A Comparison of Climate-Driven Deep Learning Ensemble and SWAT+ Models for Daily Streamflow Simulation in the Niger River Basin, West Africa

Jefia Idolor[1], Aderemi Adediji[1], and Peter Adebayo Idowu[1]

[1]Obafemi Awolowo University

April 29, 2024

Abstract

Streamflow monitoring is very important for planning and management of water resources in watersheds, and their prediction accuracy is crucial for decision-making. The Niger River Basin is a transboundary resource, shared by nine West African Countries and Algeria and, a large portion of the population rely on the basin for rain-fed agriculture and hydropower. Over the years, the basin's streamflow regime has been altered due to climate change, drought, desertification and establishment of Dams. This research describes a novel Deep Learning framework comprised of Bidirectional-Long Short-Term Memory (LSTM) requiring Antecedent Precipitation Index (API) and meteorological variables, preprocessed using Normal Quantile Transform (NQT) as input drivers and, compared with the Soil and Water Assessment Tool (SWAT+) for streamflow prediction. NQT-API-LSTM which considers catchment wetness and seasonality, was forced with reanalyzed climate (1979–2021) while, SWAT+ was driven with biophysical data and reanalyzed climate (2010–2020). The very high performance of both NQT-API-LSTM and SWAT+ models showed the models were reliable and can predict regulated flows with reasonable certainty. However, NQT-API-LSTM outperformed SWAT+ at Lokoja watershed and, realistically captured the influence of seasonal climate and regional groundwater dynamics from upstream catchments including the Sahara Desert on the Benue, Guinean, Sahelian and Sudan Flood. Overall, NQT-API-LSTM could be used successfully for watershed-scale streamflow prediction without the need for continuous ground support data, a benefit for sparsely gauged West African River Basins, while SWAT+ could be used as an alternative, particularly, to evaluate the watershed's response to land use/land cover changes.

Hosted file

`Idolor Adediji Idowu Manuscript.docx` available at https://authorea.com/users/775215/articles/874202-a-comparison-of-climate-driven-deep-learning-ensemble-and-swat-models-for-daily-streamflow-simulation-in-the-niger-river-basin-west-africa

A Comparison of Climate-Driven Deep Learning Ensemble and SWAT+ Models for Daily Streamflow Simulation in the Niger River Basin, West Africa

J. J. Idolor[1], A. Adediji[2], and P. A. Idowu[3]

[1]Institute of Ecology and Environmental Studies, Obafemi Awolowo University.

[2]Department of Geography, Obafemi Awolowo University.

[3]Department of Computer Science and Engineering, Obafemi Awolowo University.

Corresponding author: Jefia Idolor (jjidolor@gmail.com)

Key Points:

- A climate-driven deep learning ensemble (NQT-API-LSTM) is proposed and compared to SWAT+ model for watershed-scale streamflow prediction
- NQT-API-LSTM outperformed SWAT+ and reproduced the streamflow patterns of the Benue, Guinean, Sahelian and Sudan Flood in Niger River Basin
- The basin is heavily influenced by seasonal climate and regional groundwater dynamics from upstream catchments including the Sahara Desert

Abstract

Streamflow monitoring is very important for planning and management of water resources in watersheds, and their prediction accuracy is crucial for decision-making. The Niger River Basin is a transboundary resource, shared by nine West African Countries and Algeria and, a large portion of the population rely on the basin for rain-fed agriculture and hydropower. Over the years, the basin's streamflow regime has been altered due to climate change, drought, desertification and establishment of Dams. This research describes a novel Deep Learning framework comprised of Long Short-Term Memory (LSTM) requiring Antecedent Precipitation Index (API) and meteorological variables, preprocessed using Normal Quantile Transform (NQT) as input drivers and, compared with the Soil and Water Assessment Tool (SWAT+) for streamflow prediction. NQT-API-LSTM which considers catchment wetness and seasonality, was forced with reanalyzed climate (1979–2021) while, SWAT+ was driven with biophysical data and reanalyzed climate (2010–2020). The very high performance of both NQT-API-LSTM and SWAT+ models showed the models were reliable and can predict regulated flows with reasonable certainty. However, NQT-API-LSTM outperformed SWAT+ at Lokoja watershed and, realistically captured the influence of seasonal climate and regional groundwater dynamics from upstream catchments including the Sahara Desert on the Benue, Guinean, Sahelian and Sudan Flood. Overall, NQT-API-LSTM could be used successfully for watershed-scale streamflow prediction without the need for continuous ground support data, a benefit for sparsely gauged West African River Basins, while SWAT+ could be used as an alternative, particularly, to evaluate the watershed's response to land use/land cover changes.

1 Introduction

Streamflow is a major component of the hydrological processes in the hydrologic cycle, and it is required for assessment of the distribution, pattern, characteristics and behaviour of river networks in a watershed. At watershed scale, streamflow serves a crucial role in quantitative and qualitative monitoring and, control of water resources (Danandeh, 2018). Streamflow data from watersheds are required for the effective management of water resources (Ni et al., 2020), irrigation timing and scheduling (Vogel et al., 2015), hydraulic engineering design of infrastructures such as dams and reservoirs (Amirhossien et al., 2015; Awchi, 2014), river behaviour analysis (Fryirs & Brierley, 2013) and flood frequency analysis (Jimoh, 2007). Accuracy in estimation of the timing and volume of streamflow serves as decision-support tools for policy makers and water resources managers in developing effective water resources management schemes such as commissioning hydropower dams, timing and allocation of surface water for irrigation schemes, inland waterways transportation, construction of bridges and curvets, flood control and drought monitoring.

Hydrological models are representations of the physical, chemical and biological characteristics of the drainage basin catchments and, are used for simulation of basin behavior and the natural hydrological processes (Duan et al., 2019). Generally, hydrological models simulate the interactions between the input variables (such as climate data and terrain attributes) and the system (such as the drainage basin catchments) to estimate an output (such as streamflow, water level, percolation, soil moisture contents and evapotranspiration). Hydrological models are employed in estimating low flows which are necessary in watershed management, and forecasting peak flows which are necessary for flood mitigation (Pfannerstill et al., 2014). The main challenge in the implementation of hydrological models is the diverse parameters required for calibrating the model, in other to represent all hydrological processes in a drainage basin more accurately and reduce flood risk errors due to overestimating peak flows and, prevent water availability problems due to low flows underestimation (Jimeno-Sáez et al., 2018).

In recent times a number of hydrological models have been developed for simulating river discharges and associated hydrologic components, as well as assessing rainfall-runoff relationships and, the water balance in drainage basins (Makwana & Tiwari, 2017). Conceptual hydrological models utilize a number of mathematical formulations in describing the various processes of the hydrological cycle to simulate streamflow in a watershed (Noori & Kalin, 2016). The Soil and Water Assessment Tool (SWAT), is a sophisticated numerical model developed by Arnold et al. (1998) for simulation of the hydrological processes across several climatic and ecological regions. SWAT is a conceptual semi-distributed model that has gained increasing popularity within the last two decades for large scale regional hydrological simulation (Grusson et al., 2017). SWAT model have been employed in several studies to estimate the streamflow regimes in various watersheds by utilizing diverse spatial and temporal hydrometeorological and remote sensing data (Jimeno-Sáez et al., 2018).

SWAT model can assess and simulate streamflow including nutrients and sediments transport. SWAT model has been evaluated and validated in drainage basins within the United States of America and watersheds across the world for hydrologic modeling, pollutant loss, and climate change research (Arnold et al., 1998). SWAT model's major components are: hydrology; land use; plant growth; reservoirs; soil; and weather (Arnold et al., 1998). In recent times in the United States of America, SWAT model is increasingly being adopted for evaluation of the efficacy of the conservation policy of the United States Department of Agriculture (USDA) (Mausbach & Dedrick, 2004), for simulation of the Total Maximum Daily Load (TMDL) in catchments (Borah et al., 2006), for evaluation of hydrological processes at the Upper Mississippi River Basin, the entire United States of America, and a number of other hydrological purposes (Arnold et al., 1998). SWAT model has been successfully utilized for modeling the nitrate-nitrogen loadings and water quality of the raccoon river watershed (Jha et al., 2007). Adeogun et al. (2018) successfully simulated sediment transport and yield, identified and prioritized areas susceptible to erosion at the Upper Area of Lake Kainji at the Lower Niger River Basin in Nigeria and proposed better sediments management plan using SWAT model. Demirel et al. (2009) reported improvement in daily streamflow simulation accuracy of in data-scarce regional watersheds using SWAT model.

Deep Learning (DL) methods are increasingly becoming accepted as an alternative to the conventional distributed hydrological models, in simulating complex hydrological processes and predicting streamflow and water level more accurately. DL is capable of resolving large and complex tasks such as image classification, nonlinear simulations, time series forecasting, object detection and pattern analysis by discovering the nonlinear relationships between input data and the outputs (Hussain et al., 2020). Deep Learning architectures are composed of Artificial Neural Networks (ANN), and are data-driven approaches with the capability of simulating complex system dynamics. In recent times, ANNs have been used successfully in research for modeling complex systems, due to its inherent characteristics such as: being a nonlinear; self-adaptive data-driven approach; that consist of universal functional approximators; with the capacity to generalize (Haykin, 1999; Zhang et al., 1998).

In the last few years, a number of typical ANN architectures and variants that provide versatile nonlinear solutions for environmental and water resource challenges have been developed, such as: Convolutional Neural Networks (CNN); Deep Neural Networks (DNN); Recurrent Neural Networks (RNN); Gated Recurrent Unit (GRU); Long Short-Term Memory (LSTM) and; Transformers. These Machine Learning (ML) and DL architectures have been employed to simulate and forecast streamflow in watersheds. For instance, Hussain, et al. (2020) utilized extreme learning machine (ELM) and CNN for predicting the daily, weekly and monthly streamflow for a single step in Gilgit River Basin, Pakistan. According to Hussain, et al. (2020), the performance metrics indicated that ELM outperformed CNN model with an R^2 score of 0.99 for daily streamflow forecasting between 1980 and 2008. Jimeno-Sáez et al. (2018) applied ANN and SWAT models to estimate discharge in Miño-Sil and Segura watersheds located in Peninsular Spain with differing climatic conditions. It was reported that ANNs and SWAT showed good performance in modeling the daily streamflow of both watersheds. However, SWAT displayed better skills in predicting low flows, while ANNs showed better skills in simulating peak flows in the two drainage basins (Jimeno-Sáez et al., 2018).

Also, Fu et al. (2020) forecasted streamflow in Kelantan River catchment at the northeast region of Malaysia Peninsula by utilizing LSTM model. When compared with DNN models, LSTM models showed better performance in forecast accuracy irrespective of the characteristic steady dry season flow, or highly variable monsoon flow. According to Fu et al. (2020), the LSTM model showed expert skills in streamflow estimation in Kelantan River. Van et al. (2020), used CNN, LSTM and traditional ML models to forecast daily discharge at Can Tho and Chau Doc sub-catchments of the Vietnamese Mekong Delta. The CNNs and LSTMs models showed excellent performance in predicting daily rainfall–runoff. However, the CNN model showed better accuracy for streamflow simulations at both stations. It was reported that there was no significant contribution from rainfall because the LSTM and CNN models only considered lagged flows at gauge station. And concluded that CNN and LSTM models had better performance than conventional methods and, can be adopted as alternatives in other to increase the accuracy in simulation of hydrological parameters, especially in regulated upstream flows. ANN and SWAT were employed for forecasting daily streamflow in Pracana Basin and it was reported that ANN outperformed SWAT model in predicting high flows. SWAT hydrological model inefficiency in simulating high flows, despite having better mean squared error value, was attributed to the model formulation (Demirel et al., 2009). Furthermore, ANN, random forest (RF), Gaussian linear regression model (GLM), Gaussian generalised additive model (GAM), multivariate adaptive regression splines (MARSs) and 1D-CNN was used by Singh et al. (2023) for streamflow prediction in Sutlej River Basin and, concluded that RF outperformed other models in predicting streamflow. Ghorbani et al. (2016) employed support vector machine, multilayer perceptron (MLP) and radial basis function (RBF) for daily streamflow prediction whereas, Guo et al. (2011) used support vector machine and ANN for simulation of streamflow and concluded that support vector machine (SVM) showed better performance in predicting streamflow. Most recently, Xu et al. (2023) applied transfer learning (TL) Transformer (TL-Transformer), TL-LSTM, TL-MLP, Transformer, LSTM and MLP for flood modeling in data-sparse regions in the Yellow River, China.

Most physically-based hydrological models are computationally expensive and require large datasets of hydroclimatic and biophysical attributes for calibration and validation purposes (Jimeno-Sáez et al., 2018). Even the widely recognized SWAT hydrological model also requires large datasets comprising of land use, soil, terrain attributes, climate variables and management or decision variables in representing the watersheds hydrological conditions, for estimation of its CN (curve number) parameter. Hydrological modeling has remained a persistent challenge in the field of operational hydrology, due to the need to minimize the subjectivity of arbitrarily selected parameters to represent the physical conditions (Ali et al., 2010). However, these models are constrained by the limitation of required data for the model's parameter calibration, especially in sparsely gauged and un-gauged watersheds where data might be missing or inadequate.

The Antecedent Precipitation Index (API), is commonly employed to estimate runoff from storm events in watersheds where ground-support data is scarce or unavailable. It serves a crucial role in the estimation of the response of runoff to rainfall, particularly in catchments where runoff generation is heavily influenced by groundwater and adheres to the principles of the 'Variable Source Area' concept (Hewlett & Hibbert, 1967). Considerable research attention has been directed toward API (Descroix et al., 2002) which suggests subjectivity, in determining the API for representing physical conditions (Ali et al., 2010). Antecedent precipitation refers to the amount of prior rainfall, affecting the runoff yields of a specific storm event. API represents a measure of soil moisture index or catchment wetness and frequently remains a parameter determined subjectively and implemented arbitrarily in modeling runoff response to rainfall (Heggen, 2001). Recent studies have explored the use of API for simulating runoff yields and streamflow from storm events (Ali et al., 2010; Descroix et al., 2002; Ghosh et al, 2021). A number of recent studies have reported improved river discharge and stage forecasting by including API in the ANN model structure (Dawson & Abrahart, 2007). API, being a numerical value derived from rainfall depth, can be compared to or used as a proxy for soil moisture. It is a derived variable that can be incorporated into the modeling framework either as a conventional 'input driver' or as an expert 'output hint'. According to Xia et al. (1997), API can improve the effectiveness of nonlinear forecasting models, depending on their sophistication.

The last few years have seen an increased interest in process-based hydrological models for streamflow simulation in large West African River Basins (Aich et al., 2015; Poméon et al., 2018; Schuol et al., 2008). To the best of our knowledge, there are relatively few studies that has looked specifically at climate-driven deep learning approaches for modeling hydrological processes at watershed-scale. Thus, in this study, we proposed a novel NQT-API-LSTM ensemble to be used alongside SWAT+ model for daily streamflow simulation in the Niger River Basin, a large West African watershed with extremely heterogenous climatic conditions. A comparison of the performance of NQT-API-LSTM and SWAT+ model was made at the downstream gauging station at Lokoja. While, the efficiency of NQT-API-LSTM in simulating the Guinean, Sahelian and Sudan Flood events has been assessed at the Sahelian (Niamey) and Sudan (Jiderebode) sections of the basin.

2 Methodology

2.1 The Study Area

The study area is the Niger River Basin (NRB) spatially delimited to West Africa within the boundaries of Benin, Burkina Faso, Cameroon, Chad, Côte d'Ivoire, Guinea, Mali, Niger and Nigeria and, Algeria (North Africa) as shown in Figure 1. Geographically, it stretches between the meridians of 11°35'16.99" West and 15°51′44.74″ East, from Futa Jallon Highlands in Guinea to Chad; and between the parallels of Latitudes 4°21'19.60" to 23°54'20.41" North of the equator, from the Hoggar Mountains in Southern Algeria to the Gulf of Guinea. The headwaters of the

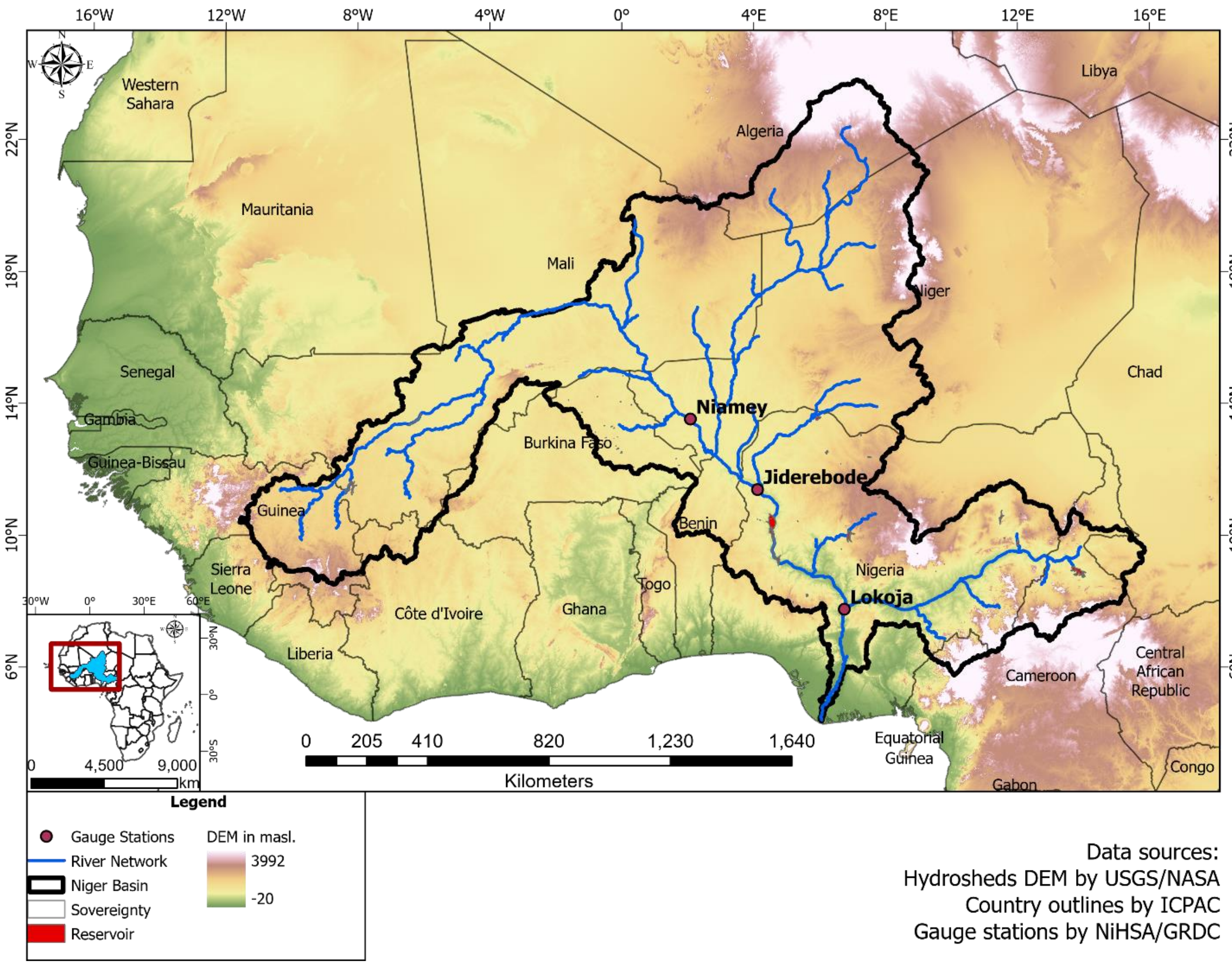

Figure 1. Niger River Basin along with Reservoirs, Discharge Stations, DEM and River Network (6th, 7th and 8th Order Channels)

Niger River System originates in the Futa Jallon Highlands in Guinea, and flows north-eastward, and during the monsoon forms an extensive floodplain in Mali known as the Inland Delta (Inland dú Niger). On leaving the delta, the river meanders in Mali, eventually flowing southeast through Niger, Benin Republic to Nigeria, and converges with the Benue River at Lokoja and, its waters, including its sediments and other associated loads such as exotic species are discharged into the Niger Delta by extensions into the Atlantic Ocean (Lienou *et al.*, 2010). There are 58 large dams and a total of 260 Dams and Reservoirs with a total volume of 4.2×10^{10} m^3, providing various water resources schemes (irrigation, water supply and hydroelectricity) within the river basin which have significantly altered the streamflow regime (Lienou *et al.*, 2010). The dams are irregularly distributed and, mostly concentrated in a few parts of NRB, like Burkina-Faso (where primarily small-sized dams are found) and Nigeria (where dams of all sizes, including large ones, exist). The capacity of existing dams ranges from 25×10^{-3} million m^3 at locations such as Camp de chasse in Tapoa, Niger to 1.6×10^{10} m^3 (Kainji, Nigeria). NRB has a total land area of 2,240,738.61 km^2 and, its stream channel subsystem consists of eight (8) orders ranked from 1st Order to 8th Order with the main channel, the Niger River ranked as the 8th Order and its largest tributary channel, the Benue River ranked as the 7th Order of NRB. The Niger River, which stretches approximately 4,200 km, is Africa's third longest river and, ranks as the second largest river in Africa in terms of discharge volume (Oguntunde *et al*., 2014; Okpara *et al*., 2013). NRB encompasses all major climatic regions of West Africa and, the regions are characterized based on their ecological zones and differing climatic characteristics. The five climatic regions are the Saharan, Sahelian, Sudan, Guinean, and Guineo-Congolian regions. The Saharan to mid Sahelian regions of the basin has the driest climatic regime, with average annual rainfall amounts less than 250 mm per year. While, the Guineo-Congolian region is the wettest with rainfall amounts between 2000 mm and 5000 mm. In terms of aridity, NRB encompasses all dryland climate subtypes (Hyper-arid, Arid, Semi-arid, Dry Subhumid zones) and non-dryland climate subtype (Humid zone). These climatic zones range from hyper-arid at the Saharan region in the Northern Niger Basin to humid at the Guineo-Congolian region in the Southern Niger Basin. The climate within NRB is influenced by the Intertropical Discontinuity (ITD) which by extension influences the hydrological processes of the river system (Thompson *et al*., 2017). NRB rainfall scheme is strongly seasonal and, determined by the Atlantic Monsoon oscillations from May to November. The magnitude of the Atlantic monsoon event varies greatly between the northern and southern NRB, but varies uniformly between the eastern and western parts of NRB (Lienou *et al*., 2010). The basin exhibits two distinct seasonal rainfall patterns: a unimodal wet season which occurs in northern NRB and; bimodal wet seasons in southern NRB with a short dry spell between the wet seasons. Three stream gauges were selected based on data availability and includes: Niamey gauging station located within the arid Sahelian region of the Middle NRB; Jiderebode gauging station located within the semi-arid Sudan region of the Lower NRB and; Lokoja gauging station located within the humid Guinean region of the Lower NRB.

The upstream area and major portions of the Upper Niger Basin is characterized by an ancient geologic landscape of metaigneous rocks followed by metasedimentary crystalline rocks. These impermeable rocks limits groundwater occurrence, with small aquifer systems occurring in areas where these rocks are either fractured or are weathered. At the Upper NRB groundwater do not contribute to the Niger River due to extremely low groundwater recharge from the headwaters (Fontes *et al*., 1991). The landscapes of the western bank of Niger River at the Middle NRB are characterized by the Liptako-Gourma Massif granitic basement (Descroix *et al*., 2012), while the sedimentary basin of Iullemeden lies on the right bank of the Niger River, from the Northern Segou through Gondo depression of the Eastern Dogon region and the Inland Delta (Andersen *et al*., 2005). The Iullemeden, is a multi-layered aquifer system consisting of the Continental Terminal dated Eocene to Pliocene overlain by Quaternary and recent dune-like Holocene ergs or alluvium deposits with aquifer’s groundwater hydrologically connected to the Niger River. The Continental Terminal, is an unbroken aquifer of about 100 m thickness covering tens of thousands of square kilometers, composed mainly of silty sandstones, clays and sand, with high-quality water. It is the most significant aquifer in the basin, and borders the Niger River System at Goundam, Timbuktu, and Gourma Rharous in Mali, Hoggar in Algeria and, extending through Bourem in Gao region in Mali to Niamey, and Gaya in Niger. Its northern stretch includes the Azaouâd, Taoudenni, Azaouâk and Tilemsi sedimentary basins. The Continental Shale Band aquifer lies underneate the Eocene to Cretaceous layers of the Continental Terminal formations and borders the Niger River at the Northern axis of Benin and also within the arid regions of Mali and Niger. At the Lower NRB in Nigeria, the watercourse flows along the Continental Terminal whose eastern axis borders the basin at Jos Plateau in Nigeria and, continues along the Quaternary alluvial deposits on both the right and left banks of the river at Jebba, and extending through the Benue valley to Cameroon and Chad. The river then flows alongside artesian aquifers and Cretaceous deposits that continue to Onitsha. At Onitsha, Tertiary marine layer, then spans across the Cretaceous layer, which are overlaid by saline Quaternary sediments from the coastal region of the Niger Delta. Outside this sedimentary basin, crystalline basement complex rock materials dated Precambrian, a constituent of the pan-African shield encloses NRB (Andersen et al., 2005; Persits et al., 1997).

The hydrological regime of NRB is heavily influenced by groundwater base flow, which is affected by annual rainfall and soil permeability. During the dry season, most of the contributions occur within the alluvial plains (Andersen *et al*., 2005). At Benin, the Iullemeden discharges into the main channel and its tributaries, and continues along the watercourse downstream. In Nigeria, the Rima and Sokoto rivers, ranked as 6th and 7th order respectively, which are the main rivers which drains the Iullemeden (IAEA, 2017), flows into the Niger River, just before Jiderebode gauging station. The Niger River Basin System is characterized by four major flood events that occurs at various sections of the basin, based on the climatic type which includes: the Benue; Guinean; Sahelian and; the Sudan Flood. The Benue Flood event is observed at Lokoja confluence, which is mainly associated with the flood waters from the Benue River, the largest tributary of the Niger River, whose source lies within the Adamawa Plateau in Cameroun, as well as the regulated upstream flows from major Dams which include Jebba, Lagdo, Kainji and Shiroro Dams. The Guinean Flood or “black flood” is the main flood from the headwaters of the Niger River in Guinea. The Sudan Flood or “white flood” is the local flood waters in Jiderebode sub-catchment, and the Sahelian Flood or “red flood” is the local flood waters in Niamey sub-catchment.

2.2 SWAT+ Model

This study simulated streamflow at Lokoja sub-catchment of NRB using the updated SWAT+ model version 2.2.0 (Bieger *et al*., 2017). The Soil and Water Assessment Tool (SWAT) is a semi-distributed, hydrological process-based river basin model, and can be calibrated to run on multiple temporal resolution (daily, monthly or yearly) depending on the time-scale of the observation (Arnold *et al*., 2012). The major constituents of SWAT+ model are: weather, hydrology, sedimentation, crop growth, pesticides, soil temperature and properties, nutrients, and agricultural management. SWAT model considers the watershed's heterogeneity by subdividing it into sub-basins derived from the land use/cover, drainage (river networks), soil properties and terrain attributes (such as reservoirs and slope). These sub-basins are subsequently partitioned into hydrologic response units (HRUs), representing distinct land areas characterized by distinctive combinations of landscape, soil, land use/cover and slope.

SWAT+ estimates the components of the water balance by considering the influence of climate forcing. The equation representing the water balance is expressed as:

$$SW_t = SW_o + \sum_{i=1}^{t} (R_i - Q_i - ET_i - Pe_i - QR_i) \tag{1}$$

Where SW_o and SW_t represents the initial and final soil water content (mm); the index t represents time (days); R_i, ET_i, Q_i, Pe_i and QR_i represents precipitation, evapotranspiration, surface runoff, percolation and baseflow (all units in mm) (Arnold *et al*., 1998).

2.2.1 Input Datasets for SWAT+ Model

Hourly ERA5 reanalysis climate data for the period of 1979 to 2020 in 0.25° × 0.25° grids of (approximately 25 km resolution), which includes temperature, precipitation, dew-point temperature, solar radiation, u-wind and v-wind components, were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) (Muñoz, 2019) and also retrieved from

Microsoft Planetary Computer (MPC) data catalog using STAC (Spatio-Temporal Access Catalog) API (Application Programming interface). The geospatial datasets used in this study include: HydroSHEDS 3 arc-second resolution (approximately 90 m) hydrologically conditioned DEM (Lehner et al., 2008; Lehner, 2022), 2 km resolution Harmonized World Soil Database (HWSD) soil data obtained from the Food and Agricultural Organization (FAO, 2012), ESA WorldCover 10 m resolution land use/land cover 2020 v100 dataset, made available by the European Space Agency (ESA) (Zanaga et al., 2021) and retrieved from MPC data catalog using the STAC API. Streamflow data was obtained from the Nigerian Hydrological Services Agency (NiHSA) and the Global Runoff Data Centre (GRDC, 2024). Reservoir data was provided by HydroSHEDS HydroLAKES database version 1.0 (Lehner et al. 2016).

2.2.2 SWAT+ Model Data Preprocessing

In view of computational cost and memory efficiency, HydroSHEDS Hydrologically conditioned DEM was resampled from 90 m (3 arc-seconds) resolution to 282 m resolution. While the 10 m resolution ESA Land Use was resampled to 30 m. ERA5 hourly meteorological reanalysis data was resampled to daily time series. Relative humidity was derived from air temperature and dew-point temperature according to Sonntag90 method (Sonntag, 1990) and Wind intensity was derived from the zonal (v-wind) and meridional (u-wind) wind components before resampling to daily timeseries. A total of 2795 climate data grid points from the delineated Niger River Basin were used as station data for climate input in SWAT+ model. Data preprocessing was carried out on node clusters (virtual machine compute instances) linked together on the backend by Microsoft Azure Kubernetes Services and the frontend by Dask in Python programming environment. In addition Microsoft Planetary Computer was also used for Cloud Native data assimilation, data preprocessing and geospastial data analysis using the STAC (Spatio-Temporal Access Catalog) API (Application Programming interface).

2.2.3 SWAT+ Model Setup

SWAT+ model parametrization was performed using the QSWAT+ interface in QGIS software. The DEM was used to derive the stream network and delineate the basin and its sub-basin boundaries. The Soil map was overlaid on the delineated watershed and sub-watersheds to provide details about soil properties, including soil texture, hydraulic conductivity, and available water content. Next, the land use/cover map was overlaid on the sub-basins and three slope categories were defined (0 – 3 %; > 3 % – 6 % and; > 6 %). Dominant HRUs option was used to derive the Hydrological Response Units (HRU) and reservoirs were added. The properties of the reservoirs included in the SWAT+ model structure in this study is presented in Table 1. Finally, NRB was subdivided into 11 sub-watersheds and 182 HRUs. The potential evapotranspiration was determined using the Penman-Monteith method while, the curve number was calculated using the Muskingum method. The dominant land use/cover distribution for NRB were: barren (30.36 %), range grasses (28.10 %); agriculture (18.45 %); range-brush (14.76 %); forest (6.83 %); wetland (0.72 %); urban (0.72 %); wetlands water (0.27 %) and; wetlands forested or mangrove forest (0.00%). The hyper arid, arid and semi-arid climatic condition in the Upper Sahel and Saharan regions explains the dominance of barren areas in NRB.

Table 1: Properties of the Reservoirs within the Niger River Basin which are included in the SWAT+ model

Name	River	Year	Long (°)	Lat (°)	Elevation (m)	Surface area (km^2)	Storage volume (km^3)
Kainji	Niger	1968	4.56	10.32	110	1034.85	15.00
Lagdo	Benue	1983	13.85	8.89	208	623.12	7.80
Shiroro	Kaduna	1984	6.90	9.98	335	271.12	7.00
Jebba	Niger	1984	4.68	9.36	89	274.76	3.60
Dadin Kowa	Gongola	1988	11.50	10.53	246	150.56	2.86
Selingue	Sankarani	1982	-8.22	11.46	336	335.77	2.17
Goronye	Rima	1983	5.95	13.54	286	107.48	0.98
Kiri	Gongola	1982	12.01	9.76	170	68.52	0.62
Markala	Niger	1947	-6.23	13.50	282	102.32	0.18

2.2.4 Sensitivity Analysis, Calibration and Validation

The hydrometeorological daily time series were split into three periods: warm-up; calibration and; validation. The period spanning from 2007 to 2009 was chosen as the warm-up phase and followed immediately by the calibration phase spanning from 2010 to 2007. While, the validation phase span from 2018 to 2020. Sobol method was used for the sensitivity analysis while, automatic calibration of the model's sensitive parameters was done using the Latin hypercube algorithm in SWAT+ Toolbox v1.0.1. Sensitivity analysis involved identifying the parameters with the strongest influence on streamflow, by varying the model's parameters, and estimating the model's output changes in relation to its variations (Arnold *et al*., 2012). During the sensitivity analysis, 2200 iterations was carried-out to obtain the 1st order sensitivity for the basin. In conducting calibration, daily streamflow observations at Lokoja gauge station was used and, involved adjustment of the model's parameters, in other for the daily simulations to correspond closely with observations. Automatic calibration was performed using two iterations of 1500 simulations with the sensitive parameters and, readjusting the parameters prior to the next simulation. The SWAT+ sensitive parameters for calibration and their final values considered in this study is shown in Table 2.

2.3 Antecedent Precipitation Index (API)

The antecedent precipitation index (API) is a hydrological model that accounts for amounts of previous rainfall occurrence prior to new storm events. It is a soil moisture index that is used in estimating runoff response to rainfall. API is derived from daily rainfall time series using the equation expressed as;

$$API = \sum_{t=-1}^{-i} P_t k^{-t} \qquad (2)$$

where P_t is the rainfall amount on the tth day prior to the occurrence of the rainfall event (storm),

Table 2. Parameters included in SWAT+ model calibration in Lokoja

Parameter	Sensitive Parameter	Rank	Range	Fitted Value
bd	Moist bulk density (g/cm^3)	1	0.9 – 2.5	-6.74 %
awc	Available water capacity (mm H_2O/mm soil)	2	0.01 – 1	-7.98 %
cn2	SCS runoff curve number	3	35 – 95	4.24 %
alpha	Baseflow alpha factor (1/day)	4	0 – 1	-6.12 %
perco	Percolation Coefficient	5	0 – 1	2.31 %
revap_co	Groundwater "revap" coefficient	6	0.02 – 0.2	12.36 %
epco	Plant uptake compensation factor	7	0 – 1	-6.44 %
esco	Soil evaporation compensation factor	8	0 – 1	2.25 %
chk	Hydraulic conductivity	9	-0.01 – 500	5.43 %
flo_min	Minimum aquifer storage to allow return flow (m)	10	0 – 50	2.58 %

and k is a constant (Kohler & Linsley, 1951).

API for this study was calculated using the daily rainfall (P) data for a period of 43 years (1979 – 2021) with a constant (k) of 0.98 that was within the limits for k (0.80 – 0.98) recommended by Viessman & Lewis (1996).

2.4 Data and Data Preprocessing Techniques

The choice of input variables has a significant influence on streamflow simulation and forecasting. According to systems theory, system variables can be categorized into three distinct categories: input; output and; state variables. State variables are representations of a measure of some intrinsic qualities of the system's condition and, these variables usually exhibit spatial variations and changes over time. In a Drainage Basin Systems, state variables include; Discharge (streamflow), Stage (water level) and water quality parameters. Input variables are further divided into two groups: control variables and random variables. Control variables are often referred to as management or decision variables such as irrigation extraction rates, and reservoir or dam storage and release schedule. While random variables exhibit statistical randomness such as temperature and precipitation. The Output variables represent the future state(s) being predicted such as streamflow. In most machine learning applications for streamflow forecasting, state variables frequently serve as both input and output, with input and output datasets representing the state(s) lag time and lead time respectively. Existing research has focused on a combination of previous state(s), as univariate or multivariate inputs, to predict future state(s) as output (Van et al., 2020; Xu et al., 2023), but has failed to explore the significant influence of meteorological variables on the simulated flow. Thereby limiting the Deep Learning model's ability to learn the causal relationship, behavior and pattern of the meteorological variables and their climatic drivers influencing the hydrological processes in the watershed. Thus, in this study, reanalysed meteorological variables such as rainfall and temperature and, API model outputs were utilized as input drivers in the LSTM model development.

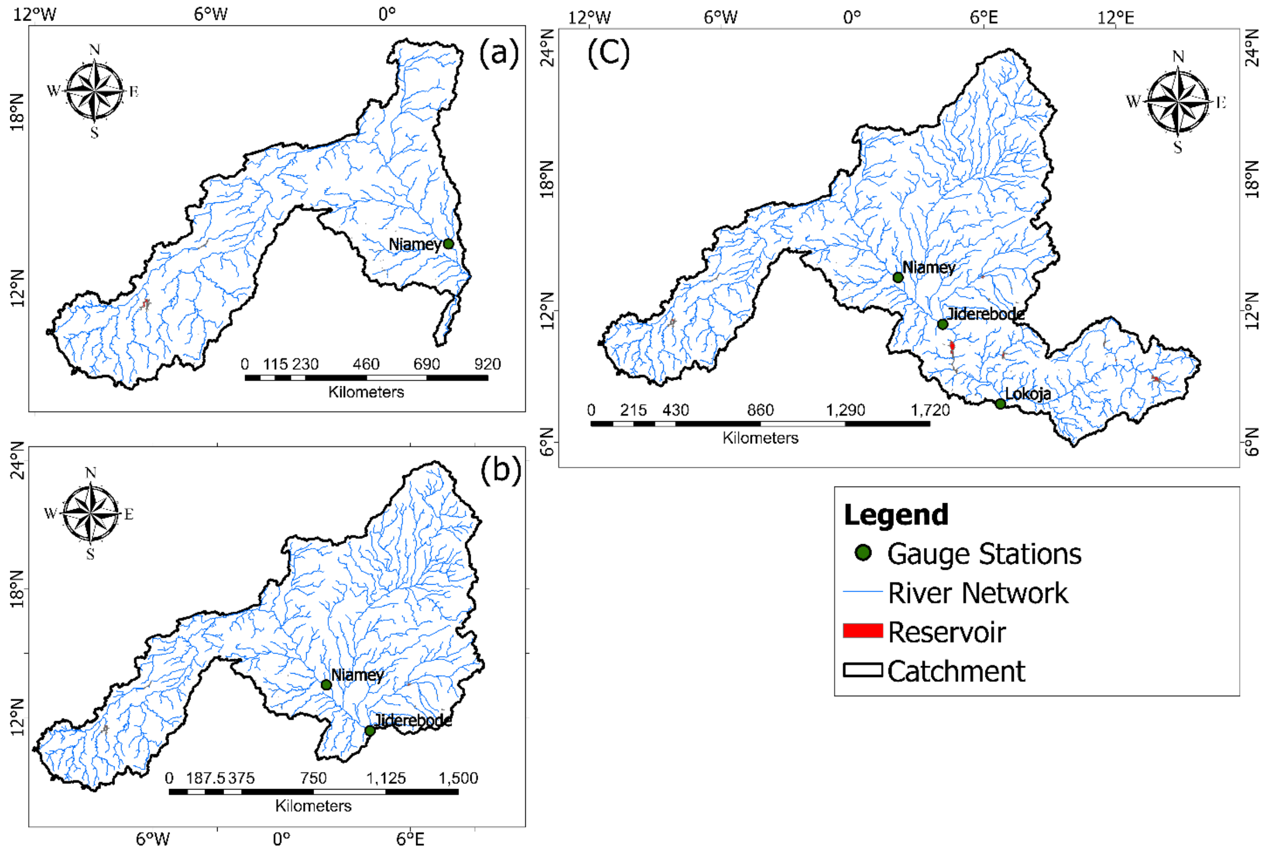

Figure 2. Catchments, reservoirs and drainage network within the Niger River Basin extending to: (a) Niamey gauge station; (b) Jiderebode gauge station and; (c) Lokoja gauge station

2.4.1 Feature Engineering

Hourly ERA5 reanalysis meteorological data spanning from 1979 to 2021 in 0.25° × 0.25° grids, was resampled to daily time series. The rainfall and temperature data required by the LSTM models was determined using the basin areal average. Since the gridded meteorological data was uniformly distributed, Arithmetic Areal Averaging method, which is based on equal contribution of all grid cells within the watershed was used to determine the basin areal average. In other for the model to accurately describe the spatial characteristics and pattern of the basin's climate, the DEM was delineated into sub-basins using ArcGIS Pro software. And, HyBAS HydroBasins level 4 shapefiles (Lehner and Grill, 2013) were merged within the boundaries of the delineated basins in other to encourage reproducibility. Due to the enormously heterogeneous geomorphology and climatic conditions of the study area (NRB), a single mesoscale modeling framework was developed for each sub-watershed as shown in Figure 2. The Arithmetic Areal Averaging method was applied to the active sub-basins rainfall and temperature data, to generate the input variables (features) for the LSTM models. Features (input variables) for the LSTM models to estimate daily streamflow (target) was generated from the precipitation and temperature time series using time delay embedding such as lag observations and rolling window statistics operation. Firstly, API was calculated for the selected sub-catchments. Other input variables considered in the feature space were: daily mean temperature (T_t); daily rainfall (P_t); lag rainfall (P_{t-n}); lagged temperature (T_{t-n}), rolling total rainfall (R_n) and; rolling mean temperature (T_m). LSTM model development is dependent on the spatial-temporal relationships between streamflow and climate dynamics. Similar to previous studies, the time delay embedding for the input variables (features) were determined using cross-correlation analysis to assess the temporal relationships between rainfall, temperature and streamflow (Amirhossien *et al*., 2015; Jimeno-Sáez *et al*., 2018).

2.4.2 Feature Selection

Feature space that comprises of large numbers of features (input variables) or highly correlated features, may lead to unacceptably high variance and reduction in prediction accuracy. Sparsity constraints can be applied on the feature space to prune uninformative covariates which do not influence the outputs. The Least Absolute Shrinkage and Selection Operator (LASSO) is a regression model introduced by Tibshirani (1996), which allows both continuous shrinkage and variable selection by utilizing an $L1$-norm sparsity constraint to enforce the coefficients of least important covariates to zero and retains only important features. The LASSO formulation is shown as follows:

$$\sum_{i=1}^{n}(y_i - \sum_{j} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{3}$$

where x_{ij} are the standardized features (input variables) and y_i are the response variables (targets) for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, p$, β_j represent the coefficient of the j-th feature. $\lambda \sum_{j=1}^{p} |\beta_j|$ is the the $L1$ penalty, also known as the Lasso penalty, and it is controlled by the hyperparameter λ, which adjusts the strength of the penalty term. In Lasso regression, the goal is to minimize the cost function by reducing the absolute values of the feature coefficients. As a feature's coefficient increases, so does the cost function value.

In other to find the best tuning parameter λ, according to procedures described by Tibshrani (1996). Features were standardized in other to be mean centered (mean = 0) with unit variance (standard deviation = 1), and split into training and test sets. Then 99 discrete λ-values that range from 0.1 to 9.9 with a step of 0.1 ($\lambda_1 = 0.1$, $\lambda_2 = 0.2$, $\lambda_3 = 0.3$, ……, $\lambda_{99} = 9.9$) and 5-fold cross-validation with a grid-search was applied to the training set. Which, randomly splits all of the training set data up into 5 sets (y_1, y_2, y_3, y_4 and y_5), then LASSO minimization was applied 5 times, each time fitted on 4 sets (y_1, y_2, y_3 and y_4) and tested on the hold-out set (y_5) chosen randomly, to obtain the regression coefficients (β_i) for a specific λ-value (for example $\lambda_1 = 0.1$). The resulting coefficients, estimates the residuals values of the hold-out set (y_i), and the MSE (mean square error) was computed for each hold-out set ($y_{i,i=1,\cdots,5}$), defined by:

$$MSE = mean((y_i - \hat{y}_i)^2) \tag{4}$$

where y_i represents the response variable and $\hat{y}_i$ is the residual. Then, the average of the MSE for each $y_{i,i=1,\cdots,5}$, was computed. The same procedure was repeated for the remaining 98 λ-values, generating a total number of 495 optimization iterations (99×5). After discovering the best performing tuning parameter, the absolute values of the LASSO coefficients for each predictor variables were obtained. Important features were selected ($\beta_i > 0$) while, non-influential features ($\beta_i = 0$) were dropped. Finally, the features selected for Jiderebode, Lokoja and Niamey watersheds are shown in Table 3.

Table 3: Feature (Input) Combinations for LSTM Models for Selected Sub-catchments of Niger River Basin

Watershed	Input Combinations	Outputs
Lokoja	API, P_{t-44}, R_{82}, T_{t-128}, T_{m201}	Q_t
Jiderebode	API, P_{t-62}, R_{194}, T_t, T_{t-130}, T_{m230}	Q_t
Niamey	API, P_{t-119}, R_{210}, T_t, T_{t-152}, T_{m9}	Q_t

2.4.3 Normal Quantile Transformations of Features

Hydrological and meteorological variables such as streamflow and precipitation are often asymmetric, because these variables have positive values and range from 0 to ∞. There is also the problem of seasonal variations, serial dependence and non-stationarity of the exogenous variables of the time series. It is crucial to transform the distribution of hydrometeorological variables and force them to follow a symmetric distribution, in other to satisfy the essential assumption of normality, which represents a fundamental concept applicable to most statistical and machine learning models (Moran, 1970; Goovaerts, 1997; Murphy, 2022). The Normal Quantile Transform (NQT) has found extensive application in various hydrological and meteorological contexts to perform nonlinear transformations of the Cumulative Distribution Function (CDF) into the CDF of the Standard Normal Distribution (Moran, 1970; Bogner *et al*., 2012), with its probability density function expressed as:

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}y^2} \tag{5}$$

In this study, the quantile-to-quantile normal score transformation was applied on each feature independently to map the p-quantile of each feature data distribution to the p-quantile of the standard normal distribution according to procedures described by Deutsch & Journel (1998); Pyrcz & Deutsch (2018). Expressed as:

$$y = F_Y^{-1}\big(F_Z(z)\big) \quad \forall\, z \tag{6}$$

Where, z represents the feature with CDF $F_z(z)$, y is the normal score value with CDF $F_Y(y)$ and F_Y^{-1} represents the Inverse CDF or quantile function of the output standard normal distribution. The steps includes: Firstly, the feature is split into training set, validation set and test set data; then the training set data is calibrated to generate the transformation parameters in other to prevent leakage of information; the CDF of each feature in the feature space is estimated and used to map the values of the observed variables to a uniform distribution; then, the associated inverse CDF or quantile function (F_Y^{-1}) is used to map the obtained values to the normal distribution and; extreme values or outliers of the validation set and test set data that fall below or above the fitted range of the training set data are mapped to the bounds of the output distribution. NQT is a robust data preprocessing technique that smooths out datasets with unusual distributions and is insensitive to outliers.

2.4.4 Features and Target Normalization

The streamflow (target) daily time series of the hydrometric gauge stations were checked for missing data and, the missing data were filled by linear interpolation expressed as:

$$\hat{y} = y_1 + (x - x_i)\frac{(y_{i+1} - y_i)}{(x_{i+1} - x_i)} \tag{7}$$

where x_i and y_i represent the first coordinates, while x_{i+1} and y_{i+1} denote the second coordinates. x represents the point at which interpolation is performed, and $\hat{y}$ corresponds to the interpolated value.

Most Deep Learning algorithms are not scale / shift invariant so it is important for the values of features and target to be within the same range. Transforming features and target to similar scale improves the performance of gradient descent, speeds up learning and, leads to faster convergence of the neural networks. Scaling is performed in other to prevent features with large values from dominating over small ones. Features (input variables) and targets (streamflow) were standardized in other to be mean centered with unit standard deviation. The procedures include: Firstly, the mean ($\bar{x}$) and standard deviation (σ) of the features and target of the training set were computed, as expressed by equations 8 – 9:

$$\bar{x} = \frac{1}{2}\sum_{i=1}^{N} x_i \tag{8}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{9}$$

Then the training set, validation set and test set of both the features and target were scaled as expressed by the equation:

$$x_i' = \frac{x_i - \bar{x}}{\sigma} \tag{10}$$

where, $\bar{x}$ and σ represents the mean and standard deviation of the variable x, x_i represents the features and targets that are transformed (scaled) into x_i'.

2.5 Long short-term memory (LSTM)

Since the introduction of RNNs by William and Zipser in the late 1980s, RNNs and their variants have received a lot of research attention in recent times. RNN can capture nonlinear short-term temporal dependencies, in RNN architecture input sequences are mapped to a sequence of hidden states, which maps to an output. The success of RNN is hindered by the problem of vanishing and exploding gradients (Hochreiter & Schmidhuber, 1997), which reduces the ability to capture non-stationary long-term temporal dependencies.

LSTM, displayed in Figure 3a, has the advantage over RNN to capture multiple non-stationary time dependencies and also long-term temporal dependencies due to the replacement of recurrent units in RNN with memory cells. LSTM cells consists of three gates: an input gate, a forget gate, and an output gate. These gates enable changes to be applied to a cell state vector and propagated iteratively in other to memorize and retrieve information over long time periods. The LSTM cell formulation are shown as follows:

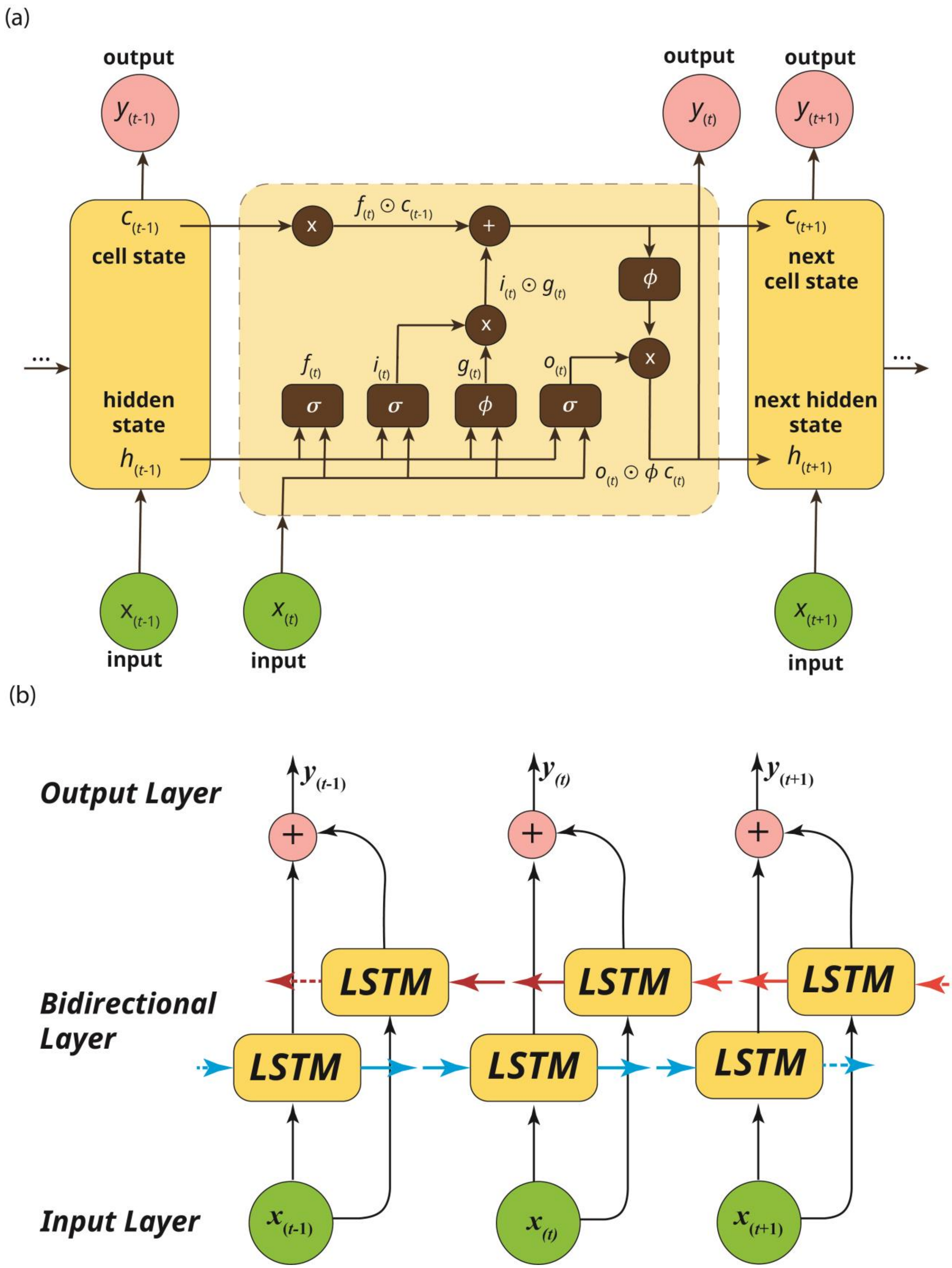

Figure 3. (a) Architecture of a LSTM Unit; (b) Bidirectional LSTM Layer

$$\begin{gathered} i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \\ f_i = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \\ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \\ g_t = \phi(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ c_t = f_t \odot c_t + i_t \odot g_t \\ h_t = o_t \odot \phi\,(c_t) \end{gathered} \tag{11}$$

where i is the input gate, f is the forget gate and o is the output gate, c is the cell activation, g is the input modulation gate, h is the hidden vector; the term W and b represents the gate matrix and bias; while ϕ represents a *tanh* function element-wise application and; $\odot$ is the Hadamard product.

2.5.1 Bidirectional LSTM

Bidirectional LSTM networks, are cutting-edge neural network architectures, which integrate LSTM gating mechanisms with optimized cell state representations, which are propagated in both forward and reverse directions. Bidirectional LSTM, displayed in Figure 3b, takes into account dependencies in both time directions by including expected correlations in future time-steps and, as a result of reverse state propagation, anticipated future correlations can influence the current outputs of the network. Unlike RNN and unidirectional LSTM networks, bidirectional LSTMs have the capability to detect, store, extract and resolve with greater precision multidimensional temporal dependencies. This study utilized the Bidirectional LSTM capabilities in assessing the correlation between prior streamflow observations and future forecasts within the history window in resolving the current streamflow outputs.

2.5.2 LSTM Model Training

The Deep Learning (DL) method used for streamflow forecasting was Bidirectional LSTM. The models were trained in Python environment with TensorFlow v2 Deep Learning framework using NVIDIA Tesla K80, M60 and T4 GPUs (Graphics Processing Units). Catchment seasonality was explicitly represented in the Bidirectional-LSTM model formulation by transforming the feature matrix to tensors in the feature space, to provide a 90-day history window with 1-day horizon. The LSTM model architecture and hyperparameters considered for this study are presented in Tables 4 – 5. Input features and targets were split into three sets of data: training set; validation set and; test set. Inputs for Lokoja catchment were split into training set (2010 – 2017): validation set (2018 – 2020): test set (2021) while, Niamey catchment was split into training set (80 %): validation set (15 %): test set (5 %), and Jiderebode catchment was split into training set (80 %): validation set (10 %): test data (10 %). The LSTM models were trained with the training datasets while the validation datasets were used to generate the validation cost function (error function) for updating the weights of the neural networks during backpropagation and the test datasets was used to test model performance.

Table 4: LSTM Model Architecture

Layer type	Input layer	Output layer	Parameters
LSTM Bidirectional layer 1	90	300	186000
LSTM Bidirectional layer 2	300	100	140400
Dense layer 1	100	20	2020
Dropout		20 %	0
Dense layer 2	20	1	21
Total Parameters			**328441**

Table 5: LSTM Model Hyperparameters

Hyperparameters	Value
Evaluation interval	150
Activation function	Tanh
Optimizer	Adam
Loss function	Mean Squared Error (MSE)
Epoch	100
Monitor	Validation loss
Regularization	Dropout, Early stopping
Patience	10

2.6 Hydrological Model Evaluation Criteria

The performance of the LSTM and SWAT models were evaluated using NSE (Nash-Sutcliffe Efficiency), KGE (Kling Gupta Efficiency), KGE′ (Adjusted Kling Gupta Efficiency), R^2 (Coefficient of Determination), RMSE (Root Mean Square Error) and PBIAS (percent of model bias).

KGE is expressed mathematically as:

$$KGE = 1 - \sqrt{(r-1)^2 + \left[\left(\frac{\sigma_s}{\sigma_o}\right) - 1\right]^2 + \left[\left(\frac{\mu_s}{\mu_o} - 1\right)\right]^2} \qquad (12)$$

where r represents the correlation coefficient between simulations and observations, σ represents the standard deviation, μ denotes the mean, and the indices s and o correspond to the simulations and observation values, respectively. KGE ranges from -∞ to 1 (Gupta *et al*., 2009).

The modified Kling Gupta Efficiency (KGE′), the second objective function for evaluation of the performance of the hydrological models is expressed mathematically as:

$$KGE' = 1 - \sqrt{(r-1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{13}$$

$$\beta = \frac{\mu_s}{\mu_o}$$

$$\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o}$$

where *KGE′* represents the modified KGE-statistic, *r* represents the correlation coefficient between simulations and observations; *b* and *γ* are the bias ratio and variability ratio respectively; *μ* represents the mean, CV_s represents the coefficient of variation, *σ* represent the standard deviation, and the indices *s* and *o* corresponds to the simulations and observations respectively. *KGE′* ranges from -∞ to 1 (Kling *et al*., 2012).

The Nash-Sutcliffe Efficiency (NSE) serves as the third objective function for assessing the performance of the hydrological models is expressed mathematically as:

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}(O_i - E_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}\right] \tag{14}$$

where O_i represents the *i*-th observation of the variable under evaluation, E_i corresponds to the *i*-th simulation of the same variable, $\bar{O}$ denotes the mean of the observed variables, and *n* signifies the total number of observations. NSE ranges from -∞ to 1 (Nash & Sutcliffe, 1970).

PBIAS, the fourth objective function for assessing the performance of the hydrological models is expressed mathematically as:

$$PBIAS = \frac{\sum_{i=1}^{n}(O_i - E_i) \times 100}{\sum_{i=1}^{n}(O_i)} \tag{15}$$

where O_i and E_i represents the *i*-th observation and simulation data respectively, and *n* represents the total number of observations. PBIAS ranges from -∞ to ∞ (Gupta *et al*., 1999).

R^2 (Coefficient of Determination), the fifth objective function for evaluation of the degree of collinearity between observations and simulated estimates is expressed mathematically as:

$$R^2 = \frac{[\sum_{i=1}^{n}(O_i - \bar{O}) \times (E_i - \bar{E})]^2}{[[\sum_{i=1}^{n}(O_i - \bar{O})^2]^{0.5} \times [\sum_{i=1}^{n}(E_i - \bar{E})^2]^{0.5}]^2} \tag{16}$$

where O_i and E_i are the *i*-th observation and simulation data respectively, $\bar{O}$ is the mean of the observation, $\bar{E}$ represent the mean of the simulation and *n* represents the total number of observations. R^2 ranges from 0 to 1.

RMSE (root-mean square error), the sixth objective function for evaluation of the closeness of the simulated to observed streamflow is expressed mathematically as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - E_i)^2}{n}} \tag{17}$$

where O_i and E_i are the i-th observation and simulation data respectively, and n represents the total number of observations. RMSE ranges from 0 (perfect fit) to ∞ (no fit) depending on the relative range of the simulated to observed data.

Indicators of the best performing hydrological models are KGE, KGE′, NSE and R^2 of 1 and, PBIAS and RMSE of 0. These coefficients are used for assessment of the goodness of fit of simulation and observed streamflow, and their performance rating is presented in Table 6.

Table 6: Hydrological Model Metrics for Daily Time Series

Performance Rating	NSE	KGE	KGE′	PBIAS (%)
Very good	NSE ≥ 0.7	KGE ≥ 0.7	KGE ≥ 0.7	| PBIAS | ≤ 25
Good	0.5 ≤ NSE < 0.7	0.5 ≤ KGE < 0.7	0.5 ≤ KGE < 0.7	25 < | PBIAS | ≤ 50
Satisfactory	0.3 ≤ NSE < 0.5	0.3 ≤ KGE < 0.5	0.3 ≤ KGE < 0.5	50 < | PBIAS | ≤ 70
Unsatisfactory	NSE < 0.3	KGE < 0.3	KGE < 0.3	| PBIAS | > 70

4 Results and Discussion

4.1 Results of Performances of NQT-API-LSTM and SWAT+ models

A comparison of the performances of SWAT+ and NQT-API-LSTM is shown in Table 7. SWAT+ model showed very high efficiency at Lokoja watershed, the NSE values for calibration was 0.71 and validation was 0.72, KGE values for calibration was 0.85 and validation was 0.81, KGE′ values for calibration was 0.85 and validation was 0.83, PBIAS values for calibration was 4.23 and validation was 8.74. The high values of the various efficiency criteria showed that SWAT+ can be classified as a very good model and, suitable for accurately simulating daily streamflow in large regional basins such as the Niger River Basin with regulated flows due to presence of hydrological modifications such as Dams and Reservoirs.

NQT-API-LSTM models showed very high efficiency at Jiderebode, Lokoja and Niamey watersheds. At Lokoja sub-catchment the values of NSE for calibration/training was 0.94, validation was 0.93 and testing was 0.86, KGE for calibration/training was 0.96, validation was 0.85 and testing was 0.87, KGE′ for calibration/training was 0.97, validation was 0.90 and testing was 0.90 and PBIAS for calibration/training was -1.75, validation was 10.01 and testing was -7.52. For Jiderebode watershed, the values of NSE for calibration/training was 0.81, validation was 0.91 and testing was 0.89, KGE for calibration/training was 0.85, validation was 0.93 and testing was 0.85, KGE′ for calibration/training was 0.83, validation was 0.94 and testing was 0.90 and PBIAS for calibration/training was -11.81, validation was 2.76 and testing was 7.12. While, Niamey sub-catchment, the values of NSE for calibration/training was 0.64, validation was 0.82 and testing was 0.83, KGE for calibration/training was 0.73, validation was 0.90 and testing was 0.83, KGE′ for calibration/training was 0.69, validation was 0.91 and testing was 0.87 and PBIAS for calibration/training was -23.56, validation was -2.12 and testing was 5.35.

Table 7: Streamflow Evaluation Metrics for LSTM and SWAT+ Model

Streamflow	Model	Subbasin	NSE	KGE'	KGE	PBIAS (%)	R^2	RMSE
Calibration/Training	SWAT+	Lokoja	0.71	0.85	0.85	4.23	0.71	3242.91
	Bi-LSTM	Lokoja	0.94	0.97	0.96	-1.75	0.94	1417.71
		Niamey	0.64	0.69	0.73	-23.56	0.64	374.10
		Jiderebode	0.81	0.83	0.85	-11.81	0.81	352.32
Validation	SWAT+	Lokoja	0.72	0.83	0.81	8.74	0.72	3762.55
	Bi-LSTM	Lokoja	0.93	0.90	0.85	10.01	0.93	1988.38
		Niamey	0.82	0.91	0.90	-2.12	0.82	261.04
		Jiderebode	0.91	0.94	0.93	2.76	0.91	280.44
Testing	Bi-LSTM	Lokoja	0.86	0.90	0.87	-7.52	0.86	1994.96
		Niamey	0.83	0.87	0.83	5.35	0.83	319.91
		Jiderebode	0.89	0.90	0.85	7.12	0.89	349.64

The high performance of NQT-API-LSTMs at all sub-catchments (Jiderebode, Lokoja, Niamey) also showed that NQT-API-LSTM can be classified as very good models for simulating and forecasting daily streamflow in large regional basins with heterogeneous climatic, topographic and geological conditions. Also, NQT-API-LSTM outperformed SWAT+ model and exhibited expert skills in predicting the influence of regulated upstream flows in downstream catchments of the basin, due to presence of dams and the hydrological settings of the Inland Delta.

The hydrographs of simulated and observed daily streamflow characteristics by SWAT model for the period 2010 to 2020 and NQT-API-LSTM models for the period 2010 to 2021 for Lokoja gauging station is shown in Figure 3 while, a subset of the simulated and observed daily streamflow characteristics for Jiderebode and Niamey gauging stations simulated by NQT-API-LSTM models for the period 2010 to 2019 is displayed in Figure 4. The daily simulated streamflow hydrograph at Lokoja station of NRB showed that SWAT model underestimated low flows, overestimated peak flows slightly, and observed peak flow lagged behind the simulated peak flow. NQT-API-LSTM was able to capture the underlying streamflow behaviour and pattern more accurately, and showed superior performance in estimating low flow but

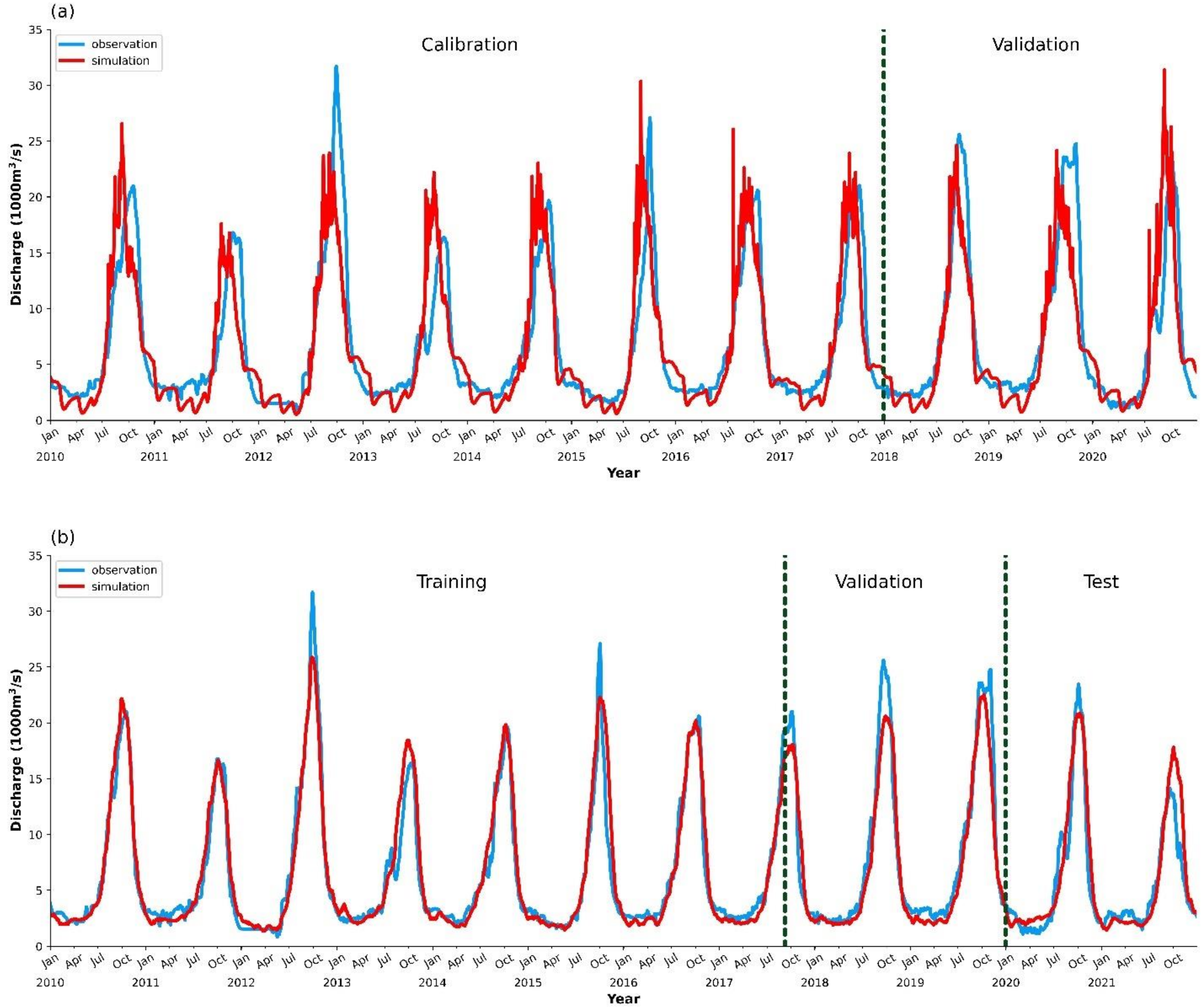

Figure 4: Hydrograph of Simulated Streamflow at Lokoja gauging station for the period 2010 to 2020: (a) SWAT model; (b) NQT-API-LSTM.

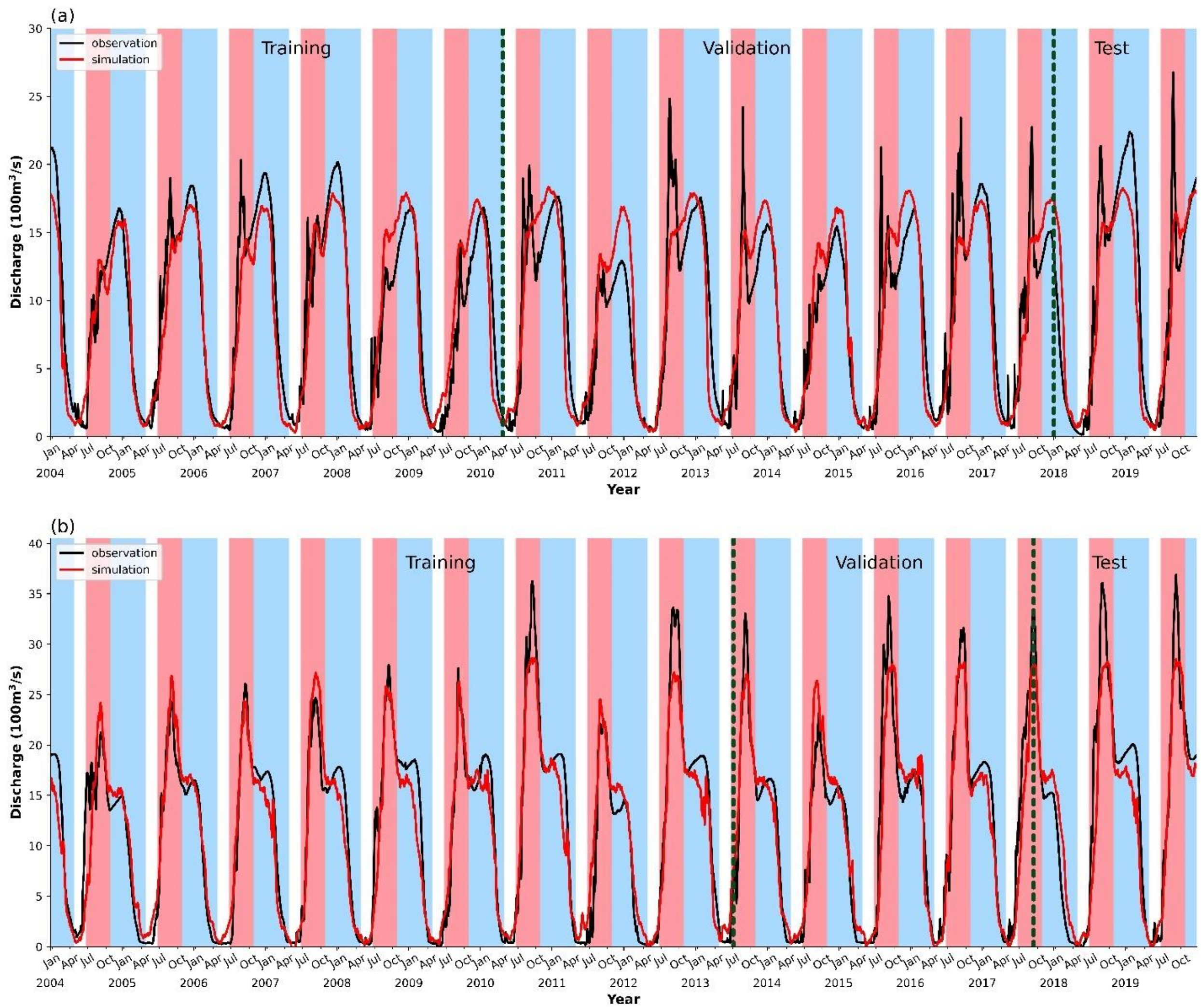

Figure 5: Hydrograph of NQT-API-LSTM Simulated Streamflow for the period 2004 to 2019: (a) Niamey station, Sahelian flooding depicted in red while Guinean flooding is indicated in blue; (b) Jiderebode station, Sudan flooding depicted in red, and Guinean flooding is indicated in blue.

underestimated peak flow slightly at Lokoja watershed. NQT-API-LSTM ensemble showed very high performance in simulating the streamflow pattern at Jiderebode sub-catchment, and showed very good skills in estimating low flow but underestimated peak flow. The NQT-API-LSTM model displayed expert skills in simulating the two flood events at Jiderebode station which include: the Guinean Flood (black flood) and; the Sudan Flood (white flood). The NQT-API-LSTM underestimated the white flood peak but showed better performance in simulating the black flood peak at Jiderebode sub-catchment. NQT-API-LSTM showed very good skills in simulating the streamflow pattern of the black flood and red flood events at Niamey sub-catchment, and also showed very good performance in estimating low flow. The NQT-API-LSTM underestimated the Sahelian peak (red flood) but showed better performance in simulating the Guinean peak (black flood) at Niamey sub-catchment.

4.2 Discussion

Generally, SWAT+ and NQT-API-LSTM models accurately reproduced the streamflow, with the SWAT+ model slightly overestimating peak flows. It has been reported that SWAT+ model's peak-flow inefficiency may be attributed to model formulation (Jimeno-Sáez et al, 2018). However, NQT-API-LSTM ensemble estimations were more accurate and closely matched the observed streamflow, which was reflected in the lower RMSE in calibration/training, validation and test phases. The varying performance of NQT-API-LSTMs across the various gauging stations may be attributed to differences in surface and extremely heterogenous climatic conditions from the hyper-arid region of Sahara Desert at the Northern NRB to the humid region of the Guinean Niger Basin at Lokoja. The observed and simulated discharge (> 2000 m^3/s) during the dry season, provide additional evidence of a link between groundwater return flow and streamflow, and would suggest that "Variable Source Areas Concept" is applicable at Lokoja catchment, hence the success of NQT-API-LSTM ensemble. In addition, there was also influence of regulated flows from more than 260 Dams and Reservoirs within the upstream sections of the basin. However, due to data limitations and model complexity, management or decision variables (such as reservoir storage, release schedule, volume, principal and emergency spillway as well as irrigation extraction rates) were not included in NQT-API-LSTM model formulation.

A cursory glance at Figure 3 reveals a very good fit, and identical pattern of both simulated and observed streamflow in all phases of the hydrograph at Lokoja sub-catchment, which underscores the efficacy of NQT-API-LSTMs in learning the causal relationships of the climatic drivers influencing streamflow in the watershed. An additional reason for the better performance of NQT-API-LSTM model at Lokoja watershed may be attributed to the surface attributes and humid climatic condition. Discharge of rivers in humid regions are less sensitive to climate input due to higher runoff and lower infiltration rate. The results showed that NQT-API-LSTM can help reduce streamflow overestimation which was inherent in SWAT+ model, although it was slightly underestimated by NQT-API-LSTM. Machine learning models tends to predict values closer to the mean of the distribution better than values at the extremes (high and lows). A possible reason for this discrepancy might be that peaks with high values are scarce, when compared with values of average peaks in the training data sets, and the LSTM model assigns relatively more importance to the average values rather than the high values extremes. These findings suggest a need for extreme caution in applying NQT-API-LSTMs for extrapolating beyond the bounds of the historically observed training data range, especially in anomaly detection and studies on extreme flood events. Results obtained by Jimeno-Sáez et al. (2018); Minns & Hall (1996) are consistent with our findings. The rolling total rainfall (R_n) and API, a proxy for catchment wetness, accounts for infiltration and groundwater dynamics at watershed-scale. Considering the remarkably strong regional heterogeneity in rainfall distribution, temperature dynamics, aridity and surface properties across the structure of the basin. NQT-API-LSTM ensemble outshined with its simplified ability to approximate the streamflow at the arid (Niamey), semi-arid (Jiderebode) and humid (Lokoja) environments.

Though, input data acquisition (such as meteorological variables) and, preprocessing is comparatively easier for NQT-API-LSTMs, the limitations of both NQT-API-LSTMs and SWAT+ for basin-scale modeling framework are increasingly apparent. The study basin is very large and sparsely gauged, which significantly increases the difficulty in acquiring the necessary meteorological and topography data to parameterize a physically-based model. In addition, the setup and calibration for SWAT+ model is computationally more expensive and takes considerably more time than NQT-API-LSTM. Conversely, obtaining a consistent timeseries of in-situ meteorological and hydrological data to calibrate a data-driven model is very challenging in West Africa. Since, the study area is a large regional basin, watershed delineation, data preprocessing and spatial averaging of meteorological variables incurs additional computational cost. NQT-API-LSTMs function as black-box models, and they do not provide information on the water balance and its constituent components. However, relying solely on precipitation and temperature as input variables for the models, represents a constraint in the case of NQT-API-LSTM ensemble because the interaction between rainfall and runoff is influenced by various biophysical parameters. Which confers an advantage to SWAT+ when exploring a number of scenarios concerning the basin's response to land use and land management. The general picture emerging from the results of this study is that both SWAT+ and NQT-API-LSTM models are suitable for simulating streamflow in large basins. However, it is recommended to use NQT-API-LSTMs for studies on extreme hydrologic events (such as floods), hydrological management (low-flow events) and developing scenarios for climate change impact on the hydrological processes. While, SWAT+ model is advisable for assessing the hydrological response of the watershed to land use/land cover (LULC) changes.

NQT-API-LSTM was able to capture the Guinean and Sahelian flood events at Niamey station however, the Guinean Flood was more accurately reproduced than the Sahelian Flood. These findings are less surprising if we consider the strong influence of the headwaters of the Niger River from the humid Guineo-Congolian region in Guinea, associated with higher runoff and lower infiltration rate on one hand and, the hydrological settings within the vast wetlands of the Inland Delta, leading to delay in arrival of the black flood at Niamey. This can be explained by the 90-day history window, which enabled NQT-API-LSTM to effectively understand the seasonal correlation between historical and future streamflow patterns while, considering the 3-month streamflow delay due to groundwater recharge in the aquifer system of the Inland Delta. The results confirm the findings of Aich et al. (2014), which posits that the hydrological conditions in the upstream region of the Sahelian NRB significantly influence the Guinean Flood. However, the Sahelian flooding has a shorter duration with inconsistent Peak flow that is not easily identified in some years thus, making it difficult for NQT-API-LSTM to learn its distribution and, therefore assign relatively more importance to the Guinean flood. An additional reason for lower accuracy in predicting the Sahelian flooding may be attributed to increased complexity in modeling the hydrological responses of arid and semi-arid regions, which are more sensitive to climate inputs, due to high infiltration and evapotranspiration rates and also limited and / or irregular precipitation. There are also uncertainties in the meteorological reanalysis whose deficiency could most easily reflect on the model's performance in arid and semi-arid regions.

NQT-API-LSTM was also able to capture both the Guinean and Sudan Flood events at Jiderebode station, but reproduced the Sudan Flood (white flood) more accurately than the Guinean Flood (black flood). A possible reason for this discrepancy might be that the delayed arrival of the black flood at Jiderebode in the dry season due to hydrometeorological dynamics at the Inland Delta and the upstream Sahelian basin constituted additional difficulty for the model. It might seem counterintuitive that NQT-API-LSTM ensemble reproduced the white flood more accurately than the black flood at Jiderebode station, but considering the fact that the white flood occurs during the wet season due to surface runoff and groundwater dynamics. NQT-API-LSTM was able to capture the influence of seasonal climate, catchment seasonality and wetness on monsoon streamflow. In Seasonal climates, increased rainfall during the wet season, followed by recharge and increased groundwater storage, results to elevated increase in the regional water table, which is seasonally dynamic and may vary in relation to rainfall variation and its climatic drivers (Davie & Quinn, 2019). An interesting side finding was that precipitation within the Saharan region was indirectly influencing the downstream flow at the Sudan region through the Continental Terminal of the Iullemeden. Since, the data preprocessing was applied to the entire upstream catchment from the Saharan to the Sudan regions of the basin. NQT-API-LSTM was able to learn the synergistic contributions of all climatic regions to streamflow. Overall, results of this study provide support for the validity of NQT-API-LSTM approach for simulating and forecasting streamflow in large watersheds due to its simplified model formulation requiring only meteorological variables and minimal computational resources, with the possibilities of exploring other hydrological processes including water quality and water levels, as reported in some studies (Chen et al., 2022; Cho et al., 2023; Pyo et al., 2023; Vizi et al., 2023).

5 Conclusions

In this study, we proposed a novel framework; a climate data-driven NQT-API-LSTM ensemble, and compared the performance with SWAT+, a quasi-physically based model, for daily streamflow forecasting. To validate their competency, they were applied in NRB, the largest transboundary river basin in West Africa, consisting of hyper-arid, arid, semi-arid, dry-subhumid and humid climatic conditions. The combination of API and LSTM for multivariate time series forecasting leverage on the synergy of API and deep learning techniques in surface water modeling. This approach exploits LSTMs sophisticated capability to capture complex temporal and seasonal dependencies while taking into consideration the inherent strengths of API in estimating catchment wetness, particularly in NRB where streamflow is strongly influenced by soil water or groundwater. The rolling total rainfall (R_n) also accounted for catchment wetness while, LASSO was used for selection of input variables which was transformed to a Gaussian distribution using NQT. The SWAT+ model was calibrated with daily streamflow observations using the Latin hypercube algorithm. The results indicated that NQT-API-LSTM ensemble showed better performance in simulating streamflow at Lokoja watershed and was able to reproduce the influence of rainfall and temperature variations and its climatic drivers adequately. While LSTM approach was superior to SWAT+ methods as shown in this study, SWAT+ can be used as an alternative hydrological model, especially to assess the basin’s response to land use/land cover changes. In light of the very good performance of NQT-API-LSTM, few conclusions can be drawn from the results of this study: Streamflow at the Middle and Lower Niger River Basin is heavily influenced by climate and regional groundwater dynamics at the upstream sections of the basin; the Saharan section of the basin is hydrologically active and its rainfall and temperature variations influences the seasonal dynamics of the regional groundwater table; the Black Flood was more accurately reproduced than the Red Flood at Niamey; the White Flood was simulated with greater precision than the Black Flood at Jiderebode; the model was able to predict regulated flows accurately in downstream catchments and; NQT-API-LSTM is suitable for studies on extreme events (such as floods), hydrological management (low flow events e.g. hydropower generation) and climate change impact on hydrological processes. The major advantages of the NQT-API-LSTMs are its ability to learn the basin’s response to climate change and variability remotely, without the need for spatially-explicit biophysical characteristics of the watershed. In this study, only precipitation and temperature inputs were considered. Therefore, the current study could be improved by including additional input variables that influences streamflow and, exploring the neural search space to discover more sophisticated deep learning architectures and hyperparameters.

Acknowledgments

This work was supported by the Microsoft AI4Earth Grant by providing Microsoft Azure Credits for access to computational resources hosted on Microsoft Azure Cloud Computing Services, and access to Microsoft Planetary Computer. Microsoft AI4Earth Team also supported with training, mentorship and provision of curated codes from open sources for the Cloud Native Geospatial Data Analysis and Deep Learning workflow. The first author was supported by the Nigerian Hydrological Services Agency (NiHSA) and the Upper Niger River Basin Development Authority (UNRBDA). The first author is grateful to ESRI for the ArcGIS Pro Professional Advanced User software license grant.

Data Availability Statement

The observed streamflow data for Jiderebode and Lokoja stations are confidential, and the authors do not have permission to share the data. ArcGIS Pro software would require the purchase of a license (https://www.esri.com/en-us/arcgis/products/arcgis-pro/buy#for-business). All other data analyzed in this study are from open sources and publicly available. River discharge data for Niamey station (GRDC, 2024), HydroSHEDS hydrologically conditioned DEM (Lehner et al., 2008; Lehner, 2022) is available at (https://data.hydrosheds.org/file/hydrosheds-v1-con/af_con_3s.zip), HWSD soil data (FAO, 2012), ESA WorldCover 10 m 2020 v100 (Zanaga et al., 2021) is available at Microsoft Planetary Computer (Microsoft Open Source et al., 2022), ERA5 reanalysis climate data (Muñoz, 2019) is also available at Microsoft Planetary Computer (Microsoft Open Source et al., 2022), HyBAS HydroBasins (Lehner and Grill, 2013) is available at (https://data.hydrosheds.org/file/hydrobasins/standard/hybas_af_lev01-12_v1c.zip), HydroSHEDS HydroLAKES (https://www.hydrosheds.org/products/hydrolakes/), SWAT+ model (https://swat.tamu.edu/software/plus/), SWAT+ Toolbox (https://swat.tamu.edu/software/plus/) and QGIS (https://www.qgis.org/en/site/forusers/download.html).

References

Adeogun, A. G., Ibitoye, B. A., Salami, A. W., & Ihagh, G. T. (2018). Sustainable management of erosion prone areas of upper watershed of Kainji hydropower dam, Nigeria. *Journal of King*

Saud University – Engineering Sciences, *32*(1): 5–10. https://doi.org/10.1016/j.jksues.2018.05.001

Aich, V., Liersch, S., Vetter, T., Andersson, J. C. M., Müller, E. N. & Hattermann, F. F. (2015). Climate or Land Use?—Attribution of Changes in River Flooding in the Sahel Zone. *Water, 7*, 2796–2820, https://doi.org/10.3390/w7062796

Ali, S., Ghosh, N. C., & Singh, R. (2010). Rainfall–runoff simulation using a normalized antecedent precipitation index. Hydrological Sciences Journal – *Journal des Sciences Hydrologiques, 55*(2), 266–274, http://dx.doi.org/10.1080/02626660903546175

Amirhossien, F., Alireza, F. Kazem, J., & Mohammadbagher, S. (2015). A Comparison of ANN and HSPF Models for Runoff Simulation in Balkhichai River Watershed, Iran. *American Journal of Climate Change*, *4*, 203–216, https://dx.doi.org/10.4236/ajcc.2015.43016

Andersen, I., Dione, O., Jarosewich-holder, M., Olivry, J.-C., & Golitzen, K. G. (2005). The Niger River Basin: A Vision for Sustainable Management. *The International Bank for Reconstruction and Development / The World Bank*. https://doi.org/10.1596/978-0-8213-6203-7

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: model development, *JAWRA Journal of the American Water Resources Association*, *34*(1):73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinvasan, R., et al. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE, 55*, 1491–1508. https://doi.org/10.13031/2013.42256) @2012

Awchi, T. A. (2014). River discharges forecasting in Northern Iraq using different ANN techniques. *Water Resources Management, 28*(3), 801–814. https://doi.org/10.1007/s11269-014-0516-3

Bieger, K., Arnold, J. G., Rathjens, H., White, M. J., Bosch, D. D., Allen, P. M., et al. (2017). Introduction to SWAT+, Introduction to SWAT+, A Completely Restructured Version of the Soil and Water Assessment Tool. *Journal of the American Water Resources Association (JAWRA), 53*(1), 115–130. https://doi.org/10.1111/1752-1688.12482

Bogner, K., Pappenberger, F., & Cloke, H. L. (2012). Technical Note: The normal quantile transformation and its application in a flood forecasting system. *Hydrology and Earth System Sciences, 16*, 1085–1094. https://doi.org/10.5194/hess-16-1085-2012.

Borah, D. K., Yagow, G., Saleh, A., Barnes, P. L., Rosenthal, W., Krug, E. C., & Hauck, L. M. (2006). Sediment and nutrient modeling for TMDL development and implementation. *Transactions of the ASABE*, *49*(4): 967–986. https://doi.org/10.13031/2013.21742) @2006

Chen, H., Yang, J., Fu, X., Zheng, Q., Song, X., Fu, Z., et al. (2022). Water Quality Prediction Based on LSTM and Attention Mechanism: A Case Study of the Burnett River, Australia. *Sustainability*, *14*, 13231. https://doi.org/10.3390/su142013231.

Cho, M., Kim, C., Jung, K., & Jung, H. (2023). Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction. *Water*, *14*, 2221. https://doi.org/10.3390/w14142221.

d'Herbès, J. M., & Valentin, C. (1997). Land surface conditions of the Niamey region: ecological and hydrological implications. *Journal of Hydrology*, *188–189*, 18–42. https://doi.org/10.1016/S0022-1694(96)03153-8

Danandeh M. A. (2018). An improved gene expression programming model for streamflow forecasting in intermittent streams. *Journal of Hydrology*, *563*, 669–678. https://doi.org/10.1016/j.jhydrol.2018.06.049

Duan, Q., Pappenberger, F., Wood, A., Cloke, H. L., & Schaake, J. C. (2019). Handbook of Hydrometeorological Ensemble Forecasting. *Springer, Berlin, Heidelberg*. https://doi.org/10.1007/978-3-642-40457-3

Davie, T., & Quinn, N. W. (2019). Fundamentals of hydrology, 3rd Eds, *Routledge*, ISBN 9780415858694. https://lccn.loc.gov/2018057141

Dawson, C. W., & Abrahart, R. J. (2007). Evaluation of two different methods for the antecedent precipitation index in neural network river stage forecasting. *Geophysical Research Abstracts*, 9, 07522.

Demirel, M. C., Venancio, A., & Kahya, E. (2009). Flow forecast by SWAT model and ANN in Pracana Basin, Portugal. *Advances in Engineering Software*, *40*, 467–473. https://doi.org/10.1016/j.advengsoft.2008.08.002

Descroix, L., Genthon, P., Amogu, O., Rajot, J.-L., Sighomnou, D., & Vauclin, M. (2012). Change in Sahelian rivers hydrograph: thecase of recent red floods of the Niger River in the Niamey region. *Global and Planetary Change*, *98–99*, 18–30. http://dx.doi.org/10.1016/j.gloplacha.2012.07.009

Descroix, L., Nouvelot, J. F., & Vauclin, M. (2002). Evaluation of an antecedent index to model runoff yield in the western Sierra Madre (north-west Mexico). *Journal of Hydrology*, *263*, 114–130. https://doi.org/10.1016/S0022-1694(02)00047-1

Deutsch, C. V., & Journel, A. G. (1998). GSLIB: Geostatistical Software Library and User Guide, 2nd Ed., *Oxford University Press*, USA.

FAO/IIASA/ISRIC/ISS-CAS/JRC (2012). Harmonized World Soil Database (Version 1.2); *FAO*: Rome, Italy; IIASA: Laxenburg, Austria, 2012; Available online:

https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/ (accessed on 15 March 2023).

Fontes, J.-C., Andrews, J. N., Edmunds, W. M., Guerre, A., & Travi, Y. (1991). Paleorecharge by the Niger River (Mali) Deduced from groundwater geochemistry. *Water Resources Research, 27*(2), 199–214. https://doi.org/10.1029/90wr01703

Fryirs, K. A., & Brierley, G. J. (2013). River Behaviour. In Geomorphic Analysis of River Systems (eds K. A. Fryirs and G. J. Brierley). Wiley-Blackwell. https://doi.org/10.1002/9781118305454.ch11

Fu, M., Fan, T., Ding, Z., Salih, S. Q., Al-Ansari, N., & Yaseen, Z. M. (2020). Deep Learning Data-Intelligence Model Based on Adjusted Forecasting Window Scale: Application in Daily Streamflow Simulation. *IEEEAccess*, *8*: 32632–32651. https://doi.org/10.1109/ACCESS.2020.2974406

Ghorbani, M. A., Khatibi, R., Goel, A., FazeliFard, M. H., & Azani, A. (2016). Modeling River Discharge Time Series using Support Vector Machine and Artificial Neural Networks. *Environmental Earth Sciences*, *75*(8) 685. https://doi.org/10.1007/s12665-016-5435-6

Ghosh, N. C., Jaiswal, R. K., & Ali, S. (2021). Normalized Antecedent Precipitation Index Based Model for Prediction of Runoff from Un-Gauged Catchments. *Water Resources Management*, *35*, 1211–1230. https://doi.org/10.1007/s11269-021-02775-w

Goovaerts, P. (1997). Geostatistics for Natural Resources Evaluation, *Applied Geostatistics Series*, *Oxford University Press*, USA. https://doi.org/10.1093/oso/9780195115383.001.0001

GRDC. (2024). The Global Runoff Data Centre, 56068 Koblenz, Germany. Available online: https://portal.grdc.bafg.de/applications/public.html?publicuser=PublicUser#dataDownload/Home/ (accessed on 8 April 2024).

Grusson, Y., Anctil, F., Sauvage, S., & Sánchez Pérez, J. M. (2017). Testing the SWAT Model with Gridded Weather Data of Different Spatial Resolutions. *Water*, *9*(1), 54. https://doi.org/10.3390/w9010054

Guo, J., Zhou, J., Qin, H., Zou, Q., & Li, Q. (2011). Monthly Streamflow Forecasting Based on Improved Support Vector Machine Model. *Expert Systems with Applications*, *38*(10): 13073–13081. https://doi.org/10.1016/j.eswa.2011.04.114

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*, *4*, 135–143. https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)

Haykin S. (1999). Neural Networks. A Comprehensive Foundation, 2nd Edition. *Prentice Hall PTR*, Englewood Cliffs, New Jersey, USA, ISBN 0132733501. pp 696.

Heggen, R. J. (2001). Normalized antecedent precipitation index. *Journal of Hydrologic Engineering ASCE*, *6*(5), 377–381. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:5(377)

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hewlett, J. D. and Hibbert, A. R. (1967). Factors affecting the response of small watershed to precipitation in humid areas. In: W. E. Soper and H. W. Lull(eds.), *International Symposium on Forest Hydrology*, Pengamonb, Oxford, 275-290.

Hussain, D., Hussain, T., Khan, A. A., Naqvi, S. A. A., & Jamil, A. (2020). A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin. *Earth Science Informatics*, *13*:915–927. https://doi.org/10.1007/s12145-020-00477-2

IAEA (2017). Iullemeden Aquifer System: Report of the IAEA-Supported Regional Technical Cooperation Project RAF/7/011. Integrated and Sustainable Management of Shared Aquifer Systems and Basins of the Sahel Region. *IAEA*, Vienna, Austria, 2017.

Jha, M. K., Gassman, P. W., & Arnold, J. G. (2007). Water Quality Modeling for the Raccoon River Watershed using SWAT. *Trans. ASABE*, 50(2), 479–493. https://doi.org/10.13031/2013.22660

Jimeno-Sáez, P., Senent-Aparicio, J., Pérez-Sánchez, J., & Pulido-Velazquez, D. (2018). A Comparison of SWAT and ANN Models for Daily Runoff Simulation in Different Climatic Zones of Peninsular Spain. *Water*, *10*(192), 1–19, https://dx.doi.org/10.3390/w10020192

Jimoh, O. D. (2007). Impacts of Dams on the Hydrology of River Niger at Lokoja, Nigeria. *Arid Zone Journal of Engineering, Technology and Environment*, *5*: 1–12.

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *422–425*, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011

Kohler, M. A., & Linsley, R. K. (1951). Predicting the Runoff from Storm Rainfall, *U.S. Weather Bureau Research Paper*. No. 34.

Lehner, B. (2022). HydroSHEDS Technical Documentation Version 1.4; *World Wildlife Fund US*: Washington, DC 20037, USA.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions*, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Lehner, B., & Grill G. (2013). Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems. *Hydrological Processes, 27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Lehner, M. L., Grill, B., Nedeva, G., & Schmitt, I. O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*, 13603. https://doi.org/10.1038/ncomms13603

Lienou, G., Mahe, G., Dieulin, C., Paturel, J. E., Bamba, F., Sighomnou, D., & Dessouassi, R. (2010). The River Niger water availability: facing future needs and climate change. Global Change: Facing Risks and Threats to Water Resources (Proceedings of the Sixth World FRIEND Conference, Fez, Morocco, October 2010). IAHS Publications, 340:637-645.

Makwana, J. J., & Tiwari, M. K. (2017). Hydrological Stream Flow Modelling using Soil and Water Assessment Tool (SWAT) and Neural Networks (NNs) for the Limkheda Watershed, Gujarat, India. Modeling Earth Systemms and Environment, *3*: 635–645. https://doi.org/10.1007/s40808-017-0323-y

Mausbach, M., & Dedrick, A. (2004). The Length We Go Measuring Environmental Benefits of Conservation Practices. *Journal of Soil and Water Conservation*, *59*(5): 96–103.

Microsoft Open Source, McFarland, M., Emanuele, R., Morris, D., & Augspurger, T. (2022). microsoft/PlanetaryComputer: October 2022 (2022.10.28). *Zenodo*. https://doi.org/10.5281/zenodo.7261897

Minns, W., & Hall, M. J. (1996). Artificial neural networks as rainfall-runoff models. *Hydrological Sciences Journal*, *41*, 399–417. https://doi.org/10.1080/02626669609491511

Mohammadi, B., Moazenzadeh, R., Christian, K., & Duan, Z. (2021). Improving streamflow simulation by combining hydrological process-driven and artificial intelligence-based models.

Environmental Science and Pollution Research, *28*, 65752–65768. https://doi.org/10.1007/s11356-021-15563-1

Moran, P. (1970). Simulation and Evaluation of Complex Water Systems Operations, *Water Resources Research*, *6*, 1737–1742. https://doi.org/10.1029/WR006i006p01737

Muñoz, S. J. (2019). ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac (Accessed on 12-March-2022).

Murphy, K. P. (2022). Probabilistic Machine Learning: An Introduction. *The MIT Press*, ISBN 9780262046824 pp31-71.

Nash, J. E., & Sutcliffe, J. V. (1970). River Flow Forecasting through Conceptual Models Part I – A Discussion of Principles. *Journal of Hydrology*, *10*, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Ni, L., Wang, D., Singh, V. P., Wu, J., Wang, Y., Tao, Y., & Zhang, J. (2020). Streamflow and rainfall forecasting by two long short-term memory-based models. *Journal of Hydrology*, *583*, 124296. https://doi.org/10.1016/j.jhydrol.2019.124296

Noori, N., & Kalin, L. (2016). Coupling SWAT and ANN models for enhanced daily stream flow Prediction. *Journal of Hydrology*, *533*: 141–151. https://doi.org/10.1016/j.jhydrol.2015.11.050

Oguntunde, P. G., Abiodun, B. J., & Lischeid, G. (2014). A numerical modelling study of the hydroclimatology of Niger River Basin, West Africa. *Hydrological Sciences Journal*, *61*(1):94-106. https://doi.org/10.1080/02626667.2014.980260

Pfannerstill, M., Björn, G., & Fohrer, N. (2014). Smart Low Flow Signature Metrics for an Improved Overall Performance Evaluation of Hydrological Models. *Journal of Hydrology*, *510*: 447–458. https://doi.org/10.1016/j.jhydrol.2013.12.044

Okpara, J. N., Tarhule, A. A., & Perumal, M. (2013). Study of Climate Change in Niger River Basin, West Africa: Reality Not a Myth. *Climate Change – Realities, Impacts Over Ice Cap, Sea Level and Risks, Edited by Bharat Raj Singh, In Tech*. pp3-37.

Persits, F. M., Ahlbrandt, T. S., Tuttle, M. L., Charpentier, R. R., Brownfield, M. E., & Takahashi, K. I., (1997). Maps showing geology, oil and gas fields and geological provinces of Africa: *U.S. Geological Survey Open-File Report 97-470-A*, https://doi.org/10.3133/ofr97470A

Poméon, T., Diekkrüger, B., Springer, A., Kusche, J., & Eicker, A. (2018). Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins—A Remote Sensing Approach. *Water*, *10*(451). https://doi.org/10.3390/w10040451.

Pyo, J., Pachepsky, Y., Kim, S., Abbas, A., Kim, M., Kwon, Y. S., et al. (2023). Long short-term memory models of water quality in inland water environments. *Water Research X*, *21*, 100207. https://doi.org/10.1016/j.wroa.2023.100207.

Pyrcz, M. J., & Deutsch, C. V. (2018). Transforming Data to a Gaussian Distribution. In J. L. Deutsch (Ed.), Geostatistics Lessons. Retrieved from http://geostatisticslessons.com/lessons/normalscore

Schuol, J., Abbaspour, K. C., Srinivasan, R., & Yang, H. (2008). Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. *Journal of Hydrology*. *352*, 30–49. https://doi.org/10.1016/j.jhydrol.2007.12.025

Singh, D., Vardhan M., Sahu, R., Chatterjee, D., Chauhan, P., & Liu, S. (2023). Machine-learning- and deep-learning-based streamflow prediction in a hilly catchment for future scenarios using CMIP6 GCM data. *Hydrology and Earth System Sciences, 27*, 1047-1075. https://doi.org/10.5194/hess-27-1047-2023

Sonntag, D. (1990). Important new values of the physical constants of 1986, vapour pressure formulations based on the ITS-90, and psychrometer formulae. *Zeitschrift für Meteorologie*, *40*(5), 340–344.

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Thompson, J. R., Crawley, A. and Kingston, D. G. (2017). Future river flows and flood extent in the Upper Niger and Inner Niger Delta: GCM-related uncertainty using the CMIP5 ensemble. *Hydrological Sciences Journal*, *62*(14):2239-2265. https://doi.org/10.1080/02626667.2017.1383608

Van, S. P., Le, H. M., Thanh, D. V., Dang, T. D., Loc, H. H., & Anh, D. T. (2020). Deep Learning Convolutional Neural Network in Rainfall–Runoff Modelling. *Journal of Hydroinformatics*, *22*(3), 541–561. https://doi.org/10.2166/hydro.2020.095

Viessman Jr. W., & Lewis, G. L. (1996). Introduction to Hydrology, 4th ed. *Harper Collins*, New York. p165.

Vizi, Z., Batki, B., Rátki, L., Szalánczi, S., Fehérváry, I., Kozák, P., & Kiss, T. (2023). Water level prediction using long short-term memory neural network model for a lowland river: a case study on the Tisza River, Central Europe. *Environmental Sciences Europe*, *35*(92). https://doi.org/10.1186/s12302-023-00796-3

Vogel, R. M., Lall, U., Cai, X., Rajagopalan, B., Weiskel, P. K., Hooper, R. P., & Matalas, N. C. (2015). Hydrology: the interdisciplinary science of water. *Water Resources Research*, *51*, 4409–4430. https://doi.org/10.1002/2015WR017049

Xia, J., O'Connor, K. M., Kachroo, R. K., & Liang, G. C. (1997). A non-linear perturbation model considering catchment wetness and its application to river flow forecasting. *Journal of Hydrology*, *200*(1-4), 164–178. https://doi.org/10.1016/S0022-1694(97)00013-9

Xu, Y., Lin, K., Hu, C., Wang, S., Wu, Q., Zhang, L., & Ran,G. (2023). Deep transfer learning based on transformer for flood forecasting in data-sparse basins. *Journal of Hydrology*, *625*(A), 129956. https://doi.org/10.1016/j.jhydrol.2023.129956

Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., et al. (2021). ESA WorldCover 10 m 2020 v100. https://doi.org/10.5281/zenodo.5571936.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7