Yanbu Journal of Engineering and Science

Manuscript 1217

Evaluation of Gridded Precipitation Products and Machine-Learning Models for Rainfall Prediction in Arid Regions - Case Study in Qassim Region, KSA

Ragab Rabeiy

Follow this and additional works at: https://yjes.researchcommons.org/yjes

The journal in which this article appears is hosted on Digital Commons, an Elsevier platform.

YANBU JOURNAL OF ENGINEERING AND SCIENCE 2026;23:1–15

Scan the QR to view the full-text article on the journal website

RESEARCH ARTICLE

Evaluation of Gridded Precipitation Products and Machine-Learning Models for Rainfall Prediction in Arid Regions - Case Study in Qassim Region, KSA

Ragab Rabeiy

Department of Civil Engineering, College of Engineering, Qassim University, Buraydah 51452, Saudi Arabia

ABSTRACT

Receiving rainfall predictions in dry regions such as Saudi Arabia is challenging due to limited observational stations and the temporal and spatial variability of rainfall events. This study explores the performance of nine remote sensing rainfall datasets over the Qassim region of central Saudi Arabia. High-resolution rainfall data from ground-based sources were used to evaluate the accuracy of CHIRPS, TRMM, GPM, PERSIANN, CMORPH, CFSR, TerraClimate, TerraClimate_Monthly, and ERA5 datasets. Performance metrics such as R^2, RMSE, MAE, and bias were calculated to quantify the accuracy of each dataset. Analysis revealed that ERA5 generally demonstrated better accuracy than the other datasets with a mean R^2 and RMSE, MAE, and bias values of 0.59, 3.91, 2.89, and 0.43, respectively. The algorithms were trained, validated, and applied to forecast daily, 5-day, 10-day, and monthly total rainfall. Root mean square error (RMSE) values decreased by up to 52%, 50%, 56%, and 45% for daily, 5-day, 10-day, and monthly rainfall forecasts, respectively. To improve the accuracy of future predictions, multivariable models were developed that incorporated mean air temperature with precipitation predictors. Five machine-learning algorithms were tested: Gradient Boosting Regressor (GBR), Histogram Gradient Boosting Regressor (HGBR), Random Forest Regressor (RFR), Extreme Gradient Boosting Regressor (XGBR), and an advanced version of XGBR. For the 70/30 train-validation split over time, the models based on boosting methods (HGBR and XGBR) demonstrated the highest predictive ability and yielded the best statistical performance ($R^2 = 0.90$) with the lowest errors. Combining a suite of reanalysis datasets with advanced machine learning strategies enables effective rainfall prediction in arid regions that lack adequate observational data. Such a strategy facilitates rainfall estimation in ungauged catchments, enabling water resource management and flash-flood risk assessment in the desert domain.

Keywords: Rainfall prediction, Gridded precipitation datasets, Machine learning, ERA5 reanalysis, Arid climate, Qassim region

1. Introduction

Precision precipitation information is critical for hydrological applications, including water resources planning, drought assessment, and flash-flood hazard assessment. Although the Kingdom of Saudi Arabia (KSA) is an arid region, rainfall occurs episodically and is spatially inhomogeneous. The rain-gauge network in KSA is not adequate to support reliable hydrological applications, particularly at local scales, owing to the high spatial variability of precipitation over very short distances [1]. Hence, remote sensing precipitation products and reanalysis datasets have been increasingly exploited. In arid regions such as KSA, the accuracy of these products is challenged, particularly in the detection of light precipitation events and convective systems due to limitations in satellite measurements and model dynamization. As a result, the accuracy of rainfall occurrence (wet-day frequency) and rainfall amount (rainfall intensity) distribution is often questioned [2]. Satellite-based (CHIRPS, TRMM, GPM, PERSIANN, CMORPH, and

Received 30 April 2026; revised 22 June 2026; accepted 16 July 2026.
Available online 29 July 2026

E-mail address: r.elsayed@qu.edu.sa (R. Rabeiy).

https://doi.org/10.53370/3122-3303.1217
3122-3303/© 2026 The Author(s). Yanbu Journal of Engineering and Science. This is an open access article under the CC BY 4.0 Licence (https://creativecommons.org/licenses/by/4.0/).
Published by Yanbu Industrial College, Saudi Arabia. The journal is hosted on Digital Commons, an Elsevier Platform.

others) and reanalysis datasets (CFSR, TerraClimate, ERA5, and others) are widely used to augment available rain-gauge information for hydrological applications. Despite some errors in validation analyses, these datasets have been found useful for hydrological applications, due to their ability to provide data over extended spatial and temporal scales. However, the accuracy of these products shows large variations in different climate regimes and for different temporal scales, especially for heavy precipitation events [3–5]. The Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) is a quasi-global precipitation product that combines observations from infrared satellite data with in-situ station data to form a "smart" interpolation of rainfall. The data have a resolution of 0.05° (5 km) and begin in 1981. CHIRPS is useful for drought monitoring and a variety of climate assessments. It has been applied to several hydrologic and drought-related studies [6]. In Helmi & Abdelhamed (2022), [7] evaluations of KSA arid region performance showed CHIRPS is able to capture broad seasonal performance but exhibits generally low daily correlation and high intensity-dependent biases.

ERA5 is a global atmospheric data product. It is based on an upgraded data assimilation system and on the latest generation of ECMWF's weather and climate prediction models. It provides a physically consistent data record of meteorological weather conditions from 1979 to the present [8]. Reanalysis products offer better performance than satellite retrievals in data-sparse regions where monthly to seasonal rainfall is often constrained by model physics and various observing systems on these basins. In the study of Helmi & Abdelhamed (2022), [7], an assessment of ERA5 precipitation products' performance over the Arabian Peninsula is presented, and it is shown that although they perform relatively well for climatological purposes, there are limitations in terms of performance at the daily event scale. The Integrated Multi-Satellite Retrieval for GPM (IMERG) is the official precipitation estimate for the Global Precipitation Measurement (GPM) mission. It combines data from microwave and infrared-sensing satellites and is calibrated and interpolated for maximum accuracy. As the newest precipitation-measuring mission, it offers improvements in estimating rainfall amounts and distributions globally, but has its own limitations in arid regions, where it is sensitive to the intermittency of individual storms and to sensor detection thresholds [9]. The performance of the Integrated Multi-satellite Retrievals for GPM (IMERG) product is evaluated over Saudi Arabia and other Arabian arid regions at a monthly scale and compared with daily scale performance. The performance of IMERG is also investigated in terms of heavy precipitation events and the effects of topography [10]. The Tropical Rainfall Measuring Mission (TRMM) multi-satellite precipitation analysis (TMPA) is a multi-sensor TRMM precipitation analysis that merges data from several orbiting platforms with current and past data from manually calibrated platforms, as well as gauge data. The performance of TMPA has been validated and reported globally using a suite of data. Even though performance is relatively poor at the daily time scale, performance improves at longer periods, such as monthly and annual [11]. Several studies have used TRMM data and analysis to investigate the climatology and extremes of rainfall in the Kingdom of Saudi Arabia (KSA). Those studies have found intensity-dependent biases [11, 12].

CFSR (Climate Forecast System Reanalysis) is a coupled atmosphere-ocean-land-sea ice reanalysis produced by NCEP for 1979 to present. The CFSR provides a consistent historical climate data record estimate of many climate variables, including precipitation, using a single, continuous model simulation that incorporates numerous observational datasets. Complete precipitation data should also be available from this source. Although certain characteristics, such as differences in model physics and assimilation systems, are expected to result in differences compared to ERA5, CFSR can be used in regions with low gauge density where traditional gauge-based datasets lack adequate coverage [13]. Reanalysis products such as CFSR have been shown to perform better as the area-averaged increases but typically retain systematic wet/dry bias errors [7]. CMORPH (Climate Prediction Center MORPHing technique) is a global precipitation product developed by morphing passive microwave rainfall estimates using motion vectors derived from infrared imagery data to produce high-resolution satellite-based precipitation information. While the product effectively captures the movement of storms, issues with sampling errors and light or warm precipitation that affect validation performance are found, particularly over arid regions [13]. Based on evaluations of performance over a range of climatic regions across the Kingdom of Saudi Arabia, it was found that the daily correlations with corresponding reference sites were generally low; however, the correlations improved with time, particularly on a monthly time scale. Performance also exhibited mixed characteristics with respect to intensity classification [7]. TerraClimate is a monthly data product at 4 km spatial resolution that includes precipitation and a set of climate/water-balance variables. The product is interpolated using climatologically aided interpolation, which combines information from a high-resolution climatology and a

coarser time-varying data product. The data is most suitable for monthly water-balance applications, such as drought indicators and water analyses, and is not as well-suited for representing individual rainfall events daily. Monthly gauge data can also be used to create monthly gridded products, but these are prone to larger errors related to the intensity distribution and the fraction of wet days, particularly in regions with few gauges [14].

PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) uses a machine-learning/neural-network approach to estimate precipitation from infrared satellite data [15]. While PERSIANN products are broadly utilized, such as in PERSIANN-CDR, for near real-time rainfall estimation at high spatial and temporal resolutions, their daily performance is particularly problematic over dry regions globally due to intensity-dependent biases and false alarms [7]. Although point-scale measurements from rain gauges are the most reliable data type, the scarcity of gauges in arid regions makes them inadequate for basin assessments. Many studies have demonstrated that the performance of models in predicting precipitation is highly sensitive to temporal scale, especially at daily time steps. Besides the error in absolute amount associated with the timing error at the daily time scale, the other error sources at the daily time scale are related to the scale mismatch (grid cell versus point gauge location) and the drizzle effect (inaccurate prediction of small amounts of rainfall), especially in arid regions. An analysis of the errors at the basin-regional scale revealed that the scale mismatch error and drizzle effect are the dominant sources of error for precipitation prediction in arid regions [16, 17].

In contrast, monthly data are better for reducing uncertainties stemming from random errors, temporal shift differences, and improving correspondence with actual observations as well as for determining climatological rainfall trends [18, 19]. Although there are several studies focused on arid and semi-arid regions using such data, most of them have relatively low correlation (or NSE) for daily data due to issues of point-to-grid mismatches caused by the short lifespan of convective events, while the performance improves significantly for monthly data [12, 20]. Among them, IMERG and reanalysis data, especially ERA5, usually exhibit relatively better performance in terms of large-scale patterns and monthly total [10]. However, their timing accuracy for events and the accuracy of light precipitation can still be an issue. In addition to these aspects of total amount, the characteristics of rainfall intensity also have an impact on flood analysis, which is different from those for general climate analysis [7]. Specifically, the so-called “drizzle effect”, where some products overestimate light precipitation and incorrectly represent heavy precipitation, often show larger errors. Discussion points regarding previous studies include differences in topography, season, gauge density, time scale (monthly versus daily), and evaluation methodology [10]. Some studies include bias correction into their evaluation approach, while others use detection and occurrence indices in addition to the correlation coefficient and NS. Our approach is meaningful in that each product is evaluated over the period during which it overlaps with the gauge record. This aspect is particularly important in arid regions where data length is limited.

Machine-learning (ML) methods have attracted considerable attention in hydrological and climatological modeling, especially for modeling complex, non-linear relationships among atmospheric variables. In hydrology, traditional statistical models are generally inadequate for modeling the stochastic behavior of precipitation, which exhibits high variability and strong temporal and spatial randomness. In arid and semi-arid regions, rainfall is typically sparse, intermittent, and multiscalar, being influenced by a multitude of climate modes. In contrast to traditional statistical models, ensemble-based and tree-based machine-learning approaches have been shown to outperform conventional models for estimating and predicting rainfall by modeling complex, non-linear relationships between large numbers of input variables effectively [21].

Predicting things in hydrology can be done by employing various techniques, such as Random Forest Regression (RFR). RFR is an interesting method because, despite its high accuracy in predicting, it does not need to be excessively complex, and it can handle as many variables as necessary. RFR constructs many decision trees and generates predictions by using bootstrap aggregation and picking features randomly, improving accuracy and stability in predictions. Applications of RFR in hydrology have been consistently superior to traditional regression methods for rainfall-runoff modeling [22, 23]. The usual suspects, including KNearestModel, Random Forest Regressor, Gradient Boosting Regressor (GBR), Gradient Boosting in R, etc., did an awesome job also. Also, boosting-based methods like GBR are very effective for time series prediction, where the key is to gradually improve the accuracy of predictions using multiple decision trees, where each subsequent tree focuses on correcting the residual errors from the previously generated predictions, finally leading to achieving high accuracy of predictions [24]. Building on this method, Histogram Gradient Boosting Regressor (HGBR) further improves computation

efficiency by pre-binning continuous variables into histograms before training, achieving faster execution without significant loss of accuracy. This is particularly useful for large climatological datasets where computation cost is often a limiting factor [25]. In addition to LightGBM, another well-known boosting approach is Extreme Gradient Boosting (XGBR), which includes regularization, parallelization, and improved tree learning to avoid overfitting [26]. In various meteorology and hydrology applications, XGBoost has been successfully employed for rainfall nowcasting, satellite precipitation correction, climate variable reconstruction, and other tasks, frequently achieving higher accuracy and higher efficiency when compared with other ensemble approaches [24, 27].

The main objective of this research work is therefore to evaluate some of the currently available gridded data sources of precipitation and to develop some machine learning-based methods to enhance the accuracy of rainfall forecasts for the Qassim region in the Kingdom of Saudi Arabia. In addition to the above main objective, this research work also has four specific objectives. The first one is to compare the performance of some satellite and reanalysis rainfall products with ground observations. The second one is to characterize the rainfall and to identify the pattern of wet-day frequency. The third one is to apply a bias correction and a multivariable regression model using some precipitation and temperature variables as predictors of rainfall. The fourth one is to evaluate some machine learning methods and to identify the best-performing one to estimate rainfall in arid regions.

2. Materials and methods

2.1. Study area

The study area is located in the Qassim region of the Kingdom of Saudi Arabia (KSA) within the vast Najd Plateau that extends from latitudes 24° N to 27° N and longitudes 41° E to 45° E, covering approximately 65,000 km^2. The region is characterized by an arid desert climate with semi-arid to arid seasonal variations that experience extremely hot summer seasons with mean maximum temperatures exceeding 45 °C in July and August, and very mild winter seasons with mean maximum temperatures ranging from 25 °C in January to 30 °C in February. The region's elevation is gently undulating, with maximum height exceeding 900 m above sea level. The low relief means that the area has no permanent flow, except for the occasional ephemeral streams that are characterized by low frequency and high magnitude of flow due to rare seasonal events of rainfall (Fig. 1). Seasonal rainfall occurs in the region at different degrees of magnitude, with an average annual rainfall ranging from 50 mm to 150 mm. Seasonal variation in rainfall is marked, with most of the rainfall occurring from November to April, with high interannual variation in magnitude. The region has experienced several severe flash flood events, such as those that occurred in September 2009 and September 2010. These storms originate as convective systems that travel over the area from a northwesterly direction, resulting in a short-term high magnitude of flow in the shallow streams.

2.2. Data sources

2.2.1. Ground-based meteorological data

Daily rainfall observations from a ground meteorological station in the Qassim region in Saudi Arabia have been used as reference data. The station data contains almost five decades (1969 - 2016) of time series of daily, monthly, and annual observations of rainfall. Only after critical processing of the observations, containing missing and erroneous values and outliers, was this station data made available for comparing the quality of various gridded precipitation products with data based on satellite and numerical weather prediction model information.

2.2.2. Gridded precipitation datasets

Gridded rainfall datasets were downloaded from the ClimateEngine platform (https://app.climateengine.org), which provides access to multiple satellite-based and reanalysis climate products. Nine widely used precipitation products were considered CHIRPS, ERA5, PERSIANN, TRMM, CFSR, GPM, TerraClimate, TerraClimate_Monthly, and CMORPH. For comparison purposes, daily rainfall data for grid cells containing station locations, monthly and annual sums or accents of rainfall, as well as average air temperature (°C) for corresponding month/years were obtained from ERA5, available through the ClimateEngine platform developed by VictoriaMR. Hourly atmospheric variables (including ERA5) are available at 0.25° from the European Centre for Medium-Range Weather Forecasts (ECMWF). For this purpose, the daily mean temperature was computed.

2.3. Data processing and evaluation

To assess the performance of gridded precipitation products or datasets to capture rainfall in arid environments, several statistical performance indicators were calculated to quantify differences between gridded model rainfall and ground-measured rain. These performance indicators included the coefficient of determination (R^2), the Root Mean Square Error

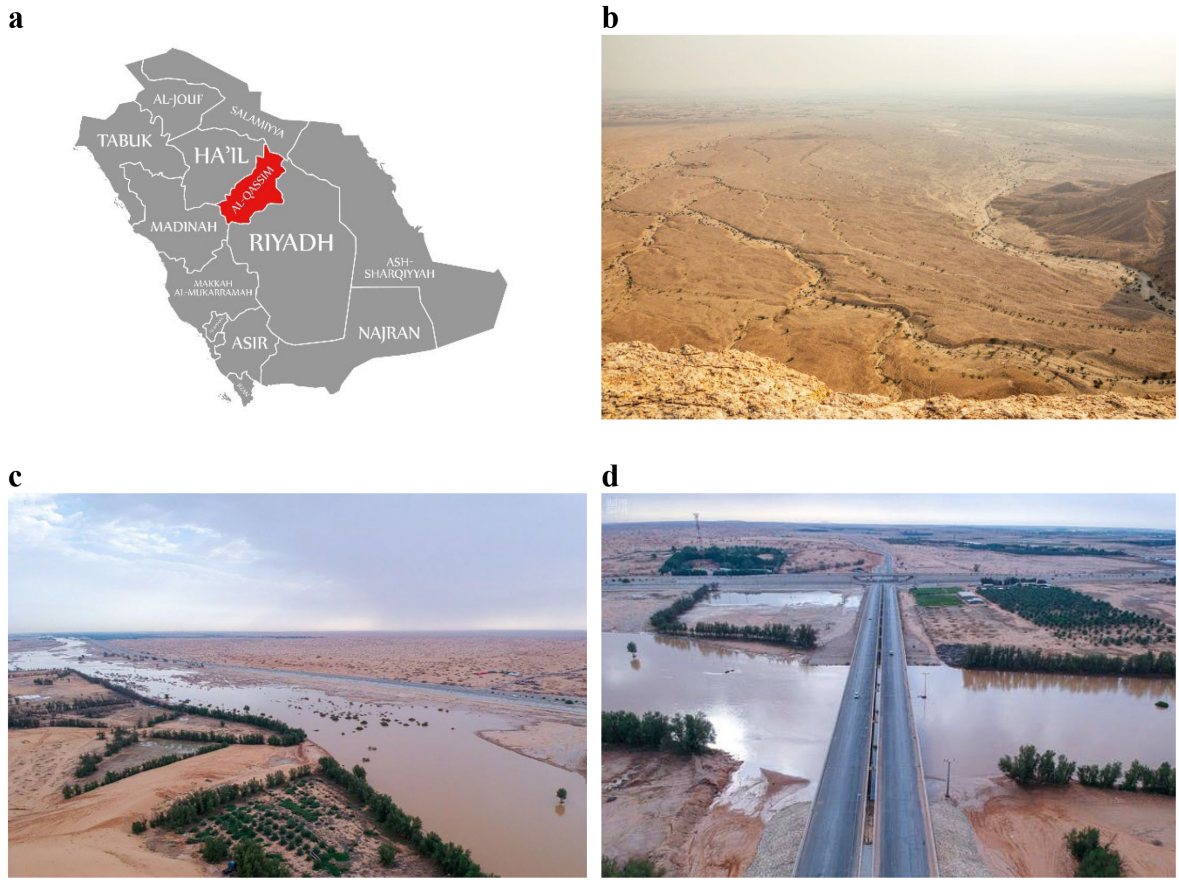


Fig. 1. Location and general characteristics of the Qassim region, KSA: (a) Map of Qassim Region, (b) Najd Plateau Desert Landscape, (c) and (d) Satellite imagery of the main wadi during flash flood.

(RMSE), Mean Bias Error (MBE), Nash–Sutcliffe Efficiency (NSE), and Pearson correlation coefficient (r) at both daily and monthly time scales. In addition, scatter plots with 1:1 reference line were created to visually assess the degree of error in the gridded datasets. Bias assessments for high-intensity rainfall events were also performed, and the results were analyzed for flood prediction and assessment.

3. Results and discussion

Annual total rainfall time series for the study area are presented in Fig. 2. The total annual rainfall from observed daily rainfall data for the period of record (1969 – 2016) clearly shows very large year-to-year variability, with some years receiving as little as 25 mm of rainfall and other years receiving more than 300 mm of rainfall. Some years had very large amounts of rainfall, including the early 1970s, mid-1990s, and around 2008 – 2009. Most years had less than 150 mm of rain, and only a few years were above 250 mm. The time series does not show any long-term increase or decrease in the time series across the record period. The rainfall is clearly dominated by a strong signal of episodic events, characterized by alternating years of very dry and very wet conditions, with minimal long-term trend. Such a regime is typical of arid to semi-arid climates worldwide and has important implications for the management of water resources and flood risk. Importantly, it indicates that a small number of very wet years can have a large impact on long-term average hydrological outcome, with drier years, which are much more common, having much less impact.

Monthly and annual average rainfall from ground measurements are compared to corresponding satellite-based estimates for nine different datasets in Fig. 3. Rainfall in Qassim region is highly seasonal, exhibiting large amplitude during some months while remaining low during others. In general, the majority of the precipitation occurs between March and April and November and December, while the summer months remain very dry due to Qassim's arid climate. ERA5 generally has the smallest error when compared to other datasets for monthly averages,

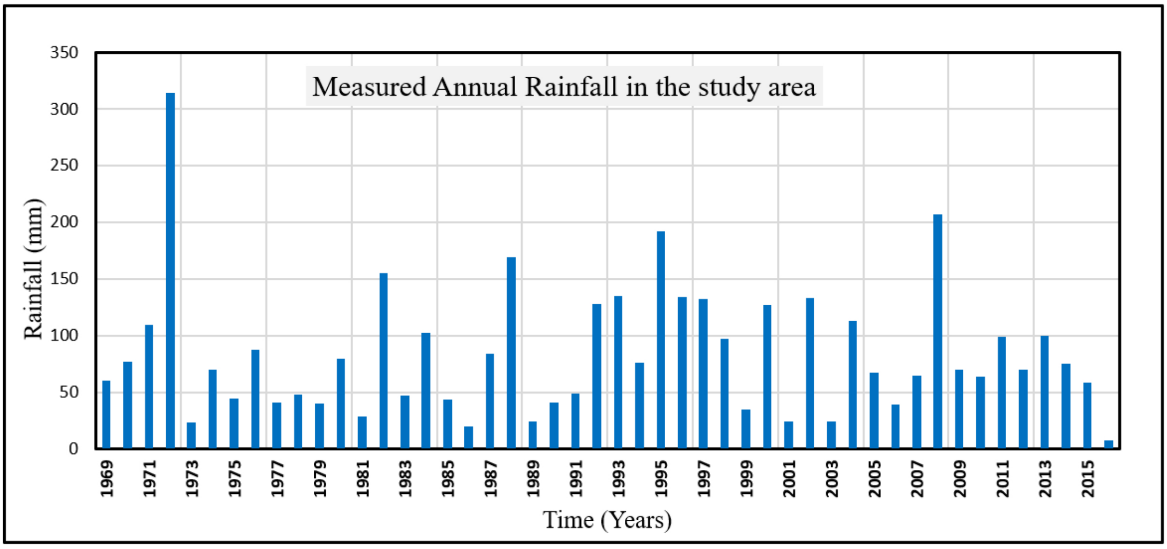


Fig. 2. Annual rainfall totals recorded at the ground meteorological station in the Qassim Region, Saudi Arabia, during the period 1969–2016.

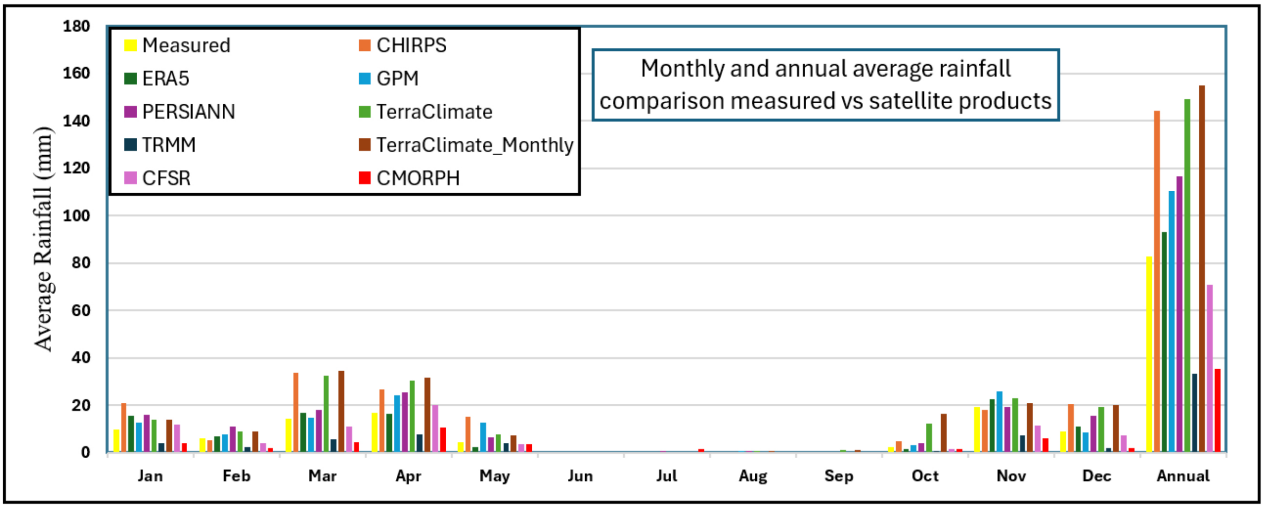


Fig. 3. Comparison of monthly and annual average rainfall derived from ground measurements and multiple satellite products (CHIRPS, ERA5, PERSIANN, TRMM, CFSR, GPM, TerraClimate, TerraClimate_Monthly, and CMORPH).

as well as the annual total. Other products, such as GPM and PERSIANN, capture the large seasonal variations but exhibit overestimation errors during the highest months of precipitation. CHIRPS and TerraClimate exhibit large overestimation errors in annual totals, while TRMM and CMORPH generally underestimate both monthly and annual rainfall totals. Annually, ERA5 remains the best of these datasets with the smallest error. In general, the reanalysis-based datasets perform better than the satellite-based datasets for precipitation estimates in arid regions of the study area.

Fig. 3 shows that ERA5 is best in mimicking the observed rainfall on both monthly and annual time scales, a finding that is in conformity with existing works over arid regions. This is because of the fundamental advantage of reanalysis datasets containing information from atmospheric models and ground observations over purely satellite-derived products. In arid regions, previous studies showed ERA5 outperforms other satellite-based precipitation products in stability and accuracy [28, 29]. Moderate performance was attained by the remotely sensed products, especially for near-real-time rainfall estimation. The results indicated that GPM and PERSIANN generally tend to overestimate rainfall, particularly in the peak rainy months. Maghsood et al., and Nguyen et al., [30, 31] reached a similar conclusion. Overestimation occurs, mainly in strong precipitation events, due to the challenges in discriminating between light precip-

Table 1. Wet-day frequency analysis based on daily rainfall and eight gridded rainfall datasets.

Method	0.5–1	1–5	5–10	10–25	25–50	50–100
CHIRPS	18	406	287	163	22	1
ERA5	120	242	75	45	2	0
GPM	143	289	73	44	2	1
PERSIANN	490	931	160	57	4	0
TerraClimate	0	91	56	108	85	38
TRMM	61	172	28	9	0	0
CFSR	0	0	0	0	0	0
CMORPH	131	174	32	14	1	0
Measured	57	265	109	101	18	3

itation and background noise, especially in infrared- and passive-microwave-based retrievals, which are mostly affected by the atmospheric conditions. On the other hand, CHIRPS and TerraClimate (TRC) overly estimated rainfall (especially during the main rainy season). This is typical for arid and data-poor regions around the world where the small number of rain gauges are used for bias correction of these products which leads to overestimation. Ashouri et al., and Funk et al., [32, 33] pointed out that the scarcity of rain gauges in desert regions affects blended rainfall products negatively and tends to increase estimated totals, whereas satellite-based products such as TRMM and CMORPH tend to underestimate rainfall, especially for light precipitation events, which are common in arid regions and sensitive to the poor performance of satellite sensors in detecting weak, convective precipitation. Similar conclusions were drawn by Alijanian et al., and Wolters et al., [34, 35], showing that satellite-based rainfall products do not capture small-scale events common in dry climates and therefore, any data used for hydrological or flood-risk studies should be thoroughly validated. ERA5 performed best of the data sets studied here, and other data products can be used if appropriately bias corrected for use in dry climates. These findings are useful for future flood hazard and risk assessment studies, water resources management, as well as impact studies of climate change.

Table 1 compares the percentage of wet days from ground-measured daily rainfall and eight satellite precipitation products. Wet days are further grouped into intensity classes: 0.5–1, 1–5, 5–10, 10–25, 25–50, and 50–100 mm/day. These classes distinguish between light and heavy rainfall events typical of arid environments. Results generally indicate that most satellite products overestimate the frequency of light rainfall (0.5–5 mm/day), particularly in arid environments. Three products (CHIRPS, TerraClimate, and PERSIANN) had the largest counts of low-intensity wet days. This overestimation of light rainfall is well-documented in arid and semi-arid regions and is particularly a limitation of infrared-based retrievals that are highly sensitive to cold cloud tops that may or may not produce measurable rainfall [36, 37]. Both ERA5 and GPM show best performance for medium (5-25 mm/day) rainfall rates, particularly for aggregated observations, with the highest correlation with measurement data for rainfall intensities corresponding to moderate, synoptic-scale, and organized convective precipitation. Previous studies have demonstrated improvement in rainfall retrieval for such an aggregated scale for several reanalysis products and multi-sensor missions that provide global rainfall estimates [5, 8]. Heavy rainfall events with rainfall amounts greater than 25 mm per day are extremely rare in arid regions. Although all considered datasets have biases, some heavy precipitation events are overestimated by TRMM and CMORPH, while others are not captured by some datasets. These errors can be attributed to sampling errors and retrieval uncertainties, particularly for short-duration convective events prevalent in desert environments [11, 38]. The frequency patterns of wet-day errors in this assessment are consistent with what has been observed in previous studies in arid and semi-arid regions and are related to the “drizzle effect,” where satellite products typically overestimate the frequency of wet days while underestimating the amount of rainfall, as discussed by Satgé et al., [39]. Both ERA5 and GPM are noted to be the most reliable products for detecting rainfall occurrence and for the representation of rainfall intensity distributions in dry climates [1, 36].

While satellite-based rainfall products are important sources of information for rain gauge-free areas of the world, their data should not be routinely used in hydrologic applications without assessment of possible errors. In this paper, the error associated with wet-day frequency is examined in terms of rainfall intensity. Evaluation of wet-day frequency as a function of rainfall intensity in classes provides a much more detailed perspective on the performance of individual products than standard accuracy statistics, which simply provide an overall wet-day bias and root mean square error. Correct estimation of wet-day frequency and intensity is critical for the assessment of water balance and flood potential in arid regions. Overestimation of light precipitation can introduce biases into soil moisture and runoff estimates, while errors in heavy precipitation events result in even greater errors for estimates of flood potential. Assessments based on several widely used satellite rainfall products indicate ERA5 and GPM generally provide the best estimates of wet-day frequency and intensity. Other products contain significant biases and should be used with caution, possibly with calibration to available local gauge data.

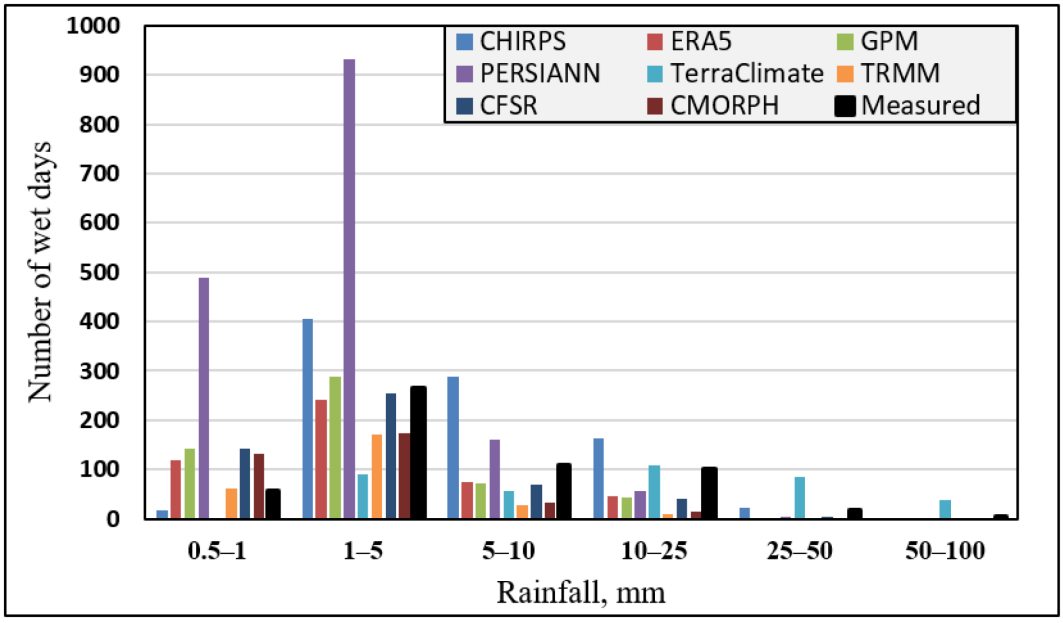


Fig. 4. Comparison of wet-day frequency distribution for ground-measured observations and eight gridded precipitation datasets.

Table 2. Comparison of daily precipitation estimates from eight gridded products and gauge measurements.

Method	R^2	Bias	Error%	ME	MAE	RMSE	NSE	n (days)
CHIRPS	0.03	0.58	-73.36	-0.17	0.53	2.86	-0.72	12935
ERA5	0.19	0.83	-20.41	-0.04	0.32	2.01	0.09	6726
GPM	0.23	0.76	-32.0	-0.07	0.33	2.03	0.13	5844
PERSIANN	0.06	0.65	-54.05	-0.12	0.45	2.15	-0.20	11613
TerraClimate	0.00	0.00	-154104	-11.98	12.00	25.5	-18995	566
TRMM	0.06	2.47	60	0.12	0.24	2.05	0.05	6726
CFSR	0.07	1.09	8.37	0.02	0.32	2.21	-0.10	6726
CMORPH	0.06	2.72	63.3	0.13	0.24	2.05	0.05	6726

Table 2 presents daily statistics between satellite/reanalysis rainfall and ground observations for different products at 0.1° grid scale. Due to the arid nature of the region, with little rainfall, infrequent events, and high spatial and temporal variability (i.e., most rainfall occurs as short-period convective events over a very small area), the agreement between any remotely sensed/reanalysis and in situ rainfall is expected to be poor and only fair on a day-to-day basis at a point. Therefore, the focus is placed on the more meaningful monthly scale. The results do show that the daily satellite/reanalysis rainfall estimates have a very weak to poor relationship with in situ measurements, as indicated by R^2 values ranging from 0.00 to 0.23. The best daily performance is that of GPM with R^2, NSE, RMSE, and MAE values of 0.23, 0.13, 2.03 mm/day, and 0.33 mm/day, respectively, which outperforms ERA5 with R^2, NSE, RMSE, and MAE values of 0.19, 0.09, 2. Figs. 3 and 4 provide a monthly scale evaluation of the considered satellite rainfall products relative to in situ measurements. Performance at daily time scales fundamentally differs from that at monthly time scales, especially in arid/semi-arid regions. Daily rainfall at a single station is unlikely to be representative of the mean rainfall in the corresponding grid cell on a particular day, due to the stochastic nature of rainfall with limited areal extent. Satellite sampling still suffers from timing/location errors, and retrieval algorithms have inherent limitations with light precipitation, warm clouds, and certain atmospheric conditions. However, on monthly time scales, random errors tend to cancel out, and the physical consistency of the water balance becomes more critical than the accuracy of event timing.

In contrast to the most random errors and decreased sampling noise observed in daily comparisons, the month-scale data revealed some increase in skill for some of the datasets, as shown in Table 3. ERA5 provided the best agreement with the gauge data (R^2 = 0.49; NSE = 0.47; RMSE = 12.27 mm/month; MAE = 5.33 mm/month), indicating that it can better capture month-to-month variability in rainfall. GPM was second best (R^2 = 0.46; NSE = 0.42; RMSE = 13.28 mm/month), indicating an improvement in the

Table 3. Monthly rainfall validation metrics for eight satellite and reanalysis precipitation datasets relative to ground measurements.

Method	R^2	Bias	Error%	ME	MAE	RMSE	NSE	n (months)
CHIRPS	0.29	0.58	-73.36	-5.13	8.93	16.32	0.03	425
ERA5	0.49	0.83	-20.41	-1.30	5.33	12.27	0.47	221
GPM	0.46	0.76	-31.99	-2.12	5.57	13.28	0.42	192
PERSIANN	0.33	0.70	-43.38	-3.05	7.07	15.38	0.15	390
TerraClimate	0.29	0.58	-71.64	-5.01	9.41	19.94	-0.50	567
TRMM	0.36	2.47	59.48	3.80	5.35	15.04	0.20	221
CFSR	0.33	1.09	8.37	0.53	5.63	14.08	0.30	221
CMORPH	0.20	2.72	63.29	4.04	5.70	15.73	0.13	221

capabilities of the retrievals from this newer platform. While CFSR showed moderate ability at representing month-scale totals ($R^2 = 0.33$; NSE = 0.3; Bias = 1.09), the bias was relatively close to 1.0, and this dataset is generally best suited for representing long-term totals rather than month-scale totals. While both TRMM and CMORPH showed positive error percentages, the bias ratios revealed that both datasets tended to overestimate monthly totals during the overlap period. In fact, the bias for TRMM was found to be 2.47, which is consistent with other dry region validation studies that have noted a tendency for these types of algorithms and data to overestimate totals due to both structural and sampling issues. While CHIRPS and TerraClimate are two of the most widely used datasets for drought monitoring and long-term hydrologic studies, both datasets showed very poor to weak performance for this study area. In contrast, PERSIANN was found to have intermediate skill ($R^2 = 0.33$; NSE = 0.15) but with a bias of 0.7, indicating underestimation. On the basis of these monthly statistics, ERA5 is found to be the best of the tested datasets for representing month-scale gauge-based rainfall in arid regions, with GPM the second best performing data. Monthly ERA5 rainfall data outperform their GPM-based counterparts in terms of both RMS error and NSE performance. The bias of the ERA5 model was generally close to 1 (on average 0.95) while its error had smaller RMS values compared to GPM (averaging 3.9 mm against 8 mm, respectively). While GPM performed well in terms of the linear relationship with the observed data, high dispersion and inefficiency characterized the performance over arid environments. Therefore, ERA5 was used to predict rainfall in other locations based on their elevation and latitude.

To better understand the performance of ERA5 and GPM monthly rainfall estimates over arid regions, Fig. 5 displays scatter plots of monthly measured rainfall (mm) compared to gridded precipitation from ERA5 and GPM against corresponding ground measurements. Moderate correlation ($R^2 = 0.5$) exists between both datasets and measured rainfall. However, ERA5 clearly exhibits a better correlation with measured rainfall ($R^2 = 0.49$) compared to GPM ($R^2 = 0.46$). Large dispersion is noted, particularly at higher rainfall amounts; large numbers of observations lie below the 1:1 line at high rainfall. Both datasets generally underestimate monthly rainfall amounts, especially at high intensities. Conversely, a strong agreement exists at lower rainfall (i.e., <20 mm) with both datasets demonstrating improved performance in arid regions for predicting rain-free months as well as light rainfall events. Overall, ERA5 performed marginally better, but both have deficiencies in the capture of high-intensity rainfall.

To improve rainfall prediction models, mean air temperature was chosen as a supplementary variable and incorporated into multivariable prediction models with precipitation. Monthly rainfall amounts from ERA5, along with the monthly mean air temperature, served as predictors in the machine learning regression algorithm. Five machine learning models were evaluated, namely Gradient Boosting Regressor (GBR), Histogram Gradient Boosting Regressor (HGBR), Random Forest Regressor (RFR), Extreme Gradient Boosting Regressor (XGBR), and an advanced XGBR configuration (XGBR_Advanced) Model development employed a chronological split of 70% of the data for training and 30% for validation to preserve the temporal structure of the rainfall series. Several engineered features were incorporated, including one-month lag rainfall (Lag1), two-month lag rainfall (Lag2), a three-month rolling rainfall average (Roll3), and an interaction term between ERA5 rainfall and temperature. A logarithmic transformation of the target rainfall variable was applied to reduce data skewness and improve model stability. To ensure physical consistency, all negative rainfall predictions generated by the machine-learning models were constrained to zero. Since no future meteorological information was used, the proposed framework was not intended for rainfall forecasting or nowcasting. Rather, it was designed as a statistical bias-correction and downscaling approach to reconstruct gauge-observed monthly rainfall from

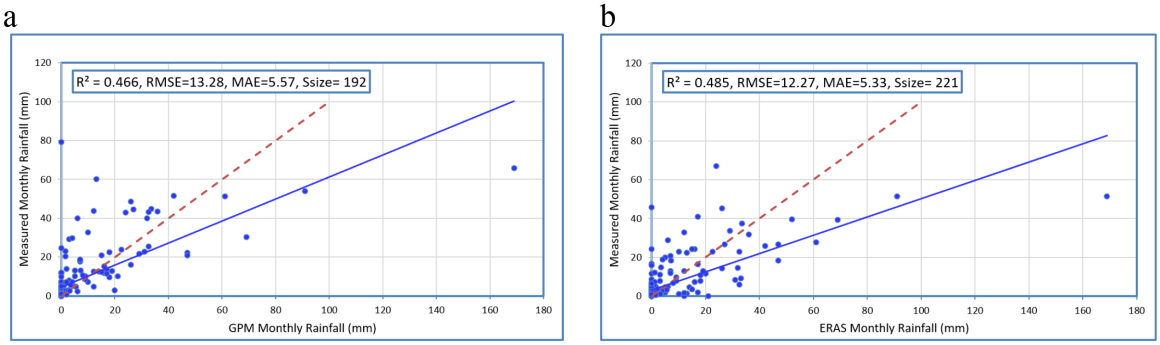


Fig. 5. Scatter plots compare measured monthly rainfall with two gridded precipitation datasets: (a) GPM monthly rainfall and (b) ERA5 monthly rainfall.

Table 4. Comparative performance metrics of machine-learning models (GBR, HGBR, RFR, XGBR, and XGBR_Advanced) for rainfall prediction during training and validation phases.

Model	Split	R^2	MSE	RMSE	MAE	Bias	Bias_%	Obs_sum	Pred_sum
GBR	Train	0.88	40.3	6.3	1.6	0.9	-5.9	1069	1006
GBR	Validation	0.85	18.6	4.3	1.9	0.9	-5.4	329	311
HGBR	Train	0.93	23.0	4.8	1.0	1.0	-4.5	1069	1021
HGBR	Validation	0.91	12.1	3.5	1.6	1.0	-3.8	329	317
RFR	Train	0.49	174.3	13.2	3.6	0.6	-42.8	1069	611
RFR	Validation	0.53	60.8	7.8	3.3	0.6	-42.7	329	189
XGBR	Train	0.94	20.6	4.5	0.9	1.0	-4.7	1069	1019
XGBR	Validation	0.90	12.6	3.6	1.6	0.8	-15.9	329	277
XGBR_Advanced	Train	0.86	47.8	6.9	1.5	0.9	-9.4	1069	969
XGBR_Advanced	Validation	0.90	13.5	3.7	1.6	0.8	-17.8	329	271

ERA5 rainfall and temperature data. The models were assessed based on several metrics, including the coefficient of determination (R^2), mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and volumetric bias. The above performance metrics provide an overall assessment of the models' accuracy and bias, summarized in Table 4.

As indicated in Table 4, the Histogram Gradient Boosting Regressor (HGBR), achieved the highest validation performance of $R^2 = 0.91$, RMSE = 3.5 mm/month, and MAE = 1.6 mm/month, with the highest training performance $R^2 = 0.93$, which showed a little overfitting. The Extreme Gradient Boosting Regressor (XGBR) got the highest training performance of $R^2 = 0.95$, and its validation performance reached $R^2 = 0.90$, RMSE = 3.6 mm/month, and MAE = 1.6 mm/month. The XGBR based on the advanced configuration (XGBR_Advanced) got the highest training performance of $R^2 = 0.86$ for SM, although the validation performance ($R^2 = 0.895$) was a little worse than that based on the basic configuration. The validation performance of the Gradient Boosting Regressor (GBR) reached $R^2 = 0.86$ and RMSE = 4.3 mm/month. However, the Random Forest Regressor (RFR) only reached the validation performance of $R^2 = 0.53$, RMSE = 7.8 mm/month, and it had an obvious underestimation bias of −42.68%.

Scatter plots in Fig. 6 compare observed monthly rainfall against predictions from five different machine-learning models (GBR, HGBR, RFR, XGBR, XGBR_Advanced) to assess their performance. The scatter plots are compared to a 1:1 reference line, which represents ideal predictions. Scatter around this line indicates error of prediction, whereas close grouping around the ideal line indicates better predictive performance. Predictions for all models cluster at low rainfall values (< 40 mm). Thus, accuracy at low rainfall values is more critical in arid regions where most instances of rainfall occur as infrequent events at low magnitudes. However, all of the boosting models demonstrated excellent accuracy in predicting low rainfall amounts with tightly clustered predictions that exhibit a strong linear relationship with observations near the 1:1 line. A relatively large dispersion is observed for the predictions by the best model (RFR), with an R^2 value of 0.53. While it is very good for moderate rainfall amounts, the model clearly underestimates extreme events and displays considerable random variability, especially averaging to lower than actual values. This is characteristic of bagging methods applied to very high-dimensional, sparse or highly skewed datasets. On the other end

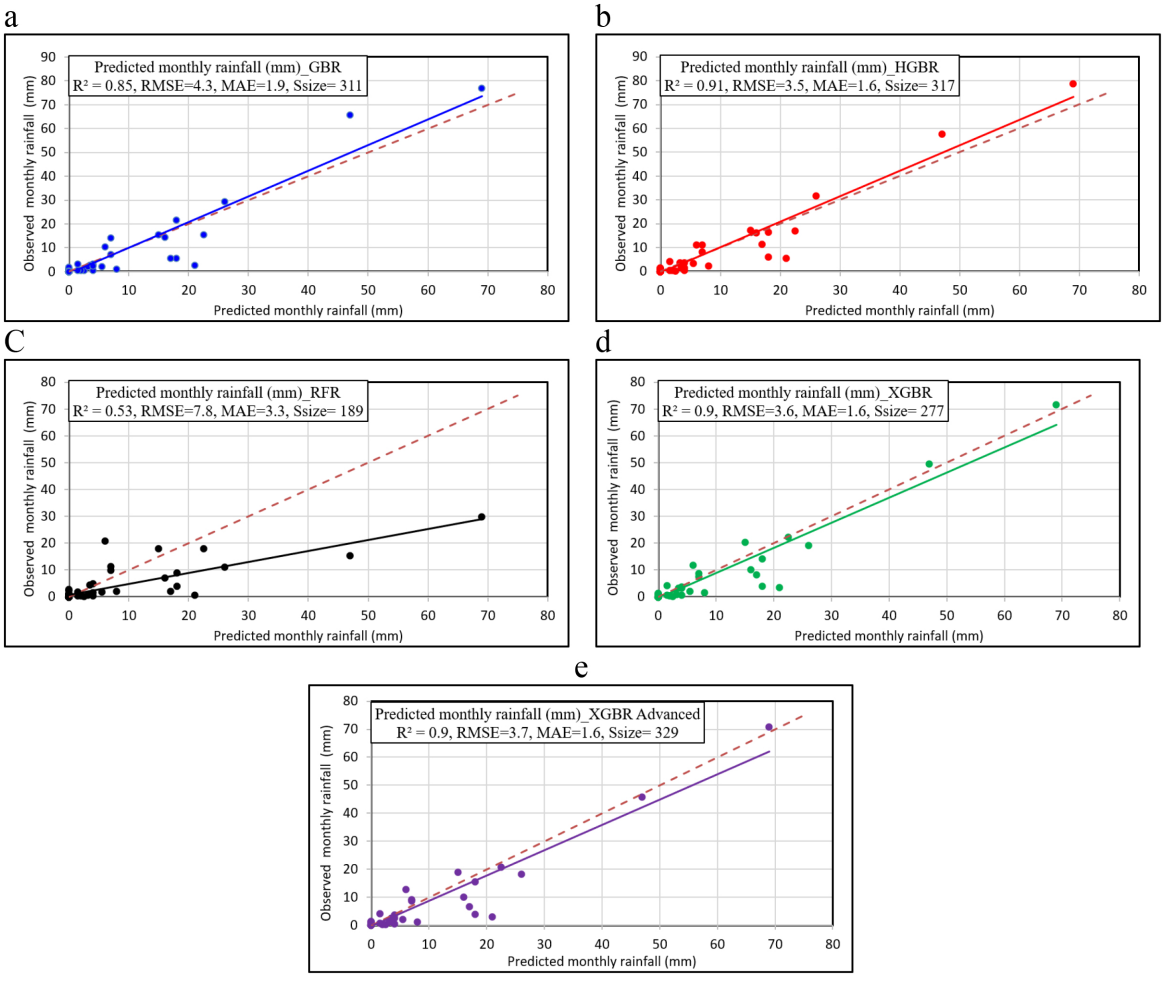


Fig. 6. Scatter plots comparing observed and validation monthly rainfall for five machine-learning models: (a) GBR, (b) HGBR, (c) RFR, (d) XGBR, and (e) XGBR_Advanced

of the spectrum, the XGBR model displays the best visual agreement with observations ($R^2 = 0.91$), with all predicted values showing a very small dispersion around the 1:1 line, and best of all models in reproducing moderate-to-high rainfall amounts. The XGBR_Advanced model follows closely and clearly outperforms all other models with an R^2 value of 0.9, showing a bit larger dispersion than XGBR. This slightly larger dispersion is presumably related to a somewhat reduced flexibility, because of the additional regularization or of the further tuning of some of the hyperparameters. From the perspective of a hydrologist, it is most important in arid basins to find a model that can realistically reproduce both frequent low-intensity events and occasional high-intensity storms. The scatter plots in the figures above show that all boosting models perform very well at predicting the frequent low rainfall amounts. However, only the XGBR and HGBR models can accurately be used to reproduce higher rainfall amounts, suggesting that these models better capture the complex, nonlinear processes driving the precipitation of convective events.

The best validation performance was obtained by HGBR and XGBR with $R^2 = 0.9$ for modeling monthly rainfall in an arid region. The intermittency, strong seasonality, and dominance of local convective events pose significant challenges to rainfall forecasting at a seasonal time scale using statistical models [46]. In this study, high predictive skill for monthly rainfall estimation in arid environments was achieved using advanced boosting-based machine-learning models. The two best-performing models, HGBR and XGBR, achieved the highest R^2 values for the validation data of 0.91 and 0.90, respectively. High validation R^2 values were also achieved in several recent studies in similar hydroclimatic regimes (Table 5). For example, a multivariate XGBoost model applied in a hyper-arid region of the UAE achieved a correlation coefficient of 0.76 [40], which is substantially lower than the corre-

Table 5. Comparative review of machine-learning models for rainfall prediction in arid and semi-arid regions.

Study	Location	Model(s)	Best R^2 / Correlation	Key Metrics
Current Study	Saudi Arabia	HGBR, XGBR	0.905, 0.902	MSE: Low
Baig et al., [40]	UAE (hyper-arid)	XGBoost (multivariate)	CC: 0.76	Stacked models: CC 0.82
Kumar et al., [41]	Urban (KSA affiliation)	CatBoost, XGBoost	0.99 (daily/weekly)	MAE: 0.00077, RMSE: 0.0010
Baljon & Sharma [27]	Saudi Arabia	Improved ML techniques	Not specified	Focus on prediction rate
Putra et al., [42]	Not specified	Gradient Boosting	0.740	RMSE: 1.765
Pakdaman et al., [43]	Southwest Asia	Random Forest	KGE: 0.68	Improved from -0.27
Assiri & Qureshi [44]	Saudi Arabia	Random Forest	KGE: 0.75	RMSE: 6.6 mm/mo
Wang et al., [45]	Saudi Arabia	ML ensemble (18 stacks)	KGE: 0.36	Best among products

lations obtained in the present study. This difference likely reflects both the inclusion of temporal memory predictors in the current modeling framework and differences in predictor structure, since stacked models by Baig et al., [40] improved correlation only to 0.82, still below the performance of HGBR and XGBR here.

Similarly, Putra et al., [42] reported a best R^2 of 0.74 using Gradient Boosting, again indicating that the present results represent a substantial improvement in predictive skill. The improvement may be attributed to the incorporation of lagged rainfall terms and nonlinear interaction features, which enhance representation of temporal persistence and nonlinear rainfall - temperature coupling. In contrast, studies using Random Forest approaches in arid regions have typically reported moderate performance. For instance, Pakdaman et al., [43] reported a Kling–Gupta Efficiency (KGE) of 0.68, while Assiri & Qureshi, [44] obtained KGE = 0.75 with RMSE = 6.6 mm/month. In the current analysis, the Random Forest regression model demonstrated only limited predictive capability (validation $R^2 = 0.527$). This finding is consistent with previous results indicating that bagging-based ensemble methods tend to perform less well than boosting methods for analyzing spatial, intermittent rainfall exhibiting highly nonlinear threshold behavior. In a recent multi-model ensemble study conducted over Saudi Arabia, the best possible KGE achieved was 0.36 using 18 stacked machine-learning models [45]. In contrast, higher performance was achieved using single boosted models, after optimal tuning, specifically in terms of selected predictors and feature engineering. It should be noted that extremely high correlations ($R^2 = 0.99$) reported by Kumar et al., [41] for CatBoost and XGBoost models were obtained at daily or weekly temporal scales, where autocorrelation and persistence effects are stronger. Monthly rainfall predictions are typically more difficult to perform due to aggregation, reducing the signal-to-noise ratio and increasing variability from that of daily predictions. Therefore, the R^2 of 0.90 obtained in this monthly study should be viewed as particularly robust and not comparable to that of monthly studies.

Comparative analysis shows that the proposed modeling framework attains state-of-the-art or improved performance for rainfall prediction in arid and hyper-arid basins. Boosting-based algorithms were shown to be particularly effective at reconstructing monthly rainfall variability by coupling well-designed predictors with physically constrained output components. The performance of the new approach suggests that it is efficacious for a variety of hydrological and water resource relevant applications, including hydrologic modeling, and flood hazard assessment. Additionally, the approach's ability to incorporate multiple types of data with little modification suggests that it can be efficiently applied in basins with sparse gauge networks.

4. Conclusion

New sets of gridded precipitation data and advances in machine-learning methods have recently been assessed for improved rainfall forecasting. The challenge of rainfall estimation in desert regions is significant due to the temporally and spatially variable nature of precipitation. The study presents an assessment and comparison of performance using a suite of techniques (simple regression, statistical post-processing, and machine-learning) for estimating rainfall in Qassim, central of the Kingdom of Saudi Arabia, using a range of satellite datasets. The findings show that such datasets and forecast scientific methods offer great potential for improving rainfall forecasts and enabling hydrological applications in arid regions lacking extensive documentary and monitoring datasets.

This comparative study evaluated the performance of a range of gridded rainfall products against independent rainfall observations. Many satellite products showed limited predictive capability at monthly scales; however, of the reanalysis-based products, ERA5 showed the greatest correspondence with ob-

servations, even though the raw ERA5 data tended to underestimate both average and extreme rainfall. Performance over arid and semi-arid regions is generally poor because convective and/or spatially variable events are difficult to represent in high-resolution dynamical models that serve as the basis for ERA5. Hence, for the development of useful rainfall estimates, it is necessary to apply bias-correction techniques to these large-scale data sets.

The inclusion of temperature as an additional predictor in multivariable rainfall nowcast models improved model performance. In multiregional diagnostics, the inclusion of temperature provided information on atmospheric moisture potential and convective instability, enabling improved representation of precipitation. This finding supports the use of physically relevant climatic variables in addition to or instead of rainfall when predicting rainfall, particularly for regions significantly influenced by thermodynamic atmospheric conditions. A comparison of the performance of various machine learning approaches was conducted to observe their predictive capability. Ensemble boosting strategies emerged as the clear winners in this comparison, as both Histogram Gradient Boosting Regression (HGBR) and Extreme Gradient Boosting Regression (XGBR) performed better than other models in terms of both R-Squared values and error metrics. Correspondingly, the bias improvement for both HGBR and XGBR better captured the complex, inherently nonlinear relationship between the spatial distribution of antecedent data and the resultant rainfall. In contrast, the performance of the Random Forest Regression (RFR) model was notably worse than the other models, likely due to the paucity of data and the need for more effective feature engineering and/or tuning. Consistent with other studies that have demonstrated that gradient-boosting methods frequently outperform other commonly used machine learning methods at environmental prediction, robust predictions of daily rainfall intensity were generated using XGBR. Importantly, a sophisticated XGBR configuration was crucial in maximising the accuracy of the predictions, with a thorough model optimization process, including hyperparameter tuning, required to achieve the best results.

The findings are in line with several regional and global studies performed across the globe. Several studies in arid and semi-arid regions have demonstrated that the use of ensemble machine-learning models outperforms traditional single-model approaches for forecasting, bias correction, and climate modeling of various variables, including rainfall. Furthermore, the improvements observed in this study by using a suite of satellite data and incorporating a range of additional predictors, along with bias-adjustment, are in line with some of the recent developments in hydrological modeling. The study contributes to the growing body of evidence supporting the potential of advanced data-driven approaches for precipitation modeling in complex and challenging climates.

From an applied perspective, the formalized approach can serve as a practical tool for rainfall estimation in ungauged basins. This is particularly relevant for water resource management, prediction, and flash-flood prediction studies in arid zones where ground meteorological networks are very sparse or incomplete. Despite these positive outcomes, there are some limitations that need to be considered. The effectiveness of the predictions made by machine learning depends on the quality, completeness, and timeliness of the input data. Even though the high-resolution gridded database and a lengthy observation period were applied, the development and validation of the models were performed using only the rainfall records collected from one location. This means that the performance evaluation might not fully reflect the spatial distribution of rainfall within the broader area, especially for the arid regions, where precipitation tends to be extremely localized. Furthermore, the validation and testing were carried out solely in the same climatic zone, which could potentially hinder the generalization of the proposed method outside this specific environment.

In the concluding remarks, this study argues that the utilization of gridded climate data and machine learning algorithms provides a relevant approach for rain forecasting in arid environments. The results also demonstrate that the boosting-based algorithms were the most accurate to be tested, especially HGBR and XGBR. The use of methods to correct bias, the incorporation of multiple predictors, and ensemble learning techniques are feasible and promising approaches for rainfall forecasting, notably in areas of limited data, such as the arid region of this study. Furthermore, the use of machine learning methods as a complement to other traditional methods for hydrological purposes may also provide interesting results for the fields of climate and water resources, mainly in arid and semi-arid environments.

Statements and declarations

Funding

The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests

The author has no relevant financial or non-financial interests to disclose.

Conflict of interest

The author declares no conflict of interest.

Author contributions

As the sole author of this study, I was responsible for all stages of the research. This included developing the research ideas, designing the methodology, collecting and processing the rainfall datasets, building and evaluating the machine learning models, and interpreting the results. I also prepared all figures and tables, reviewed the findings in light of previous studies, and wrote, revised, and finalized the manuscript.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Hasanean H, Almazroui M. Rainfall: Features and variations over Saudi Arabia, a review. Climate. 2015;3:578–626.
2. Nashwan MS, Shahid S, Dewan A, Ismail T, Alias N. Performance of five high resolution satellite-based precipitation products in arid region of Egypt: An evaluation. Atmos Res. 2020;236:104809.
3. Al-Falahi AH, Saddique N, Spank U, Gebrechorkos SH, Bernhofer C. Evaluation the performance of several gridded precipitation products over the highland region of yemen for water resources management. Remote Sens (Basel). 2020;12:2984.
4. Abbas A, Yang Y, Pan M, Tramblay Y, Shen C, Ji H, *et al.* Comprehensive global assessment of 23 gridded precipitation datasets across 16,295 catchments using hydrological modeling. EGUsphere. 2025;2025:1–31.
5. Beck HE, Vergopolan N, Pan M, Levizzani V, Van Dijk AIJM, Weedon GP, *et al.* Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol Earth Syst Sci. 2017;21:6201–17.
6. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, *et al.* The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Sci Data. 2015;2:1–21.
7. Helmi AM, Abdelhamed MS. Evaluation of CMORPH, PERSIANN-CDR, CHIRPS V2. 0, TMPA 3B42 V7, and GPM IMERG V6 satellite precipitation datasets in Arabian arid regions. Water (Basel). 2022;15:92.
8. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, *et al.* The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 2020;146:1999–2049.
9. Huffman GJ, Bolvin DT, Braithwaite D, Hsu K, Joyce R, Xie P, *et al.* NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD) Version. 2015;4:30.
10. Mahmoud MT, Mohammed SA, Hamouda MA, Mohamed MM. Impact of topography and rainfall intensity on the accuracy of imerg precipitation estimates in an arid region. Remote Sens (Basel). 2020;13:13.
11. Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, *et al.* The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol. 2007;8:38–55.
12. Tekeli AE, Fouli H. Evaluation of TRMM satellite-based precipitation indexes for flood forecasting over Riyadh City, Saudi Arabia. J Hydrol (Amst). 2016;541:471–9.
13. Saha S, Moorthi S, Pan H-L, Wu X, Wang J, Nadiga S, *et al.* The NCEP climate forecast system reanalysis. Bull Am Meteorol Soc. 2010;91:1015–58.
14. Abatzoglou JT, Dobrowski SZ, Parks SA, Hegewisch KC. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci Data. 2018;5:1–12.
15. Ashouri H, Hsu K-L, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, *et al.* PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull Am Meteorol Soc. 2015;96:69–83.
16. Sneha MR, Nair A. Evaluation of CMORPH satellite precipitation product for modelling rainfall over South Peninsular region of India. in 2nd International Conference On Computational Sciences-Modelling, Computing And Soft Computing (CSMCS 2022), vol. 2875, AIP Publishing LLC; 2023, p. 020009.
17. Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, *et al.* The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol. 2007;8:38–55.
18. Prakash S. Performance assessment of CHIRPS, MSWEP, SM2RAIN-CCI, and TMPA precipitation products across India. J Hydrol (Amst). 2019;571:50–9.
19. Najmaddin PM, Whelan MJ, Balzter H. Application of satellite-based precipitation estimates to rainfall-runoff modelling in a data-scarce semi-arid catchment. Climate 2017;5:32.
20. Mahmoud MT, Al-Zahrani MA, Sharif HO. Assessment of global precipitation measurement satellite products over Saudi Arabia. J Hydrol (Amst) 2018;559:1–12.
21. Breiman L. Random forests. Mach Learn. 2001;45:5–32.
22. Tyralis H, Papacharalampous G, Langousis A. A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water (Basel). 2019;11:910.
23. Slamet C, Imron RM, Wahana A, Maylawati DS, Zulfikar WB, Ramdhani MA. Feature engineering for tropical rainfall forecasting using random forest and support vector regression. Journal of Applied Data Sciences. 2026;7:715–28.
24. Kumar B, Kumari P, Kumar A, Prasad SK. Predicting daily rainfall amount using random forest, multivariate linear regression, and extreme gradient boosting. in 2025 10th International Conference on Signal Processing and Communication (ICSC), IEEE; 2025, p. 535–40.
25. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, *et al.* Lightgbm: A highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst. 2017;30.
26. Chen T. XGBoost: A Scalable Tree Boosting System. Cornell University 2016.

27. Baljon M, Sharma SK. Rainfall prediction rate in Saudi Arabia using improved machine learning techniques. Water (Basel). 2023;15:826.
28. Salih W, Epule TE, EL Khalki ELM, Ouatiki H, Erraki S, Achli S, *et al.* A comprehensive assessment of satellite precipitation products over a semi-arid region: focus on extreme events. Natural Hazards. 2024;120:3037–65.
29. Najmi A, Igmoullan B, Namous M, El Bouazzaoui I, Brahim YA, El Khalki EM, *et al.* Evaluation of PERSIANN-CCS-CDR, ERA5, and SM2RAIN-ASCAT rainfall products for rainfall and drought assessment in a semi-arid watershed, Morocco. Journal of Water and Climate Change. 2023;14: 1569–84.
30. Maghsood FF, Hashemi H, Hosseini SH, Berndtsson R. Ground validation of GPM IMERG precipitation products over Iran. Remote Sens (Basel). 2019;12:48.
31. Nguyen P, Ombadi M, Sorooshian S, Hsu K, AghaKouchak A, Braithwaite D, *et al.* The PERSIANN family of global satellite precipitation data: A review and evaluation of products. Hydrol Earth Syst Sci. 2018;22:5801–16.
32. Ashouri H, Hsu K-L, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, *et al.* PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bull Am Meteorol Soc. 2015;96:69–83.
33. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, *et al.* The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Sci Data. 2015;2:150066. https://doi.org/10.1038/sdata.2015.66.
34. Alijanian M, Rakhshandehroo GR, Mishra AK, Dehghani M. Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. International Journal of Climatology. 2017;37:4896–914.
35. Wolters ELA, Van Den Hurk B, Roebeling RA. Evaluation of rainfall retrievals from SEVIRI reflectances over West Africa using TRMM-PR and CMORPH. Hydrol Earth Syst Sci. 2011;15:437–51.
36. Dinku T, Funk C, Peterson P, Maidment R, Tadesse T, Gadain H, *et al.* Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society. 2018;144:292–312.
37. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, *et al.* The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Sci Data. 2015;2:1–21.
38. Joyce RJ, Janowiak JE, Arkin PA, Xie P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J Hydrometeorol. 2004;5:487–503.
39. Satgé F, Hussain Y, Molina-Carpio J, Pillco R, Laugner C, Akhter G, *et al.* Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. International Journal of Climatology. 2020;41.
40. Baig F, Ali L, Faiz MA, Chen H, Sherif M. How accurate are the machine learning models in improving monthly rainfall prediction in hyper arid environment? J Hydrol (Amst). 2024;633:131040.
41. Kumar V, Kedam N, Sharma KV, Khedher KM, Alluqmani AE. A comparison of machine learning models for predicting rainfall in urban metropolitan cities. Sustainability. 2023;15:13724.
42. Putra M, Rosid MS, Handoko D. Rainfall estimation using machine learning approaches with raingauge, radar, and satellite data. in 2022 International conference on electrical engineering and informatics (ICELTICs), IEEE; 2022, p. 25–30.
43. Pakdaman M, Babaeian I, Bouwer LM. Improved monthly and seasonal multi-model ensemble precipitation forecasts in southwest Asia using machine learning algorithms. Water (Basel). 2022;14:2632.
44. Assiri ME, Qureshi S. A multi-source data fusion method to improve the accuracy of precipitation products: A machine learning algorithm. Remote Sens (Basel). 2022;14:6389.
45. Wang X, Alharbi RS, Baez-Villanueva OM, Green A, McCabe MF, Wada Y, *et al.* Saudi Rainfall (SaRa): hourly 0.1° gridded rainfall (1979–present) for Saudi Arabia via machine learning fusion of satellite and model data. Hydrol Earth Syst Sci. 2025;29:4983–5003.
46. Elmotawakkil A, Moumane A, Sadiki A, Zahi A, Al Karkouri J, Batchi M, *et al.* Forecasting short-term rainfall patterns in arid and semi-arid regions using machine learning and deep learning models: a case study from Morocco. Theor Appl Climatol. 2025;156:436.