EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology, 2025, 26(2), 407–419
https://doi.org/10.12912/27197050/199448
ISSN 2719–7050, License CC-BY 4.0

Received: 2024.11.19
Accepted: 2024.12.21
Published: 2025.01.01

Optimizing the performance of rainflow-runoff models using multi-source satellite data-application to the region of Beni-Bahdel (Algeria)

Hiba Chahrouri[1], Abdelkader Bemmoussat[1*], Khaled Korichi[1]

[1] Civil and Environmental Engineering Laboratory, University of Djillali Liabes, BP 89. DZ-22000, Sidi Bel Abbés, Algeria
* Corresponding author's e-mail: bemmoussaek@yahoo.fr

ABSTRACT

The integration of satellite data in rainfall-runoff simulation is of paramount importance in regions where data are limited and not easily available. The data used in this study, concerning the daily precipitation and daily runoff during the period extending from 2007 to 2015, were measured and recorded in situ. This study aims primarily to optimize the performance of the modeling carried out, using data from various satellite sources in order to complete the missing data, and consequently to predict the liquid flow in the Beni Bahdel watershed, in the northwestern region of Algeria, by applying the long short-term memory (LSTM) learning model. It is important to know that the optimization of rainfall-runoff modeling is based on the use of satellite data relating to evapotranspiration, mean temperatures, minimum and maximum temperatures, net radiation, wind speed, and relative humidity. These data come from the NOAA CPC, ERA, ERA5_AG, ERA5-Land, GLDAS, CFSR and MERRA2 satellites. In addition, two statistical indicators were calculated to perform this optimization that is based on the LSTM approach that integrates remote sensing data, the coefficient of determination (R^2), and the Nash-Sutcliffe efficiency (NSE). Thus, a performance difference of about 0.30 was observed, for the NSE and R^2 coefficients, between the CFSR temperature data (NSE of 0.61and R^2 of 0.61) and the maximum and minimum temperature data from the ERA5-LAND and ERA5 satellite sources (0.92 for NSE and 0.93 for R^2), for the validation period. This significant difference suggests that the use of the minimum and maximum temperature data from the ERA5-LAND source allows achieving a rainfall-runoff modeling wih optimal performance. Indeed, the findings showed that quite high performances were achieved for the calibration period (0.93 for NSE and 0.95 for R^2) and for the validation period (0.92 for NSE and 0.93 for R^2).

Keywords: optimization, modeling, satellite data, long short-term memory, Beni Bahdel, Algeria

INTRODUCTION

It is widely acknowledged that remote sensing is an innovative information technology that plays a vital role in different fields (Huang et al., 2024). Remote sensing techniques, as indicated in a number of studies, are often used in agriculture (Khanal et al., 2020), environmental monitoring (Vasiliou & Economides, 2006), urban planning (Netzband et al., 2007), mining exploration (Ciampalini et al., 2013), as well as salinity-affected soil management (Dehni & Lounis, 2012). Likewise, in hydrology, remote sensing turns out to be highly important, particularly in rainfall-runoff modeling. The application of remote sensing allows converting multiple data collected by remote sensors into information that is exploited to better understand and predict hydrological flows. In particular, precipitation data are viewed as the essential tool that generally affects several disciplines, such as water resources management, planning, modeling and planning (Duan et al., 2019; Shekar et al., 2024; Ur Rahman et al., 2020).

A large number of studies, carried out throughout the world, have used remote sensing for the purpose of analyzing various hydrological

processes. For example, (Yu et al., 2024) conducted a study, in northern China, to develop a runoff simulation model using multi-source satellite data. For this, these researchers relied on the precipitation data from the precipitation estimation from remotely sensed information using artificial neural networks – climate data record (PERSIANN-CDR), surface temperature data from MOD11C1, a product developed by moderate resolution imaging spectroradiometer (MODIS), the evapotranspiration data, and soil moisture data from Global land evaporation amsterdam model (GLEAM). Similarly, in South Korea (Cho & Kim, 2022) conducted a study in which they exploited the local data assimilation and prediction system (LDAPS) meteorological data, while integrating precipitation, temperature data, surface pressure data, and some surface wind elements, with specific humidity and radiation data. Likewise, in Egypt (Abdelmoneim et al., 2020) used data from the tropical rainfall measuring mission (TRMM 3B42V7) and climate hazards group infra-red precipitation with station (CHIRPS) data satellite precipitation products as inputs to establish an appropriate hydrological model. As for Darand et al. (2017), they carried out a study in Iran with a view to assessing the tropical rainfall measuring mission (TRMM) multi-satellite precipitation analysis (TMPA) precipitation products. In addition, a particularly interesting study was conducted in Ethiopia by (Gebremicael et al., 2019) with the purpose of analyzing satellite precipitation data from satellite-based rainfall products CHIRPS, RFEv2 (Rainfall Estimates version 2), TRMM and PERSIANN.

It is important to point out that several Algerian researchers have used satellite-based rainfall products due to the scarcity of data in some areas. In this regard, (Lazri et al., 2020) performed a study that focused on improving precipitation estimation in northern Algeria. The remote sensing data used include multispectral images from the spinning enhanced visible and infrared imager (SEVIRI) radiometer as well as satellite precipitation data obtained from the climate prediction center morphing technique (CMORPH) and TRMM 3B42 products. In particular (Bemmoussat et al., 2021) benefited from the contribution of satellite precipitation for the rainfall-runoff hydrological modeling in the region of Hammam Boughrara (Algeria), using precipitation data from TRMM 3B43V7 and evapotranspiration data from Global land data assimilation system (GLDAS). Moreover, other studies were conducted by a number of researchers who used precipitation, wind, relative humidity, solar radiation data in the field of hydrogeology (Derdour et al., 2022; Inan et al., 2024). Nevertheless, to the best of our knowledge, the comparison of multi-source satellite data in rainfall-runoff modeling has not been sufficiently investigated in Algeria, which constitutes a good research opportunity for those working in this field.

It should be pointed out that, in the study area under consideration, we only have observed precipitation and flow data. It turns out that using neural networks for rainfall-runoff modeling represents a solution to this problem. In this case, only precipitation data are considered as input variable. However, it should be emphasized that the present work aims primarily to optimize the performance of flow simulation, while integrating input parameters from different satellite data sources, in addition to precipitation data measured and recorded in situ. These parameters were selected based on their relationship with the estimation of evapotranspiration.

MATERIALS AND METHODS

Presentation of the study region

In this article, the Beni Bahdel watershed was chosen and studied (Figure 1). It is located in the northwest of Algeria, between longitudes 1°8' W and 1°47' W, and latitudes 34°33' N and 34°41' N. The Beni Bahdel watershed covers an area of 989 Km^2 and crosses five rural municipalities, i.e. Ain Ghoraba, Beni Bahdel, Beni Snous, Azail and Sebdou, where the main activity is agriculture. In the year 1930, the Beni Bahdel concrete dam, with a capacity of 63 million m3, was built at the outlet of the watershed where the two main Wadis, namely Wadi Khemis and Wadi Sebdou, meet. It should be noted that the climate in this region is semi-arid. This area is characterized by a mountainous relief, with a predominance of karst geological formations. This favors the infiltration of precipitation (Bouguerra & Mansour, 2023).

Data used

Observed data

The observed data, which were used in this study, include information on daily precipitation (mm), and daily runoff (mm), from January

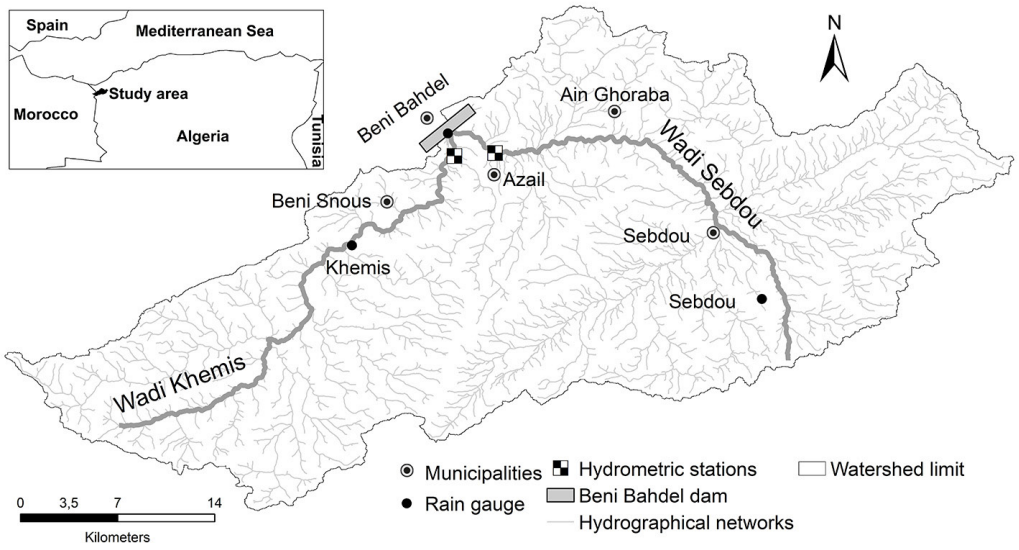

Figure 1. The study area map

1, 2007 to December 31, 2015. These data were provided by the National Agency for Hydraulic Resources (ANRH – Agence Nationale des Ressources Hydrauliques). In addition, the arithmetic means from the three rainfall stations, i.e. Khemis, Beni Bahdel and Sebdou stations, as well as the sum of the runoffs from the hydrometric stations, i.e. Wadi Khemis and Wadi Sebdou, were used for the rainfall-runoff modeling (Figure 1).

Satellite data

It is worth emphasizing that several satellite data were used as input parameters in the model, along with precipitation data, for the purpose of optimizing the performance of the flow simulation. These data include information on evapotranspiration, mean, maximum, and minimum temperature, radiation, wind speed, and humidity. All these data were downloaded for free from the Climate Engine, POWER, and Giovanni websites. These sites provide weather and climate data from the following sources:

- National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC)
- European Reanalysis Analysis 5th Generation (ERA5)
- European Reanalysis Analysis 5th Generation Agroclimatology (ERA5_AG)
- European Reanalysis Analysis 5th Generation land (ERA5 Land)
- Global Land Data Assimilation System (GLDAS).
- Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2)
- Climate Forecast System Reanalysis (CFSR) produced by National Centers for Environmental Prediction (NCEP); this system is managed by NOAA.

Hydrological modeling

Long-short term memory

Long short-term memory (LSTM) is a type of recurrent neural network (RNN) that was initially designed by (Hochreiter & Schmidhuber, 1997) and was then developed by many other researchers over time. It has been revealed that these neural networks can be used in modeling the dependencies of long time series data with high accuracy (Mao et al., 2021). In addition, LSTM can also be adopted for rainfall-runoff prediction because this approach represents a major advance to various fundamental problems in the hydrological domain (Hashemi et al., 2022). In this regard, many studies have been carried out by (Cho & Kim, 2022; Lees et al., 2021; Sabzipour et al., 2023) with the aim of comparing the performances of LSTM networks and classical hydrological models in the simulation of flows. The results proved that neural predictions are more efficient.

These positive results regarding neural networks motivated our choice to adopt the deep learning (DP) technique in our investigation. With regard to modeling, it was deemed interesting to use LSTMs as a black box, while respecting the following tasks:

Provide the model with the necessary input data such as the observed precipitation and other

complementary parameters from satellite sources. In addition, observed flow data were introduced to allow the model to train and evaluate the performance of the output flows for future periods.

Segment the study period into two phases. The first phase (2007–2013) for training and the second (2014–2015) for testing.Assign random values to the training parameters of the LSTM model. This operation is then repeated, while performing several tests, in order to improve the performance.

The LSTM optimization function

The performance of flow simulation mainly depends on the input data used. However, it can also be influenced by the optimization function of the LSTM model. This optimization function contains several hyperparameters whose values can impact the modeling performance. Before running the LSTM algorithm, it is essential to tune the values of all hyperparameters by testing several combinations because this allows obtaining the best possible performance in a reasonable time. This process is called hyperparameter optimization (Wu et al., 2019).

The hyperparameter optimization process can be approached using two different methods, i.e. a manual method or an automatic method. The manual approach has been adopted in this work. It is useful to know that this method is difficult to handle due to the number of parameters to be optimized and also to their value interval (Wu et al., 2019). Thus, to make the tuning of hyperparameters easier, it is highly recommended to set typical starting values (Table 1) in order to train and evaluate the model.

A large number of research studies have shown that the number of epochs is one of the most important parameters. According to (Greff et al., 2016), this parameter accounts for up to two-thirds of the overall performance. For this reason, special attention has been paid to this parameter during the optimization step.

Table 1. Typical parameters of LSTM (Ravindra, 2018)

Parameters	Typical starting values
Max epochs	250
Num hidden units	200
Gradient threshold	1
Initial learn rate	0.005
Learn rate drop factor	0.2
Learn rate drop period	125

Modeling steps

Figure 2 presents the modeling steps that are given below:

- Introduce the in situ measured data, namely precipitation and runoff. Precipitation should be used as input data for the LSTM model, while runoff is used for training the model and evaluating the performance of simulated runoffs.
- Choose the input data complementary to precipitation. These data are divided into five groups, and each group can contain one or more parameters. The data for these parameters come from several satellite sources.
- Introduce the initial hyperparameters.
- Run the LSTM model.
- Evaluate the performance using the NSE and R^2 indices.
- Tune the hyperparameters by testing several possible combinations, and then stop when the best values of R^2 and NSE are obtained.
- The hyperparameters that give the optimal performance are then used as initial hyperparameters for the next modeling group.
- Compare and interpret NSE and R^2 values.

Evaluation of hydrological modeling performance

Nash-Sutcliffe efficiency

The Nash-Sutcliffe efficiency (NSE) is a system of measurement that is widely used by researchers to assess the accuracy of models in flow simulation. It was first introduced by Nash and Sutcliffe in 1970 (Nash & Sutcliffe, 1970). It varies from minus infinity to 1. The value 1 indicates a perfect match between the observed and simulated values. The NSE can be calculated using the following formula:

$$NSE = 1 - \frac{\sum_{i=1}^{N}(Qobs_i - Qsim_i)^2}{\sum_{i=1}^{N}(Qobs_i - \overline{Qobs})^2} \quad (1)$$

where: $Qobs_i$ is the daily observed runoff, $Qsim_i$ is the daily simulated runoff, $\overline{Qobs}$ the average of observed runoff, and N the number of observations.

Furthermore, the NSE performance ranges are presented as follows (Ritter & Muñoz-Carpena,2013): for NSE ≥ 0.90, the model is excellent; for 0.80 < NSE < 0.90, the model is very satisfactory; for 0.60 < NSE < 0.80, the model is satisfactory; for NSE < 0.60: the model is poor.

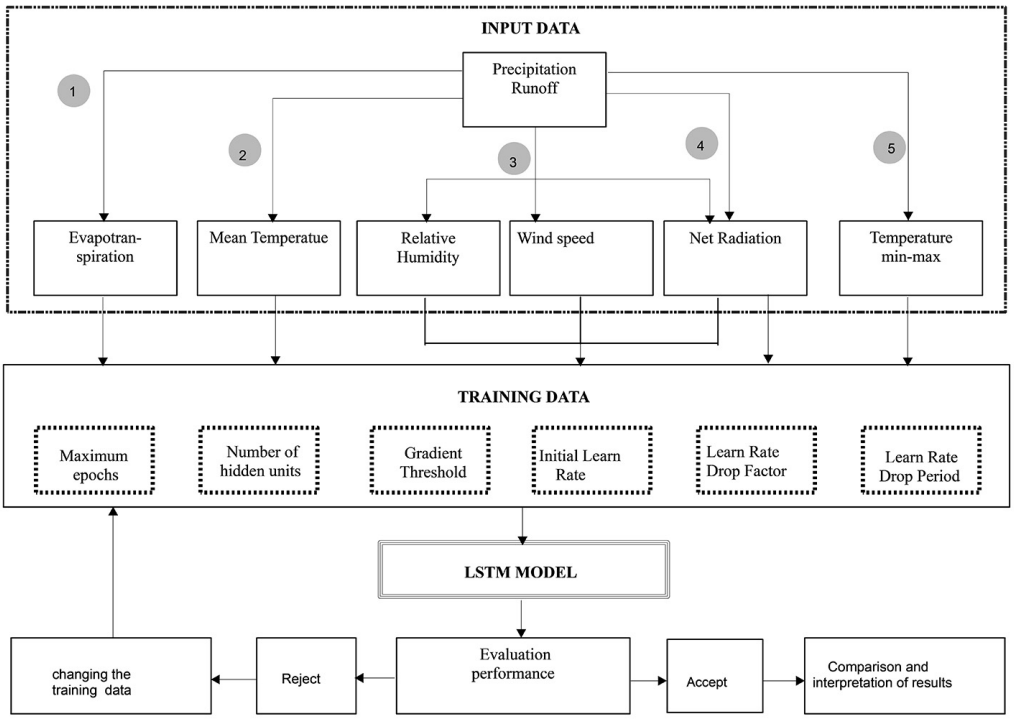

Figure 2. Modeling steps

Coefficient of determination

The coefficient of linear determination (R^2) is a qualitative indicator of a simple linear regression. It is commonly used in statistical studies; it allows evaluating the correlation between observed and simulated data (Renaud & Victoria-Feser, 2010). This coefficient can be calculated by means of the following formula:

$$R^2 = \frac{\Sigma_{i=1}^{N}(Qobs_i - \overline{Qobs}).(Qsim_i - \overline{Qsim})}{\Sigma_{i=1}^{N}(Qobs_i - \overline{Qobs})^2.\Sigma_{i=1}^{N}(Qsim_i - \overline{Qsim})^2} \quad (2)$$

where: $\overline{Qsim}$ is the average of simulated runoff.

It is worth pointing out that R^2 varies from 0 to 1. The fit is considered perfect for R^2 values equal to 1, and values greater than 0.5 are generally accepted (Santhi et al., 2001; Van Liew et al., 2003). With regard to the performance ranges of the R^2 coefficient, it was decided to adopt the same thresholds as for the NSE.

RESULTS AND DISCUSSIONS

Study of the linear correlation between rainfall and measured runoff

The graph depicted in Figure 3 shows a significant dispersion of points around the regression line, with a coefficient of determination almost equal to zero ($R^2 = 0.08$), which indicates a non-linearity between precipitation and measured runoff. These observations confirm that the relationship between these two variables is much more complex, which requires the use of a more sophisticated model that can be utilized to find a solution to this problem.

Study of the correlation between rainfall and measured runoff using the LSTM model

Analysis of the linear correlation between precipitation and runoff gave rather poor results. It was therefore deemed necessary to test the probable existence of a relationship with LSTM while considering only rainfall as input data. Consequently, very good results were obtained for both the training and test periods (training period: $R^2 = 0.92$ and NSE = 0.9; test period: $R^2 = 0.82$ and NSE = 0.79). The graphical results of the test period, which obviously are the most interesting, are presented below. Thus, (Figure 4**)** clearly shows that the points are grouped around the regression line, which confirms the existence of a significant correlation. No apparent shift is observed when the two curves depicted in Figure 5 are superimposed, which indicates a good fit between the observed and simulated data. These results

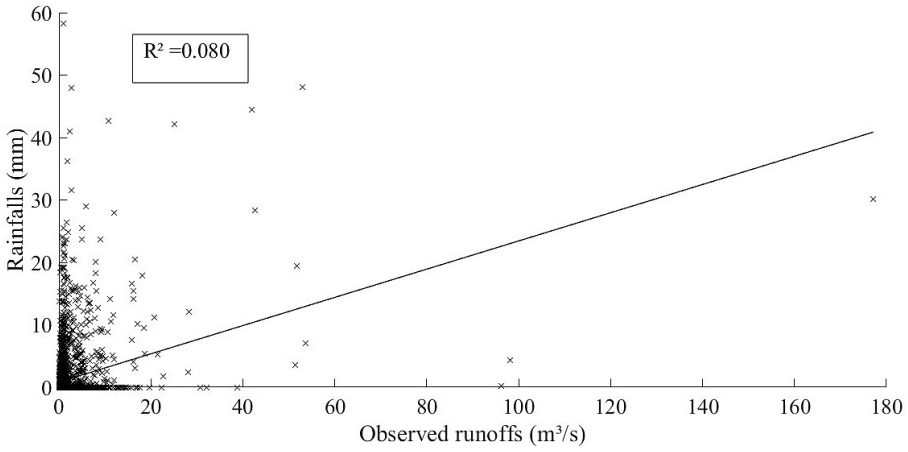

Figure 3. Linear correlation between precipitation and measured runoff, during the period between 2007 and 2015)

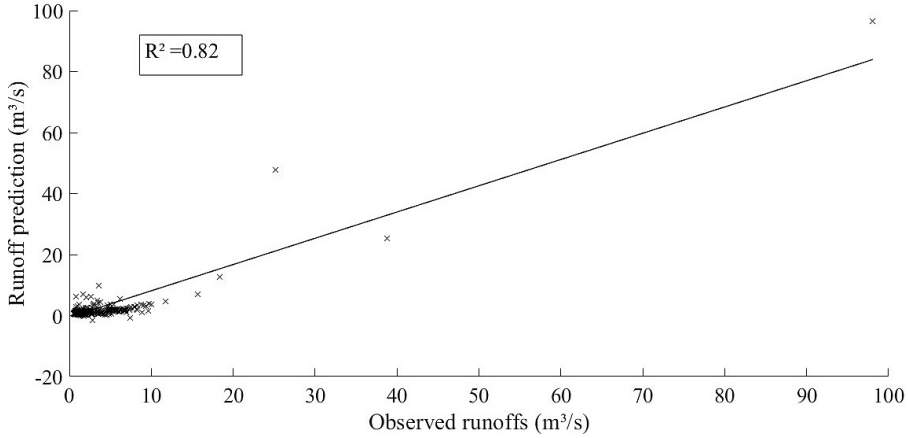

Figure 4. Linear regression between measured and simulated runoff, using the LSTM model, for the test period between 2014 and 2015

confirm that it is more advantageous to use neural networks than classical models which include at least evapotranspiration, in addition to rainfall as input data, and sometimes even other parameters, such as pedology, land cover, slope, and others. The development and adoption of deep learning (DL) is highly beneficial for hydrological modeling as it helps to detect the non-linear relationship between different variables, using only a limited number of data (Yu et al., 2024)

Rainfall-runoff modeling

The modeling, in which only precipitation was used as input, gave quite satisfactory results. However, these results were further improved

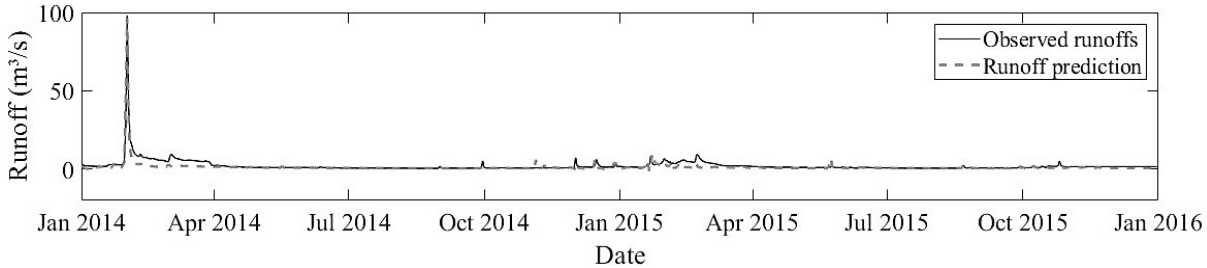

Figure 5. Observed daily runoff and simulated runoff, using only measured rainfall and observed runoff as inputs in the LSTM model

by adding other parameters (Table 2). It should also be noted that the performances are generally good, for training and prediction, with NSE values ranging from satisfactory to excellent, and R^2 values generally quite high. With regard to interpretation of the results obtained, it was agreed to focus mainly on the prediction ones which are considered essential for the evaluation, comparison and ranking of the models. These outcomes allow a better understanding of the effectiveness of the different models tested.

Interpretation of modeling results by input group

The main purpose of this section is to evaluate the performance of each dataset in order to analyze and compare the quality of modeling using input parameters from different satellite data sources.

- Evapotranspiration – the impact of different evapotranspiration data sources on rainfall-runoff modeling reveals variations in performance, with very satisfactory results obtained data from NOAA CPC and MERRA-2 systems. On the other hand, the results obtained from the ERA5-Land satellite are the weakest in comparison with the other sources. The high performances from NOAA CPC and MERRA-2 are certainly due to the quality and accuracy of the data that these satellites provide. Observations from both systems and ground stations are also integrated (Okirya et al., 2024; Yakubailik et al., 2020; Chaudhari et al., 2004). The method adopted here allows to better estimate the variations of evapotranspiration (Núñez et al., 2021; Olioso et al., 2022). It is noteworthy that evapotranspiration is a crucial factor in a region like northwestern Algeria where climatic conditions vary considerably due to the proximity to the Mediterranean Sea and the Atlas Mountains (Born & Bachner, 2003; Meddi, 2015). In addition, the NOAA CPC and MERRA-2 data are distinguished by their high spatial and temporal resolution, their advanced data assimilation techniques (Khatibi & Krauter, 2021), and their adaptability to the specific conditions of regions with sparse

Table 2. Performance results and training parameters, using different satellite data

Group	Parameters	Satellite sources	Training		Prediction	
			NSE	R^2	NSE	R^2
1	Evapotranspiration	NOAA CPC	0.92	0.94	0.84	0.85
		ERA5_AG	0.91	0.93	0.81	0.87
		GLDAS	0.88	0.89	0.73	0.78
		MERRA2	0.92	0.92	0.87	0.89
		CFSR	0.91	0.93	0.76	0.78
		ERA5	0.88	0.89	0.79	0.79
		ERA5 LAND	0.79	0.82	0.76	0.76
2	Temperature	ERA5_AG	0.95	0.96	0.91	0.91
		MERRA2	0.95	0.96	0.86	0.86
		CFSR	0.80	0.81	0.61	0.61
		ERA5	0.89	0.91	0.92	0.93
		ERA5 LAND	0.89	0.90	0.82	0.82
3	Radiation, wind speed, humidity	MERRA2	0.79	0.81	0.75	0.79
		CFSR	0.96	0.98	0.81	0.83
4	Radiation	GLDAS	0.91	0.93	0.78	0.87
		MERRA2	0.76	0.78	0.77	0.78
		CFSR	0.89	0.91	0.82	0.83
5	Temperature min-max	NOAA CPC	0.84	0.87	0.71	0.86
		ERA5_AG	0.92	0.94	0.84	0.89
		MERRA2	0.90	0.91	0.88	0.90
		CFSR	0.83	0.86	0.75	0.75
		ERA5	0.90	0.91	0.88	0.89
		ERA5 LAND	0.93	0.95	0.92	0.93

monitoring network, such as northwestern Algeria (El & Kerroumi, 2022; White et al., 2008)where the climatic conditions are quite complex and can vary significantly. Such an approach allows providing more accurate and reliable evapotranspiration estimates.

- Temperature – as concerns temperatures, the ERA5_AG data evidently achieve an excellent performance level. The very good outcomes are obviously attributed to the high resolution of this data source, the quality of reanalyses, the combination of multiple data sources, the long-term consistency of this source, its adaptability to various regions, as well as the continuous updates that ensure the accuracy and relevance of temperature data (Kassem et al., 2024). In compliance with these results, Ma et al. (2008) also found out that data from the global reanalysis dataset ERA-40 provided an excellent match with the temperature measurements that were recorded in China, although ERA-40 is a rather old version. These findings then suggest that the global weather dataset ERA5_AG is the most reliable data source because it offers better robustness and precision for temperatures compared to other sources. These findings corroborate the excellent NSE and R^2 results obtained.
- Min-max temperature – the minimum and maximum temperatures from the reanalysis dataset ERA5-Land allowed achieving an optimal flow simulation performance, which means that using min-max temperature data from the ERA5-Land source leads to superior performances compared to those obtained from the other data sources. The good results obtained are due to several reasons. First, the ERA5-Land source which offers a precise spatial resolution of 9 km (Arsenault et al., 2020) is particularly advantageous for detecting local climatic variations. These variations are crucial in a semi-arid region like northwestern Algeria where microclimatic variations are important due to the diversity of terrain and vegetation, and surface characteristics. Indeed, according to (Aimé & Remaoun, 1988), these circumstances can change rapidly over short distances. Next, the climate data provided by ERA5-Land are regularly updated and adjusted from multiple observations, which ensures greater reliability of extreme temperatures that are highly important in semi-arid environments where they can have a significant impact (Levesque, 2022). In addition, ERA5-Land adapts mainly to semi-arid regions where temperature directly affects hydrological processes, which helps to achieve more accurate runoff modeling. Finally, the extended temporal coverage and data consistency provided by ERA5-Land facilitate more efficient calibration of LSTM models (Muñoz-Sabater et al., 2021), which allows for more accurate simulations. All these observations lead to the conclusion that ERA5-Land outperforms other data sources.
- Radiation – the solar radiation results obtained from the CFSR reanalysis product show slightly better performance than other sources. The Nash and R^2 values obtained suggest that the model used simulates the flow with a fairly high accuracy using the CFSR radiation data. Moreover, the achieved performance was found quite consistent with the findings of (Khaled et al., 2014) who evaluated the accuracy of CFSR solar radiation data and compared it with that of the ground-based measurements made in the MENA region. The study found that CFSR data generally show good agreement with ground-based observations, especially in the regions of Egypt and Northwest Africa where the normalized mean bias coefficients were low and the correlation coefficients high, reaching the value of 0.982.
- Radiation, wind speed, humidity – the results of the modeling integrating radiation, wind speed and humidity, show that CFSR outperforms MERRA2 as input data source in the LSTM model. This is due to several factors. First, the spatial and temporal resolution of CFSR is slightly more precise than that of MERRA. Indeed, CFSR has a resolution of 0.5° x 0.5° (Saha et al., 2010), while MERRA-2 has a slightly lower resolution, namely 0.5° x 0.625° (Khatibi & Krauter, 2021). It should be emphasized that this accuracy is particularly important in regions, such as northwestern Algeria, where specific climatic conditions can have a significant impact on the results (Irvem & Ozbuldu, 2018; Khaled et al., 2014).

Furthermore, the diversity of data types integrated in CFSR is remarkable, including direct measurements of wind speed and humidity, from various sources, such as ocean buoys, weather

stations and aircraft (Essou et al., 2016; Saha et al., 2010). This multitude of data sources promotes a more robust modeling of dynamic parameters, unlike MERRA-2 which could have a data weighting and assimilation less adapted to these specific parameters (McClean et al., 2023).

Global interpretation by comparison of the results of the modeling of the different input parameters in the LSTM model from various satellite sources

The results obtained with the LSTM model, with the integration of different input parameters from various satellite data sources, exhibit a variability that affects the flow simulation performance. Indeed, a difference of 0.31 was recorded for NSE and 0.32 for R^2, between the temperature data from the CSFR (NSE = 0.61 and R^2 = 0.61) and the maximum and minimum temperature data from ERA5-Land and ERA5 (NSE =0.92 and R^2 = 0.93), which represents a very significant gap. In addition, the performances of the minimum and maximum temperatures from the ERA5-Land source and those of the temperature data from the ERA5 source are ranked equal, for the considered validation periods. However, when the training phase is integrated into the evaluation, the ERA5-Land dataset stands out with an excellent performance. It was indeed found that the Nash coefficient was equal to 0.93 and R^2 equal to 0.95 in the training phase, while the Nash coefficient was equal to 0.92 and R^2 to 0.93 in the validation period. Furthermore, (Figure 6) clearly shows that, for the test period, the data points are tightly clustered around the regression line, confirming an almost perfect correlation between the observed and simulated data.

It should also be noted that the superposition of the two curves (Figure 7) shows that there is almost no mismatch between them, proving the strong fit between the observed and simulated results.

The excellent performance obtained with ERA5-Land data can be explained as follows. It is observed that minimum and maximum temperatures have a direct impact on several hydrological processes, such as snowmelt, evaporation and evapotranspiration, which significantly affect the river runoff (Wang et al., 2020). In addition, these temperatures reveal a wide range of climatic conditions, which allows the model to better detect and predict watershed responses in extreme situations.

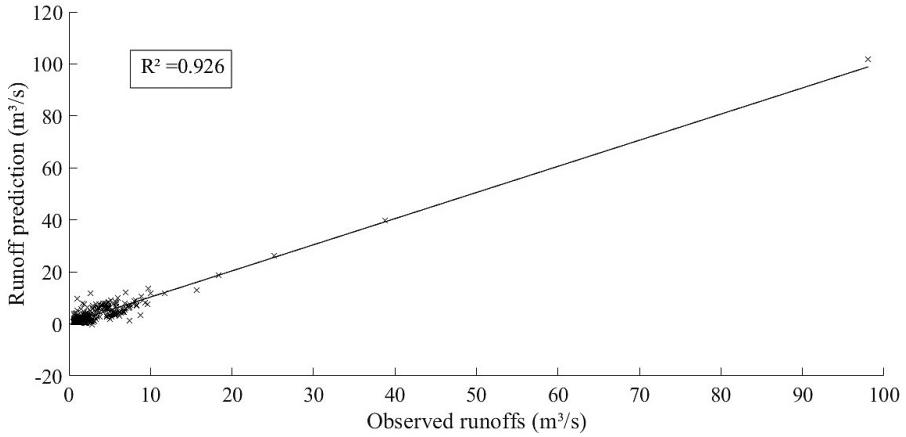

Figure 6. Linear regression between measured and simulated runoff, using the ERA5-Land dataset, for the test period from 2014 to 2015

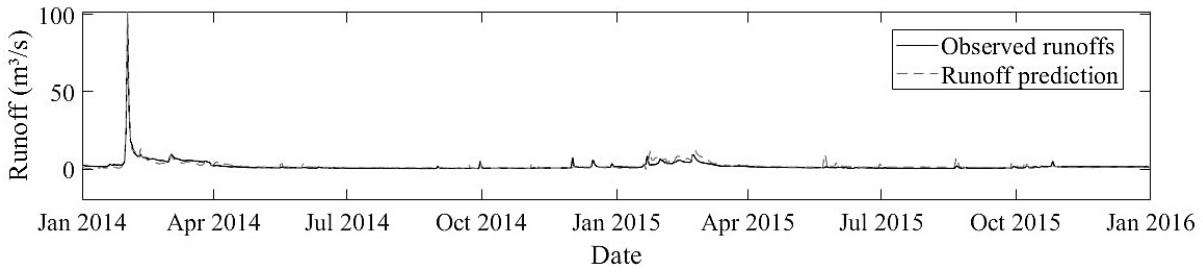

Figure 7. Observed and simulated daily runoffs, using the ERA5-Land dataset

This should increase and improve the accuracy of forecasts (Zhang et al., 2023). Likewise, the integration of minimum and maximum temperatures from remote sensing allows the LSTM model to better understand the complex dynamics of runoff in watersheds, hence maximizing its performance (Wang et al., 2024). In the same context, (Boulmaiz et al., 2020) have also consistently found that using these temperature data as inputs to the LSTM model contributes to significantly improving the modeling of the rainfall-runoff relationship. Furthermore, these temperature data are generally considered more reliable and less subject to measurement or estimation errors compared to other parameters, such as evapotranspiration or radiation (Yilmaz, 2023).

CONCLUSIONS

When satellite data as well as rainfall and runoff are used, the model shows generally good performance for training and prediction, with NSE and R^2 values ranging from satisfactory to excellent. The findings indicate that the temperature parameter presents the best performance, which corroborates its significant influence on evapotranspiration, a key factor in the water balance. In addition, the minimum and maximum temperatures from the ERA5-Land source showed excellent performance during the calibration and validation periods. They seem particularly well suited to capture the hydrological dynamics in our study area, outperforming all other parameters, such as evapotranspiration, mean temperature, radiation, as well as the combination of radiation, wind speed, and humidity, whose impact is more moderate. Based on these findings, it can be concluded that minimum and maximum temperatures are the most influential input parameters, and that the ERA5-Land data source is particularly effective for optimizing rainfall-runoff modeling with the LSTM model. However, it should be noted that the runoff simulation performance varies in accordance with the satellite data source used, because the choice of sources is highly crucial in modeling optimization. Moreover, a significant gap of nearly 0.30 was observed between the CSFR and ERA5-Land sources. The results of the present study highlight the effectiveness of satellite data in the LSTM model. They also emphasize the importance of correctly choosing the input data for rainfall-runoff modeling with recurrent neural networks. In the future, it would be interesting to explore other data sources to further improve the accuracy of hydrological forecasts in various contexts.

REFERENCES

1. Abdelmoneim, H., Soliman, M.R., Moghazy, H.M. 2020. Evaluation of TRMM 3B42V7 and CHIRPS Satellite Precipitation Products as an Input for Hydrological Model over Eastern Nile Basin. *Earth Systems and Environment* 4(4), 685–698. https://doi.org/10.1007/s41748-020-00185-3
2. Aimé, S., Remaoun, K. 1988. Climatic variability and steppisation in the Tafna basin (western Oranie) (in french). *Méditerranée* 63(1), 43–51. https://doi.org/10.3406/medit.1988.2528
3. Arsenault, R., Brissette, F., Martel, J.L., Troin, M., Lévesque, G., Davidson-Chaput, J., Gonzalez, M.C., Ameli, A., Poulin, A. 2020. A comprehensive, multisource database for hydrometeorological modeling of 14,425 North American watersheds. *Scientific Data* 7(1), 1–12. https://doi.org/10.1038/s41597-020-00583-2
4. Bemmoussat, A., Korichi, K., Baahmed, D., Maref, N., Djoukbala, O., Kalantari, Z., Bateni, S.M. 2021. Contribution of Satellite-Based Precipitation in Hydrological Rainfall–Runoff Modeling: Case Study of the Hammam Boughrara Region in Algeria. *Earth Systems and Environment* 5(4), 873–881. https://doi.org/10.1007/s41748-021-00256-z
5. Born, K., Bachner, S. 2003. Cyclogenesis and severe weather in the vicinity of the Atlas Mountains: studies using a nonhydrostatic mesoscale atmospheric model, in: EGS-AGU-EUG Joint Assembly. p. 11689. https://ui.adsabs.harvard.edu/abs/2003EAEJA....11689B/abstract
6. Bouguerra, S.A., Mansour, B. 2023. Rainfall-flow modeling using a global conceptual model: Case of the beni bahdel watershed (northwest of algeria). *Journal of Water Management Modeling*. https://doi.org/10.14796/JWMM.C500
7. Boulmaiz, T., Guermoui, M., Boutaghane, H. 2020. Impact of training data size on the LSTM performances for rainfall–runoff modeling. *Modeling Earth Systems and Environment* 6(4), 2153–2164. https://doi.org/10.1007/s40808-020-00830-w
8. Chaudhari, K.N., Sarkar, C., Patel, N.K., Parihar, J.S. 2006. an Inter-Comparison of Satellite Based Noaa Cpc Rainfall Estimates and Gauge Observations Over Selected Stations in India. *Proc of ISPRS symposium on Geospatial databases for Sustainable Development* 27–30. https://www.isprs.org/proceedings/xxxvi/part4/RS-A-3.pdf

9. Cho, K., Kim, Y. 2022. Improving streamflow prediction in the WRF-Hydro model with LSTM networks. *Journal of Hydrology* 605, 127297. https://doi.org/10.1016/j.jhydrol.2021.127297
10. Ciampalini, A., Garfagnoli, F., Del Ventisette, C., Moretti, S. 2013. Potential Use of Remote Sensing Techniques for Exploration of Iron Deposits in Western Sahara and Southwest of Algeria. *Natural Resources Research* 22(3), 179–190. https://doi.org/10.1007/s11053-013-9209-5
11. Darand, M., Amanollahi, J., Zandkarimi, S. 2017. Evaluation of the performance of TRMM Multi-satellite Precipitation Analysis (TMPA) estimation over Iran. *Atmospheric Research* 190, 121–127. https://doi.org/10.1016/j.atmosres.2017.02.011
12. Dehni, A., Lounis, M. 2012. Remote sensing techniques for salt affected soil mapping: Application to the Oran region of Algeria. *Procedia Engineering* 33, 188–198.https://doi.org/10.1016/j.proeng.2012.01.1193
13. Derdour, A., Benkaddour, Y., Bendahou, B. 2022. Application of remote sensing and GIS to assess groundwater potential in the transboundary watershed of the Chott-El-Gharbi (Algerian–Moroccan border). *Applied Water Science* 12(6), 136. https://doi.org/10.1007/s13201-022-01663-x
14. Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., Yang, L., Mekonnen, D.F. 2019. Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia. *Journal of Hydrology* 569, 612–626. https://doi.org/10.1016/j.jhydrol.2018.12.026
15. El, N., Kerroumi, I. 2022. Study of the interannual variability of temperature and precipitation in the high plateaus (in french) 6, 65–77. https://onm-blog.meteo.dz/wp-content/uploads/2024/08/Article_Kerroumi_2.pdf
16. Essou, G.R.C., Sabarly, F., Lucas-Picher, P., Brissette, F., Poulin, A. 2016. Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling? *Journal of Hydrometeorology* 17(7), 1929–1950. https://doi.org/10.1175/JHM-D-15-0138.1
17. Gebremicael, T.G., Mohamed, Y.A., Zaag, P. van der, Gebremedhin, A., Gebremeskel, G., Yazew, E., Kifle, M. 2019. Evaluation of multiple satellite rainfall products over the rugged topography of the Tekeze-Atbara basin in Ethiopia. *International Journal of Remote Sensing* 40(11), 4326–4345. https://doi.org/10.1080/01431161.2018.1562585
18. Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J. 2016. LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems* 28(10), 2222–2232. https://doi.org/10.48550/arXiv.1503.04069
19. Hashemi, R., Brigode, P., Garambois, P.A., Javelle, P. 2022. How can we benefit from regime information to make more effective use of long short-term memory (LSTM) runoff models? *Hydrology and Earth System Sciences* 26(22), 5793–5816. https://doi.org/10.5194/hess-26-5793-2022
20. Hochreiter, S., Schmidhuber, J. 1997. Long Short-Term Memory. *Pakistan Journal of Zoology* 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
21. Huang, X., Lu, K., Wang, S., Lu, J., Li, X., Zhang, R. 2024. Understanding remote sensing imagery like reading a text document: What can remote sensing image captioning offer? *International Journal of Applied Earth Observation and Geoinformation* 131, 103939. https://doi.org/10.1016/j.jag.2024.103939
22. Inan, C.A., Maoui, A., Lucas, Y., Duplay, J. 2024. Multi-Station Hydrological Modelling to Assess Groundwater Recharge of a Vast Semi-Arid Basin Considering the Problem of Lack of Data: A Case Study in Seybouse Basin, Algeria. *Water (Switzerland)* 16(1), 160. https://doi.org/10.3390/w16010160
23. Irvem, A., Ozbuldu, M. 2018. Accuracy of satellite-based solar data to estimate solar energy potential for Hatay province, Turkey. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi* 7(2), 361–369. https://doi.org/10.17798/bitlisfen.428757
24. Kassem, Y., Camur, H., Preala, T.A. 2024. Assessment of Wind Energy Potential for achieving Sustainable Development Goal 7 in the Rural Region of Jeje, Nigeria. *Engineering, Technology & Applied Science Research* 14(4), 14977–14987. https://doi.org/10.48084/etasr.7311
25. Khaled, W., Authority, E.M., Afandi, G.S. El. 2014. Evaluation of Ncep / Cfsr Solar Data Against Ground Observation Over MENA. *Open Journal of Atmospheric and Climate Change* 1, 1–12. https://doi.org/10.15764/ACC.2014.02001
26. Khanal, S., Kushal, K.C., Fulton, J.P., Shearer, S., Ozkan, E. 2020. Remote sensing in agriculture—accomplishments, limitations, and opportunities. *Remote Sensing* 12(22), 1–29. https://doi.org/10.3390/rs12223783
27. Khatibi, A., Krauter, S. 2021. Validation and performance of satellite meteorological dataset merra-2 for solar and wind applications. *Energies* 14(4), 882. https://doi.org/10.3390/en14040882
28. Lazri, M., Labadi, K., Brucker, J.M., Ameur, S. 2020. Improving satellite rainfall estimation from MSG data in Northern Algeria by using a multi-classifier model based on machine learning. *Journal of Hydrology* 584, 124705. https://doi.org/10.1016/j.jhydrol.2020.124705
29. Lees, T., Buechel, M., Anderson, B., Slater, L., Reece, S., Coxon, G., Dadson, S.J. 2021. Benchmarking data-driven rainfall-runoff models in Great

Britain: A comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. *Hydrology and Earth System Sciences* 25(10), 5517–5534. https://doi.org/10.5194/hess-25-5517-2021

30. Levesque, G. 2022. Evaluation of the use of ERA5 and ERA5-Land reanalysis data for evapotranspiration modelling (in french). *Doctoral dissertation, École de technologie supérieure*. https://espace.etsmtl.ca/id/eprint/3085
31. Ma, L., Zhang, T., Li, Q., Frauenfeld, O.W., Qin, D. 2008. Evaluation of ERA-40, NCEP-1, and NCEP-2 reanalysis air temperatures with ground-based measurements in China. *Journal of Geophysical Research Atmospheres* 113, D15115. https://doi.org/10.1029/2007JD009549
32. Mao, G., Wang, M., Liu, J., Wang, Z., Wang, K., Meng, Y., Zhong, R., Wang, H., Li, Y. 2021. Comprehensive comparison of artificial neural networks and long short-term memory networks for rainfall-runoff simulation. *Physics and Chemistry of the Earth* 123, 103026. https://doi.org/10.1016/j.pce.2021.103026
33. McClean, F., Dawson, R., Kilsby, C. 2023. Intercomparison of global reanalysis precipitation for flood risk modelling. *Hydrology and Earth System Sciences* 27(2), 331–347. https://doi.org/10.5194/hess-27-331-2023
34. Meddi, H. 2015. Variability of annual rainfall in north-western Algeria (in french). *Science et changements planétaires/Sécheresse* 20(1), 57–65. https://doi.org/10.1684/sec.2009.0169
35. Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.N. 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021
36. 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
37. Netzband, M., Stefanov, W.L., Redman, C. 2007. Applied remote sensing for urban planning, governance and sustainability. *Springer Science & Business Media*.https://doi.org/10.1007/978-3-540-68009-3
38. Núñez, P.Á., Silva, B., Schulz, M., Rollenbeck, R., Bendix, J. 2021. Evapotranspiration estimates for two tropical mountain forest using high spatial resolution satellite data. *International Journal of Remote Sensing* 42(8), 2940–2962. https://doi.org/10.1080/01431161.2020.1864058
39. Okirya, M., Du Plessis, J. 2024. Trend and Variability Analysis of Annual Maximum Rainfall Using Observed and Remotely Sensed Data in the Tropical Climate Zones of Uganda. *Sustainability* 16, 6081. https://doi.org/10.3390/su16146081
40. Olioso, A., Jacob, F., Olioso, A., Jacob, F., De, E., Olioso, A. 2002. Estimation of evapotranspiration from remote sensing measurements (in french). *la houille blanche* 88(1), 62–67. https://doi.org/10.1051/lhb/2002008
41. Ravindra, B. 2018. Forecasting solar radiation during dust storms using deep learning. *arXiv preprint arXiv:1808.10854*. https://doi.org/10.48550/arXiv.1808.10854
42. Renaud, O., Victoria-Feser, M.P. 2010. A robust coefficient of determination for regression. *Journal of Statistical Planning and Inference* 140(7), 1852–1862. https://doi.org/10.1016/j.jspi.2010.01.008
43. Ritter, A., Muñoz-Carpena, R .2013. Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology* 480, 33-45. https://doi.org/10.1016/j.jhydrol.2012.12.004
44. Sabzipour, B., Arsenault, R., Troin, M., Martel, J.L., Brissette, F., Brunet, F., Mai, J. 2023. Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. *Journal of Hydrology* 627. https://doi.org/10.1016/j.jhydrol.2023.130380
45. Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, Jiande, Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, Jun, Hou, Y.T., Chuang, H.Y., Juang, H.M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van Den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G., Goldberg, M. 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* 91(8), 1015–1057. https://doi.org/10.1175/2010BAMS3001.1
46. Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M. 2001. Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of the American Water Resources Association* 37(5), 1169–1188. https://doi.org/10.1111/j.1752-1688.2001.tb03630.x
47. Shekar, P.R., Mathew, A., Yeswanth, P. V., Deivalakshmi, S. 2024. A combined deep CNN-RNN network for rainfall-runoff modelling in Bardha Watershed, India. *Artificial Intelligence in Geosciences* 5, 100073. https://doi.org/10.1016/j.aiig.2024.100073

48. Ur Rahman, K., Shang, S., Shahid, M., Wen, Y. 2020. Hydrological evaluation of merged satellite precipitation datasets for streamflow simulation using SWAT: A case study of Potohar Plateau, Pakistan. *Journal of Hydrology* 587, 125040. https://doi.org/10.1016/j.jhydrol.2020.125040
49. Van Liew, M.W., Arnold, J.G., Garbrecht, J.D. 2003. Hydrologic simulation on agricultural watersheds: Choosing between two models. *Transactions of the ASAE* 46(6), 1539–1551. https://doi.org/https://doi.org/10.13031/2013.15643
50. Vasiliou, A., Economides, A.A. 2006. MANETs for environmental monitoring. *2006 International Telecommunications Symposium, IEEE* 813–818. https://doi.org/10.1109/ITS.2006.4433383
51. Wang, H., Zhang, J., Yang, J. 2024. Time series forecasting of pedestrian-level urban air temperature by LSTM: Guidance for practitioners. *Urban Climate* 56, 102063. https://doi.org/10.1016/j.uclim.2024.102063
52. Wang, K., Xu, Q., Li, T. 2020. Does recent climate warming drive spatiotemporal shifts in functioning of high-elevation hydrological systems? *Science of the Total Environment* 719, 137507. https://doi.org/10.1016/j.scitotenv.2020.137507
53. White, J.W., Hoogenboom, G., Stackhouse Jr, P.W., Hoell, J.M. 2008. Evaluation of NASA satellite-and assimilation model-derived long-term daily temperature data over the continental US. *agricultural and forest meteorology* 148(10), 1574–1584. https://doi.org/10.1016/j.agrformet.2008.05.017
54. Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., Deng, S.-H. 2019. Hyperparameter optimization for machine learning models based on Bayesian optimization. *Journal of Electronic Science and Technology* 17(1), 26–40. https://doi.org/10.11989/JEST.1674-862X.80904120
55. Yakubailik, O.E., Yakubalik, T. V. 2020. Analysis of accumulated precipitation based on satellite data in Central Siberia. *IOP Conference Series: Earth and Environmental Science* 548, 032025. https://doi.org/10.1088/1755-1315/548/3/032025
56. Yilmaz, M. 2023. Accuracy assessment of temperature trends from ERA5 and ERA5-Land. *Science of the Total Environment* 856, 159182. https://doi.org/10.1016/j.scitotenv.2022.159182
57. Yu, C., Hu, D., Shao, H., Dai, X., Liu, G., Wu, S. 2024. Runoff simulation driven by multi-source satellite data based on hydrological mechanism algorithm and deep learning network. *Journal of Hydrology: Regional Studies* 52, 101720. https://doi.org/10.1016/j.ejrh.2024.101720
58. Zhang, Y.-R., Shang, G.-F., Leng, P., Ma, C., Ma, J., Zhang, X., Li, Z.-L. 2023. Estimation of quasi-full spatial coverage soil moisture with fine resolution in China from the combined use of ERA5-Land reanalysis and TRIMS land surface temperature product. *Agricultural Water Management* 275, 107990. https://doi.org/10.1016/j.agwat.2022.107990