Environmental Research and Risk Assessment*, *32*(8), 2199–2212. https://doi.org/10.1007/s00477-018-1560-y

Zhang, J., Zhu, Y., Zhang, X., Ye, M., & Yang, J. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *Journal of Hydrology*, *561*, 918–929. https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.04.065

Zhang, Y., Chiew, F. H. S., Zhang, L., & Li, H. (2009). Use of remotely sensed actual evapotranspiration to improve rainfall-runoff modeling in Southeast Australia. *Journal of Hydrometeorology*, *10*(4), 969–980. https://doi.org/10.1175/2009JHM1061.1

Zhao, H., Yang, Z., Di, L., Li, L., & Zhu, H. (2009). Crop phenology date estimation based on NDVI derived from the reconstructed MODIS daily surface reflectance data. *2009 17th International Conference on Geoinformatics*, 1–6.

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, *115*, 213–237. https://doi.org/https://doi.org/10.1016/j.ymssp.2018.05.050

Zhou, L., Liu, P., Zhang, X., Cheng, L., Xia, Q., Xie, K., Liu, W., & Xia, J. (2023). Improving structure identifiability of hydrological processes by temporal sensitivity with a flexible modeling framework. *Journal of Hydrology*, *616*, 128843. https://doi.org/https://doi.org/10.1016/j.jhydrol.2022.128843

Zhu, S., Wei, J., Zhang, H., Xu, Y., & Qin, H. (2023). Spatiotemporal deep learning rainfall-runoff forecasting combined with remote sensing precipitation products in large scale basins. *Journal of Hydrology*, *616*, 128727.

Zou, S., Ruan, H., Lu, Z., Yang, D., Xiong, Z., & Yin, Z. (2016). Runoff Simulation in the Upper Reaches of Heihe River Basin Based on the RIEMS–SWAT Model. *Water*, *8*(10). https://doi.org/10.3390/w8100455