OPEN ACCESS
International Journal of Innovations in Science & Technology

Validation of Satellite-Based Gridded Rainfall Products with Station Data Over Major Cities in Punjab

Syeda Nimra Raza Geelani [1], Sawaid Abbas[1, 2, *], Muhammad Umar[1], Muhammad Usman[3], Irum Yousfani[4,5]

[1]Smart Sensing for Climate and Development, Center for Geographical Information System, University of the Punjab, Lahore, Pakistan

[2]Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong SAR

[3]Interdisciplinary Research Center for Aviation and Space Exploration, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia

[4]Department of Computer Science, Asian Institute of Technology, Thailand

[5]Bell Media Canada, Toronto, Canada

***Correspondence:** sawaid.gis@pu.edu.pk (S.A); sawaid.abbas@connect.polyu.hk (S.A)

Citation | Geelani. S. N. R, Abbas. S, Umar. M, Usman. M, Yousfani. I, "Validation of Satellite-Based Gridded Rainfall Products with Station Data Over Major Cities in Punjab", IJIST, Special Issue pp 305-318, June 2024

Received | June 04, 2024; **Revised |** June 08, 2024; **Accepted |** June 12, 2024; **Published |** June 20, 2024.

A Critical evaluation of newly developed gridded rainfall datasets is essential for their effective application. Over the past two decades, the availability of gridded rainfall measurements has increased; however, finding suitable proxies for traditional station-based measurements remains challenging. This study conducted a comparative assessment of rainfall estimates from IMERG, CHIRPS, ERA-5, and APHRODITE against meteorological station data from five cities in Pakistan: Lahore, Faisalabad, Multan, Islamabad, and Murree. The assessment covered multiple temporal scales (daily, monthly, and yearly) using daily data recorded from 2001 to 2022. Analytical metrics applied included Bias, Mean Error (ME), Root Mean Square Error (RMSE), Correlation Coefficient (CC), and Coefficient of Determination (R^2). The results revealed notable spatial and temporal patterns of agreement among the datasets. Correlations for daily data were generally weak across all gridded datasets, with APHRODITE performing the best. Monthly aggregates showed that IMERG had the highest association with ground data, followed by CHIRPS. Yearly accumulated rainfall records indicated that IMERG had the highest correlation, followed by CHIRPS. Overall, IMERG demonstrated higher consistency across stations at both monthly and yearly scales. CHIRPS exhibited lower errors (RMSE and bias) at most locations, especially Lahore, but showed higher errors in Murree at the monthly scale. The study concludes that a single satellite dataset alone may not provide sufficient accuracy over large areas; a combination of products may be required for better estimation.

Keywords: Rainfall; CHIRPS; IMERG; ERA-5; APHRODITE, Punjab.

Introduction:

The global water cycle is significantly influenced by precipitation, making it a crucial element with substantial impacts on the socio-economic development of various regions [1]. In recent decades, the scientific community has increasingly focused on understanding climate variability and change [2][3], with precipitation data playing a pivotal role in studying both regional and global climate changes [4][5]. Precipitation is essential for agricultural growth, food security, water management, and hydropower generation. It also plays a crucial role in identifying weather and climatic conditions [6][7][8], and is a key variable for assessing extreme climate events such as droughts, which can adversely affect crop production [9][10][11]. A significant challenge in such studies is the lack of long-term climate data. Despite some available data sources, observed data from ground stations in Pakistan, including Punjab province, are very limited [12][13], and the quality of this data is often insufficient for comprehensive hydro-climatological assessments. However, technological advancements have enabled hydroclimatic scientists to estimate climate variables and produce reliable data over longer periods, particularly in regions with limited ground data [14][15]. High-resolution remote sensing data, collected from satellites or aircraft, appear to be an optimal solution for addressing the shortage of ground station data [12][13]. Supplementing gauge-based measurements with satellite data has proven effective in overcoming the limitations of ground data availability [16][17][18], demonstrating significant effectiveness and cost-efficiency in recent years [19][20][21].

Recent studies suggest a declining annual trend in central Asia over the last few decades [22][23], while a slightly increasing trend in annual, seasonal, and monthly rainfall has been reported in the northwest Himalayan region from 2000 to 2022. In Pakistan, several studies have identified uneven rainfall patterns from 1980 to 2020. For example, [24] noted an increasing rainfall trend in Punjab province using various remotely sensed gridded precipitation datasets. Other studies have observed an overall increase in rainfall at annual and seasonal scales over the past four decades [22][25][26], with a declining trend in winter and post-monsoon periods but an increasing trend during monsoon and pre-monsoon periods from 1980 to 2016 [27]. Additionally, [28] evaluated the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) dataset against the Surface Precipitation Gauge (SPG) dataset for the period 1981-2018 in Pakistan. Statistical metrics such as Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Correlation (CC) were used to assess the datasets at daily, monthly, annual, and seasonal scales. On average, CHIRPS showed an underestimation of rainfall by -0.80 to -0.17 mm/day across 46 stations. The highest ME, MAE, and RMSE values were 2.0 mm/day, 1.8 mm/day, and 1.8 mm/day, respectively, at the yearly scale. CHIRPS demonstrated high accuracy at monthly and seasonal timescales and has been found to be a useful alternative to station data for understanding spatial patterns and temporal trends in precipitation in regions of Pakistan with limited station coverage [13].

Objectives:

This study evaluates the performance of four rainfall products—IMERG (Integrated Multi-satellite Retrievals for the Global Precipitation Measurement), CHIRPS (Climate Hazards Group InfraRed Precipitation), ERA-5 (ECMWF Fifth Generation Atmospheric Analysis of Global Climate Coverage), and APHRODITE (Asian Precipitation-Highly-Resolved Observation Data Integration Towards Evaluation of Water Resources)—against climate station data to address the limitations of spatially contiguous rainfall data availability in Pakistan. The evaluation was conducted using gridded rainfall datasets validated across five cities in Punjab: Lahore, Faisalabad, Multan, Islamabad, and Murree. Analytical metrics employed in this study included Bias, Mean Error (ME), Root Mean Square Error (RMSE), Correlation Coefficient (CC), Coefficient of Determination (R^2), and Probability Value of Correlation (p). To provide a comprehensive assessment of the accuracy and reliability of these rainfall estimates, the analysis covered three timescales (daily, monthly, and annual) using daily data recorded from 2001 to 2022.

Novelty Statement:

This study conducted a comparative assessment of rainfall estimates from IMERG, CHIRPS, ERA-5, and APHRODITE using meteorological station data from five cities in Pakistan. This analysis aims to identify suitable proxies for conventional station-based measurements, offering a solution to address the limitations of traditional rainfall data sources.

Material and Methods:

A brief overview of the methodology is illustrated in the flow diagram (Figure 1).

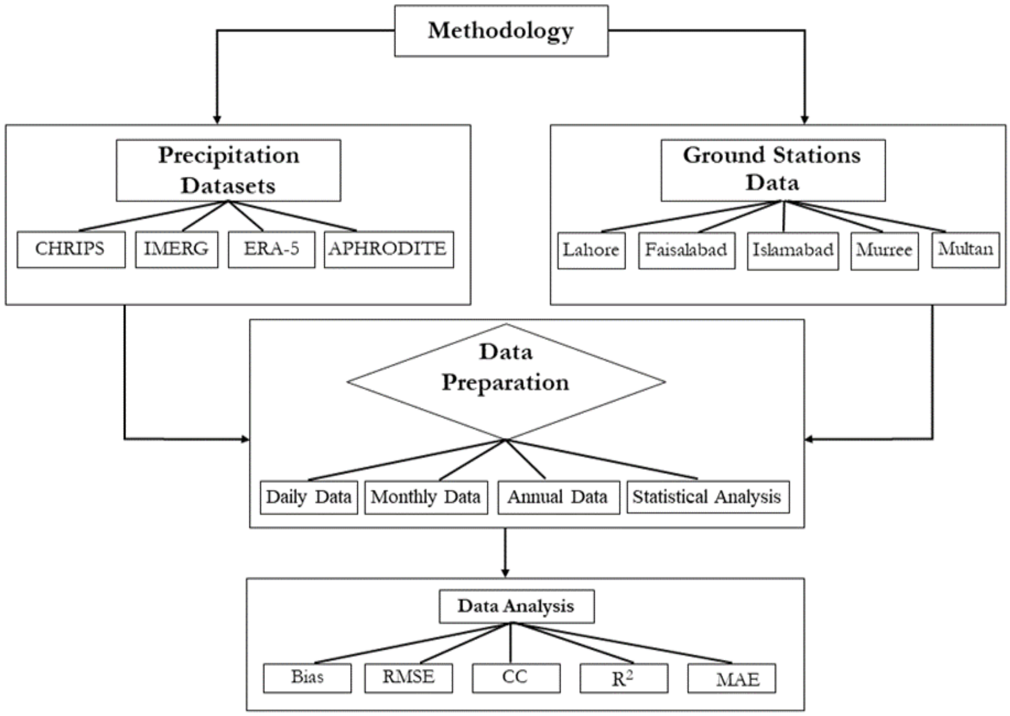

Figure 1: An overview of the methodology applied in this study

Description of Study Area and Location of Ground Stations:

Pakistan, located in Southwest Asia, covers an area of 881,913 km², situated between latitudes 24° to 37° N and longitudes 60° to 75° E (Figure 2). The country features diverse topography, ranging from the Karakoram and Himalayan mountains in the north and northwest to agricultural plains in the Indus Basin and the southern coast of the Arabian Sea [29]. As an agrarian nation, Pakistan's largest province, Punjab, is notable for its significant agricultural production, contributing over 50% of the country's agricultural commodities [30]. The regional average temperature ranges from 23°C to 26°C, with minimum and maximum temperatures between 16°C to 19°C and 29°C to 33°C, respectively. Rainfall varies significantly, with the northern part of Punjab receiving more precipitation than the south; total annual rainfall ranges from less than 300 mm in the south to over 800 mm in the north [31]. The region experiences two main monsoon seasons: the summer monsoon (June–September) and the winter monsoon (December–April). Maximum rainfall occurs during the summer monsoon due to the monsoon system originating from the Bay of Bengal, entering from the east and northeast. Winter precipitation, driven by western disturbances from the Mediterranean Sea, affects the region from the southwest and northeast [32]. Extreme variations in rainfall during these monsoon seasons contribute to the region's vulnerability to climate change [33].

Ground station selection for this study was based on data from the Pakistan Meteorological Department (PMD), ensuring coverage of key urban centers and one hill station in Punjab and the Federal Capital Territory. The selected cities—Lahore (31.52°N, 74.36°E), Multan (30.1864°N, 71.4886°E), Faisalabad (31.4504°N, 73.1350°E), Islamabad (33.6995°N, 73.0363°E), and Murree (37°N, 73.3943°E)—reflect the region's geographical and climatic diversity, each contributing significantly to the province's socio-economic fabric (Table 1). According to the PMD, Lahore has moderate to high rainfall, averaging over 617 mm annually, with a range of 332 to 903 mm over the past 22 years. Faisalabad averages 615 mm annually, while Multan's rainfall ranged from 60 to 395 mm per year from 2001 to 2022. Islamabad shows considerable variability, with annual rainfall between 580 and 1796 mm over the past 22 years. Murree, known for its scenic beauty, recorded an annual rainfall range of 1136 to 2420 mm from 2001 to 2022.

Table 1: Description of the ground station data acquired from the Pakistan Meteorological Department

Station	Longitude	Latitude	Elevation (m)	Period
Lahore	74.33	31.55	213	2001-2022
Faisalabad	73.1	31.43	184	2001-2022
Islamabad	73.13	33.67	495	2001-2022
Murree	73.38	33.92	2290	2001-2022
Multan	71.43	30.2	122	2001-2022

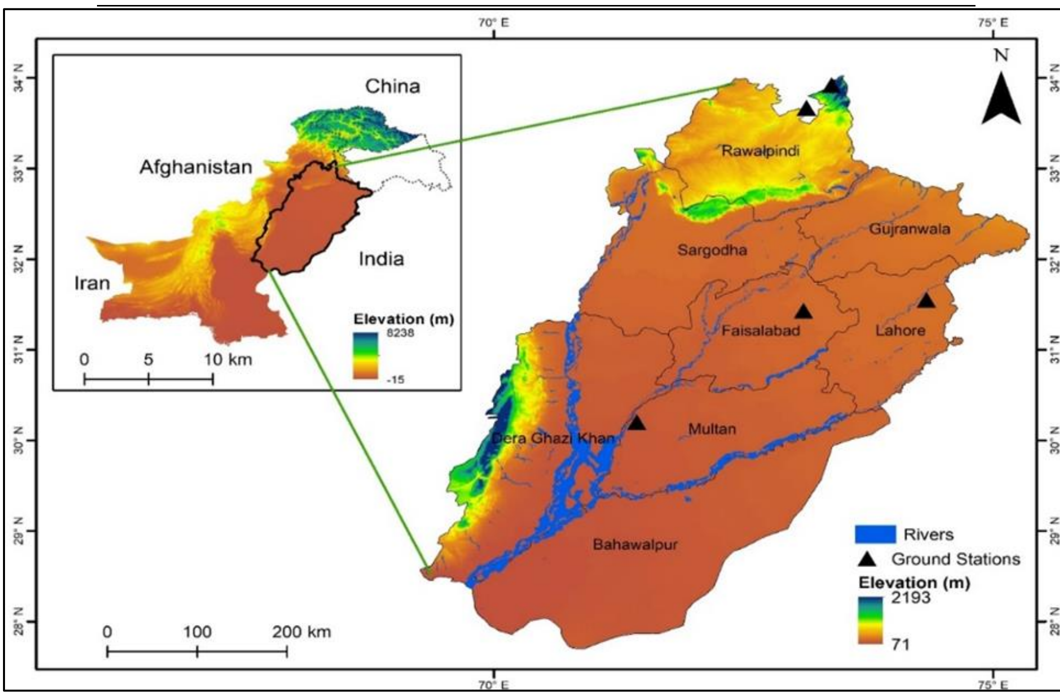

Figure 2: Topographic settings and location of the ground stations

Gridded Rainfall Measurements:

Gridded rainfall measurements were obtained from four sources: CHIRPS [9], IMERG [34], APHRODITE [35], and ERA-5 [36], covering the period from 2001 to 2022. A brief description of these datasets is provided below and summarized in Table 2.

CHIRPS:

The CHIRPS (Climate Hazards Group InfraRed Precipitation with Station Data) product offers high spatial resolution (0.05°) and extends from 1983 to the present. CHIRPS provides temporal resolutions including daily, monthly, pentad, and annual composites [20][37][38]. For this study, daily, monthly, and annual gridded data from 2001 to 2022 were utilized. Data can be accessed at CHIRPS website.

IMERG:

IMERG (Integrated Multi-satellite Retrievals for the Global Precipitation Measurement) version V06 provides rainfall estimates at a spatial resolution of 0.1° × 0.1° since 1997, with temporal aggregates of 30 minutes, 1 day, and 1 month. In this study, IMERG-Final v6 data were retrieved from NASA's GPM site.

ERA-5:

ERA-5 (Fifth Generation ECMWF Atmospheric Reanalysis of the Global Climate) is the latest global atmospheric reanalysis product from ECMWF. For this study, daily data with a spatial resolution of 0.25° × 0.25° were downloaded from Climate Data Guide and converted into monthly totals and annual composites for analysis.

APHRODITE:

APHRODITE (Asian Precipitation - Highly Resolved Observational Data Integration Towards Evaluation of Water Resources) provides high-resolution (0.25°) gridded daily rainfall datasets for Asia. Developed by the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of the Japan Meteorological Agency (MRI/JMA) since 2006 [35], APHRODITE data are available from 1961. Access the dataset at APHRODITE website and APHRODITE product page.

Table 2: Description of gridded rainfall measurement data used in the study

Datatype	Temporal Scale	Source	Year	Spatial Resolution
CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data)	Daily, 5Day, 10Day, Monthly	UCSB/ USGS	1983 - Present	0.05°×0.05°
IMERG (Integrated Multi-satellite Retrievals for the Global Precipitation Measurement)	30min	NASA	1997-2015 & 2014-Present	0.1° × 0.1°
APHRODITE (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources)	Daily	JAXA /RIHN	1951-Present & 1961-Present	0.25°×0.25° & 0.05°×0.05°
ERA-5 (Fifth generation ECMWF atmospheric reanalysis of the Global Climate Covering)	Hourly, Daily, Monthly	ECMWF	1950 - Present	0.125° × 0.125°

Data Analysis:

A comprehensive analysis was conducted to examine the relationships between the CHIRPS, IMERG, ERA-5, and APHRODITE datasets and the rainfall data recorded at the stations. The statistical indicators used for this assessment included bias, mean absolute error (MAE), root mean square error (RMSE), correlation coefficient (CC), coefficient of determination (R^2), and the probability value of correlation (p).

Bias:

Bias (Eq 1) represents the average difference between the datasets, which can be either positive or negative. A negative bias indicates an underestimation of rainfall, while a positive bias signifies an overestimation.

$$\text{Bias} = \frac{\sum_{i=1}^{N}(\text{PSi} - \text{POi})}{N} \qquad \text{(Eq 1)}$$

Mean Absolute Error (MAE):

The Mean Absolute Error (MAE) quantifies the magnitude of errors in the datasets, as represented by Eq (2).

$$MAE=\frac{\sum_{i=1}^{N}|PS_i-PO_i|}{N} \quad \text{(Eq 2)}$$

Root Mean Square Error (RMSE):

The Root Mean Square Error (RMSE) highlights larger errors more significantly than the MAE and is used to assess the average magnitude of errors, as defined by Eq (3). It is calculated as follows:

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(PS_i-PO_i)^2} \quad \text{(Eq3)}$$

Correlation Coefficient (CC):

The correlation coefficient (CC) measures the consistency between two observations, as given by Eq (4). Ranging from -1 to +1, it indicates high negative or positive associations, with values close to zero reflecting a weak relationship. The formula for CC is:

$$CC=\frac{\sum_{i=1}^{N}(P_{Si}-\bar{P}_S)(P_{Oi}-\bar{P}_O)}{\sqrt{\sum_{i=1}^{N}(P_{Si}-\bar{P}_S)2}\sqrt{\sum_{i=1}^{N}(P_{Oi}-\bar{P}_O)2}} \quad \text{(Eq 4)}$$

Results:

In this study, daily, monthly, and annual precipitation datasets from CHIRPS, ERA-5, IMERG, and APHRODITE were assessed against rain-gauge observations over a 22-year period (2001–2022) across five major cities: Lahore, Faisalabad, Multan, Islamabad, and Murree.

Table 3: Statistical indicators for the association of the datasets at the daily timescale

Ground Station	Data	Bias	MAE	RMSE	CC	R^2
Lahore	CHIRPS	1.000801	0.00147	0.131799	0.206	0.04257
	ERA-5	0.92401	-0.1395	12.50495	0.289	0.08357
	IMERG	2.21794	2.235785	200.4241	0.306	0.09383
	APHRODITE	0.837535	-0.28334	20.97266	0.626	0.39305
Faisalabad	CHIRPS	0.9056	-0.11032	9.889782	0.174	0.03029
	ERA-5	1.042978	0.050227	4.502556	0.246	0.06068
	IMERG	2.548662	1.809883	162.2446	0.26199	0.068641
	APHRODITE	1.130008	0.153446	11.35814	0.402	0.161348
Islamabad	CHIRPS	0.981185	-0.06415	5.750653	0.261	0.068275
	ERA-5	0.65691	-1.16979	104.8644	0.357	0.126884
	IMERG	1.950067	3.23932	290.3848	0.261	0.06852
	APHRODITE	1.237686	0.80124	59.30799	0.1517	0.023014
Murree	CHIRPS	1.23966	0.998514	89.51049	0.184	0.033864
	ERA-5	0.934577	-0.27258	24.43484	0.376	0.141287
	IMERG	1.715339	2.980372	267.1717	0.217	0.047237
	APHRODITE	1.103078	0.439234	32.51222	0.562	0.315611
Multan	CHIRPS	0.917942	-0.05106	4.57694	0.164	0.027194
	ERA-5	1.142331	0.088559	7.938722	0.224	0.050183
	IMERG	2.968035	1.224517	109.7703	0.21945	0.048516
	APHRODITE	1.713999	0.440873	32.63354	0.4392	0.192869

Association of the Data at a Daily Timescale:

The daily data from rain gauge stations were compared with rainfall estimates from CHIRPS, IMERG, ERA-5, and APHRODITE. The association between these datasets was assessed using metrics including Bias, MAE, RMSE, R^2, and p (Table 3). Overall, APHRODITE demonstrated the highest correlations, with the strongest correlation observed at Lahore station (0.63), followed by Murree (0.56). IMERG showed a high bias across all stations (Table 3, Figure 3). For Lahore, correlations with CHIRPS, ERA-5, and IMERG were weak, ranging from 0.2 to 0.35. In Murree, which is geographically closer to Islamabad, APHRODITE provided a slightly better correlation, just above 0.5, while other datasets had weaker correlations ranging from 0.1 to 0.37. CHIRPS slightly overestimated daily precipitation at Lahore (MAE = 0.00147), whereas APHRODITE and ERA-5 both underestimated (MAE = -0.28 and MAE = -0.139). At Multan, CHIRPS underestimated but was close to the mean of ground station data, while ERA-5 and APHRODITE overestimated (0.088 and 0.44, respectively). IMERG had the highest mean error. In Lahore, CHIRPS had a low RMSE, indicating higher accuracy, followed by ERA-5 and APHRODITE with RMSE values of 12.5 and 20.9, respectively. In Islamabad, only CHIRPS had a low RMSE, consistent with its good correlation. All datasets showed significant dispersion at Murree, with ERA-5 and APHRODITE showing relatively lower accuracy.

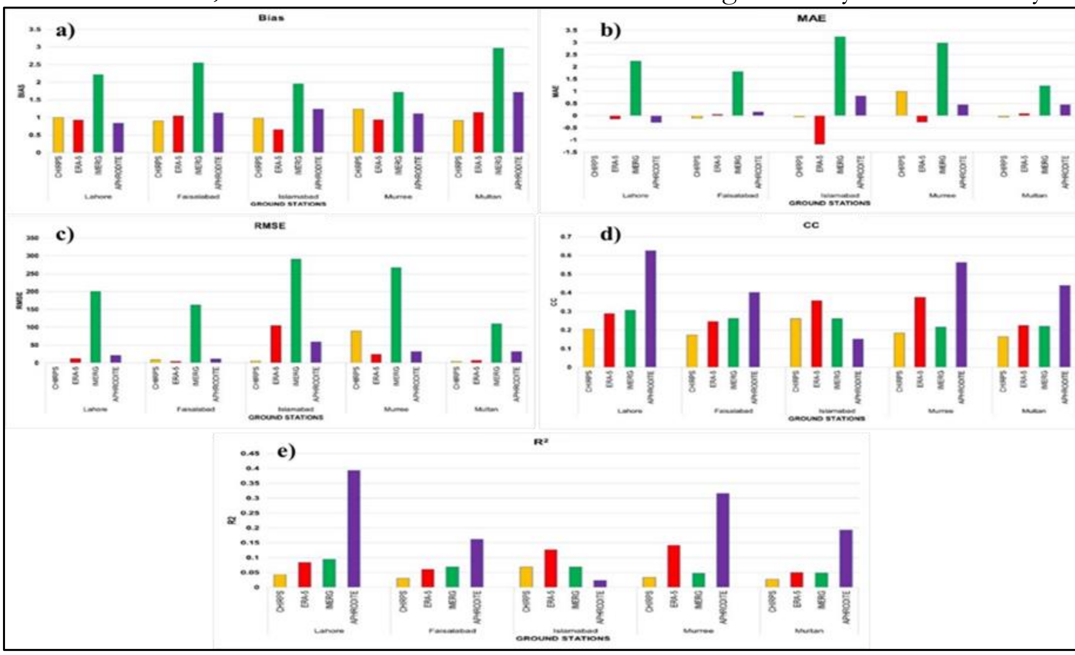

Figure 3: Statistical indicators showing the relationship between the ground station data and daily gridded precipitation datasets a) Bias, b) MAE, c) RMSE, d) CC, and e) R2 with correspondence observatories from 2001–2022.

Association of the Data at a Monthly Timescale:

All datasets exhibited strong correlations at the monthly timescale (Table 4, Figure 4). IMERG showed the highest correlation with monthly aggregates, particularly at Lahore station, where the correlation reached 0.93, followed by CHIRPS (0.90), APHRODITE (0.89), and ERA-5 (0.73). At Faisalabad station, IMERG also had the highest correlation with ground station data for monthly aggregates (0.90), followed by CHIRPS (0.80). Similarly, at Islamabad station, IMERG achieved the highest correlation (0.80). Regarding MAE, CHIRPS recorded values lower than the mean of ground station data for Multan, Islamabad, and Faisalabad, ranging from -1 to -3, indicating proximity to the ground station's monthly aggregates. Notably, a very low MAE of overestimation was observed for Lahore, while Murree exhibited a higher overestimation. ERA-5 showed underestimation at Lahore, Islamabad, and Murree, with the most significant mean underestimated error in Islamabad, exceeding -35. CHIRPS demonstrated very low RMSE values, less than 1 for Lahore, indicating high accuracy in precipitation estimation, though it showed considerable dispersion for Murree, resulting in the lowest accuracy among all cities. Despite providing the strongest monthly precipitation correlations, IMERG's high RMSE is attributed to outliers and significant differences between sub-datasets of ground stations and precipitation datasets on specific dates. APHRODITE, despite a lower correlation compared to IMERG, showed high RMSE values across all ground stations, ranging from 115.3 to 178.8, suggesting lower accuracy compared to CHIRPS and ERA-5. Although APHRODITE had a lower correlation than IMERG, its lower RMSE was due to less variance between APHRODITE sub-datasets and ground station data compared to IMERG.

Table 4: Statistical indicators for the association of the datasets at a monthly timescale

Ground Station	Data	Bias	MAE	RMSE	CC	R^2
Lahore	CHIRPS	1.000801	0.044585	0.72579	0.9	0.8102
	ERA-5	0.92401	-4.23015	68.86191	0.729	0.5317
	IMERG	2.21794	67.79912	1103.69	0.928	0.848
	APHRODITE	0.837535	-8.5768	115.3891	0.8897	0.7915
Faisalabad	CHIRPS	0.9056	-3.3455	54.46076	0.8187	0.6703
	ERA-5	1.042978	1.523117	24.79455	0.6686	0.4471
	IMERG	2.548662	54.88383	893.4441	0.9051	0.8195
	APHRODITE	0.7781	-7.86403	128.0171	0.5697	0.3246
Islamabad	CHIRPS	0.981185	-1.94532	31.66753	0.8427	0.7101
	ERA-5	0.65691	-35.4733	577.4641	0.6467	0.4182
	IMERG	1.950067	98.23085	1599.082	0.8526	0.727
	APHRODITE	1.237686	24.25411	326.3056	0.3789	0.1383
Murree	CHIRPS	1.23966	30.27945	492.9138	0.8115	0.6586
	ERA-5	0.934577	-8.26577	134.557	0.7579	0.5745
	IMERG	1.715339	90.37837	1471.253	0.857	0.7352
	APHRODITE	1.103078	13.29593	178.8784	0.6789	0.461
Multan	CHIRPS	0.918732	-1.53205	24.94001	0.7903	0.6246
	ERA-5	1.143314	2.701724	43.98087	0.6871	0.4721
	IMERG	2.97059	37.14912	604.7438	0.8778	0.7704
	APHRODITE	1.71618	13.3693	179.8655	0.5841	0.3412

Table 5: Statistical indicators for the association of the datasets at an annual timescale

Ground Station	Data	Bias	MAE	RMSE	CC	R^2
Lahore	CHIRPS	1.000801	0.513696	2.463598	0.7835	0.6139
	ERA-5	0.92401	-48.7387	233.7427	0.1342	0.018
	IMERG	2.21794	781.1637	3746.33	0.8178	0.6688
	APHRODITE	0.818419	-97.0251	388.1004	0.0316	0.778553
Faisalabad	CHIRPS	0.9056	-38.546	184.8599	0.6265	0.3925
	ERA-5	1.042978	17.54896	84.16183	0.26	0.0676
	IMERG	2.548662	632.3572	3032.679	0.6855	0.4698
	APHRODITE	1.13	52.54579	210.1832	0.3087	0.0953
Islamabad	CHIRPS	0.984776	18.07	86.66068	0.736	0.5417
	ERA-5	0.659314	404.371	1939.295	0.2302	0.0553
	IMERG	1.957203	1136.134	5448.706	0.687	0.472
	APHRODITE	1.244417	280.6183	1122.473	0.4459	0.1989
Murree	CHIRPS	1.23966	348.872	1673.131	0.6139	0.3769
	ERA-5	0.934577	-95.2361	456.7362	0.3477	0.1209
	IMERG	1.715339	1041.316	4993.976	0.5556	0.3087
	APHRODITE	1.103078	150.4102	601.6408	0.7446	0.5443
Multan	CHIRPS	0.917936	-17.8402	85.55847	0.7165	0.5133
	ERA-5	1.142324	30.94029	148.3844	0.3469	0.1204
	IMERG	2.968018	427.8342	2051.821	0.8738	0.7637
	APHRODITE	1.713983	150.9696	603.8783	0.6249	0.3905

Association of the Data at an Annual Timescale:

In the yearly correlation analysis between CHIRPS and ground station data, a good to strong correlation was observed across all ground stations, ranging from 0.61 to 0.78 (Table 5, Figure 5). Lahore exhibited the strongest correlation, followed by Islamabad, Multan, Faisalabad, and Murree. IMERG also demonstrated a good to strong relationship with all ground stations, with correlations ranging from 0.55 to 0.87. IMERG showed particularly high correlations for Lahore and Multan, with values exceeding 0.80 for both stations. APHRODITE presented slightly better correlations for Faisalabad and Islamabad. For Murree, APHRODITE showed a higher correlation (0.74) and a moderate correlation for Multan (0.62). CHIRPS exhibited minimal mean error for Lahore, but for Multan and Faisalabad, it was underestimated (-17.8 and -38.5), whereas for Murree and Islamabad, it was overestimated (348.87 and 18.07). These statistics highlight the deviations of rainfall datasets from the yearly mean of ground station data.

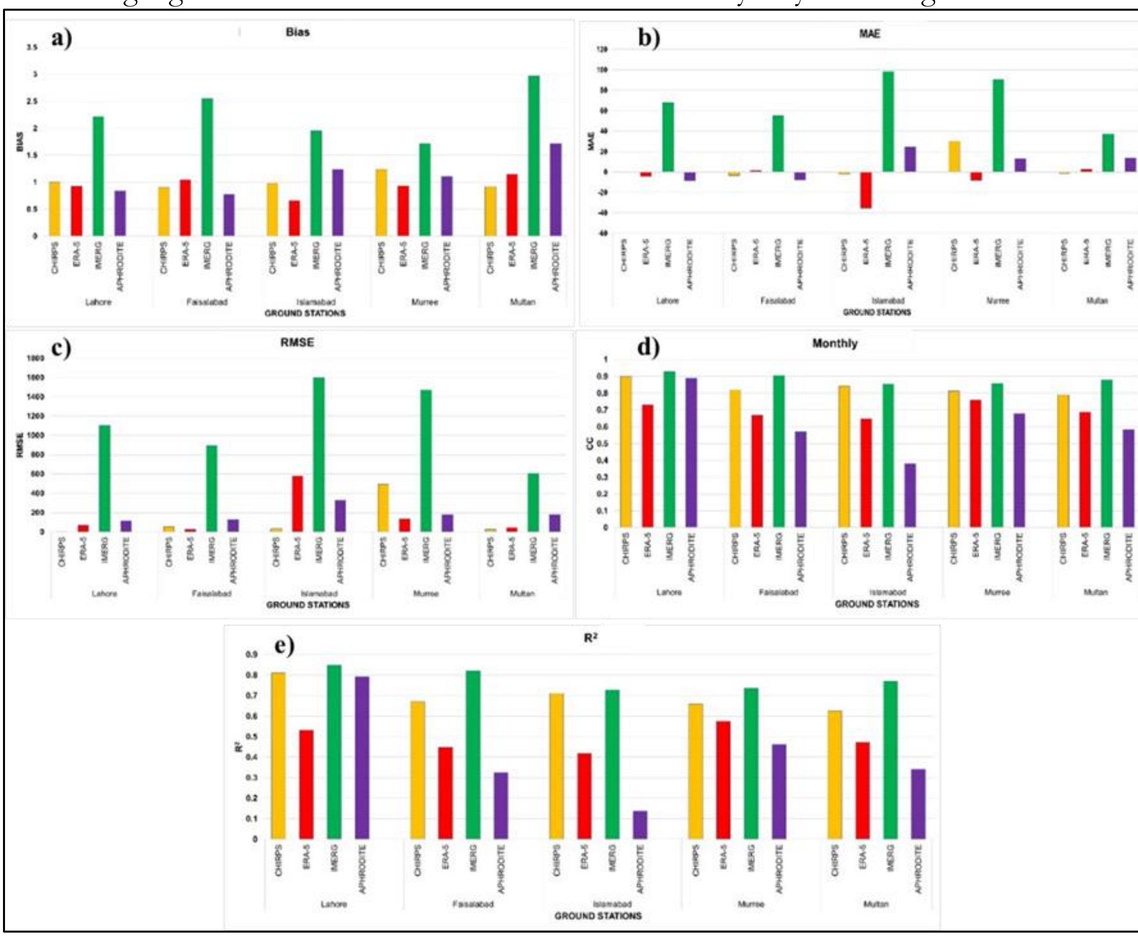

Figure 4: Statistical indicators showing the relationship between the ground station data and daily gridded precipitation datasets at a monthly timescale, a) Bias, b) MAE, c) RMSE, d) CC, and e) R2 with correspondence observatories from 2001–2022.

Discussion:

The four gridded precipitation datasets evaluated in this study—ERA5, IMERG, CHIRPS, and APHRODITE—were thoroughly analyzed using data from five ground stations in Punjab Province, Pakistan, covering the period from 2001 to 2022. Statistical techniques, including bias, mean error (ME), root mean square error (RMSE), correlation coefficient (CC), and coefficient of determination (R^2), were applied to assess these datasets on daily, monthly, and annual scales. Results revealed significant variations in dataset performance across different periods and geographic regions in Punjab. On a daily basis, IMERG exhibited limitations, particularly in Faisalabad, whereas CHIRPS demonstrated the best correlation with ground data in Lahore. IMERG showed greater dispersion from ground data compared to CHIRPS, which had lower Mean Absolute Error (MAE) values in Multan, Islamabad, and Faisalabad. However, IMERG led in monthly correlations across cities. Both IMERG and CHIRPS showed strong correlations in yearly evaluations, with IMERG exhibiting stronger correlations and CHIRPS having lower mean errors. The study's strengths lie in its comprehensive assessment over multiple timescales and climatic zones in Punjab, enhancing the validity of its conclusions. However, regional emphasis and variability in ground station data quality are noted drawbacks. This study contributes to understanding the performance of gridded precipitation datasets, aligning with previous research that highlights the reliability of monthly forecasts [39] and the relative accuracy of IMERG and CHIRPS [18][40]. The findings suggest that, except for Faisalabad, most cities exhibited better daily correlations with CHIRPS, while IMERG outperformed on monthly and annual bases. These results support earlier studies that found monthly projections to be more reliable. Moreover, CHIRPS consistently showed smaller errors compared to IMERG, highlighting its relative accuracy. Nonetheless, disparities were observed, with IMERG occasionally overestimating precipitation and ERA-5 underestimating rainfall compared to other datasets. These findings underscore the importance of considering regional and spatiotemporal factors when using gridded precipitation datasets. Future research should explore hybrid models integrating multiple datasets, expand geographic coverage, incorporate recent data to assess climate change impacts, and employ advanced statistical techniques for evaluating dataset performance, especially in extreme weather events. Such advancements will enhance the selection and application of gridded rainfall datasets for improved water resource management, disaster planning, and climate monitoring.

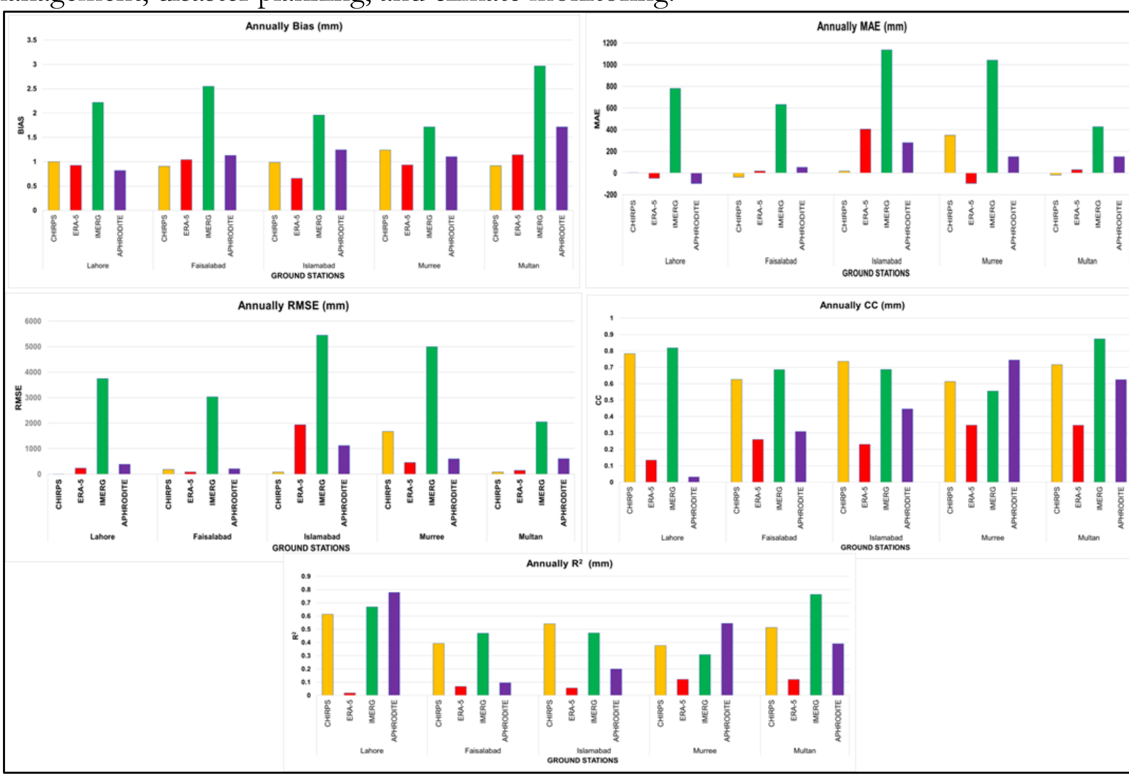

Figure 5: Statistical indicators showing the relationship between the ground station data and daily gridded precipitation datasets at an annual timescale, a) Bias, b) MAE, c) RMSE, d) CC, and e) R2 with correspondence observatories from 2001–2022.

Conclusion:

In countries like Pakistan, where rain gauge distribution is sparse, effective alternatives are needed. Satellite datasets offer a viable solution for assessing precipitation, provided they are efficient and deliver near-real-time data after error correction. This study evaluates four gridded rainfall measurements against rain gauge data from five ground stations, finding that while these datasets have higher errors for daily estimates compared to ground station data, they perform effectively for monthly and annual precipitation. Among the datasets, CHIRPS proved most suitable for both monthly and annual assessments due to its strong correlation and low RMSE

[41]. Biases vary significantly for daily satellite estimates compared to ground observatory data [42], but bias reduces with longer timescales due to data accumulation. For monthly evaluations, ERA-5 is suitable for Murree and Faisalabad due to its low RMSE, while APHRODITE is recommended for annual estimates in Murree due to its lowest variance and highest correlation. CHIRPS, however, remains highly effective for measuring both monthly and annual rainfall across all stations due to its minimal RMSE and bias.

Acknowledgement:

We extend our gratitude to the Pakistan Meteorological Department (PMD) for providing the long-term station observation rainfall data from Lahore, Faisalabad, Multan, Islamabad, and Murree, which was crucial for this study. We also appreciate the satellite-derived rainfall datasets from NASA's Goddard Space Flight Center for IMERG, the Climate Hazards Group for CHIRPS, ECMWF for ERA-5, and the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of the Japan Meteorological Agency for APHRODITE. Their contributions have been invaluable to this research.

Author's Contribution:

All the authors had different contributions to this research work and are mentioned here accordingly. Conceptualization (S.A, S.N.R.G), formal analysis (S.N.R.G), methodology (S.N and S.A), writing—original draft preparation (S.N, M.U, S.A), writing—review and editing (S.A, M.U, M.U, S.M.I), visualization (M.U and I.Y) All authors have read and agreed to the published version of the manuscript.

Conflict of Interest: The authors declare they have no conflict of interest in publishing this manuscript in this Journal.

References:

[1] K. E. Trenberth, "Changes in precipitation with climate change," Clim. Res., vol. 47, no. 1–2, pp. 123–138, 2011, doi: 10.3354/cr00953.

[2] Q. You, J. Min, W. Zhang, N. Pepin, and S. Kang, "Comparison of multiple datasets with gridded precipitation observations over the Tibetan Plateau," Clim. Dyn., vol. 45, no. 3–4, pp. 791–806, 2015, doi: 10.1007/s00382-014-2310-6.

[3] M. Luo, J. Feng, Z. Xu, Y. Wang, and L. Dan, "Evaluating the performance of five twentieth-century reanalysis datasets in reproducing the severe drought in northern China during the 1920s-1930s," Theor. Appl. Climatol., vol. 137, no. 1–2, pp. 187–199, 2019, doi: 10.1007/s00704-018-2591-5.

[4] S. Feng, Q. Hu, and W. Qian, "Quality control of daily meteorological data in China, 1951-2000: A new dataset," Int. J. Climatol., vol. 24, no. 7, pp. 853–870, 2004, doi: 10.1002/joc.1047.

[5] A. Asfaw, B. Simane, A. Hassen, and A. Bantider, "Variability and time series trend analysis of rainfall and temperature in northcentral Ethiopia: A case study in Woleka sub-basin," Weather Clim. Extrem., vol. 19, no. June 2017, pp. 29–41, 2018, doi: 10.1016/j.wace.2017.12.002.

[6] E. Sharifi, R. Steinacker, and B. Saghafian, "Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results," Remote Sens., vol. 8, no. 2, 2016, doi: 10.3390/rs8020135.

[7] M. Adnan et al., "Snowmelt runoff modelling under projected climate change patterns in the Gilgit river basin of northern Pakistan," Polish J. Environ. Stud., vol. 26, no. 2, pp. 525–542, 2017, doi: 10.15244/pjoes/66719.

[8] M. Latif, F. S. Syed, and A. Hannachi, "Rainfall trends in the South Asian summer monsoon and its related large-scale dynamics with focus over Pakistan," Clim. Dyn., vol. 48, no. 11–12, pp. 3565–3581, 2017, doi: 10.1007/s00382-016-3284-3.
[9] C. Funk et al., "The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes," Sci. Data, vol. 2, pp. 1–21, 2015, doi: 10.1038/sdata.2015.66.
[10] C. Toté, D. Patricio, H. Boogaard, R. van der Wijngaart, E. Tarnavsky, and C. Funk, "Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique," Remote Sens., vol. 7, no. 2, pp. 1758–1776, 2015, doi: 10.3390/rs70201758.
[11] F. Zambrano, B. Wardlow, T. Tadesse, M. Lillo-Saavedra, and O. Lagos, "Evaluating satellite-derived long-term historical precipitation datasets for drought monitoring in Chile," Atmos. Res., vol. 186, no. November, pp. 26–42, 2017, doi: 10.1016/j.atmosres.2016.11.006.
[12] M. Arshad et al., "Evaluation of GPM-IMERG and TRMM-3B42 precipitation products over Pakistan," Atmos. Res., vol. 249, p. 105341, 2021, doi: 10.1016/j.atmosres.2020.105341.
[13] M. Nawaz, M. F. Iqbal, and I. Mahmood, "Validation of CHIRPS satellite-based precipitation dataset over Pakistan," Atmos. Res., vol. 248, p. 105289, 2021, doi: 10.1016/j.atmosres.2020.105289.
[14] Y. Yang, G. Wang, L. Wang, J. Yu, and Z. Xu, "Evaluation of Gridded Precipitation Data for Driving SWAT Model in Area Upstream of Three Gorges Reservoir," PLoS One, vol. 9, no. 11, p. e112725, Nov. 2014, doi: 10.1371/journal.pone.0112725.
[15] N. Khan, S. Shahid, T. Ismail, K. Ahmed, and N. Nawaz, "Trends in heat wave related indices in Pakistan," Stoch. Environ. Res. Risk Assess., vol. 33, no. 1, pp. 287–302, 2019, doi: 10.1007/s00477-018-1605-2.
[16] M. Dembélé and S. J. Zwart, "Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa," Int. J. Remote Sens., vol. 37, no. 17, pp. 3995–4014, Sep. 2016, doi: 10.1080/01431161.2016.1207258.
[17] G. T. Ayehu, T. Tadesse, B. Gessesse, and T. Dinku, "Validation of new satellite rainfall products over the Upper Blue Nile Basin, Ethiopia," Atmos. Meas. Tech., vol. 11, no. 4, pp. 1921–1936, 2018, doi: 10.5194/amt-11-1921-2018.
[18] L. Bai, C. Shi, L. Li, Y. Yang, and J. Wu, "Accuracy of CHIRPS Satellite-Rainfall Products over Mainland China," Remote Sens., vol. 10, no. 3, p. 362, Feb. 2018, doi: 10.3390/rs10030362.
[19] H. Feidas, "Validation of satellite rainfall products over Greece," Theor. Appl. Climatol., vol. 99, no. 1–2, pp. 193–216, 2010, doi: 10.1007/s00704-009-0135-8.
[20] C. Funk, A. Verdin, J. Michaelsen, P. Peterson, D. Pedreros, and G. Husak, "A global satellite-assisted precipitation climatology," Earth Syst. Sci. Data, vol. 7, no. 2, pp. 275–287, 2015, doi: 10.5194/essd-7-275-2015.
[21] M. F. Iqbal and H. Athar, "Validation of satellite based precipitation over diverse topography of Pakistan," Atmos. Res., vol. 201, pp. 247–260, 2018, doi: 10.1016/j.atmosres.2017.10.026.

[22] S. Ullah, Q. You, W. Ullah, and A. Ali, "Observed changes in precipitation in China-Pakistan economic corridor during 1980–2016," Atmos. Res., vol. 210, pp. 1–14, Sep. 2018, doi: 10.1016/j.atmosres.2018.04.007.

[23] A. Banerjee et al., "Tracking 21st century climate dynamics of the Third Pole: An analysis of topo-climate impacts on snow cover in the central Himalaya using Google Earth Engine," Int. J. Appl. Earth Obs. Geoinf., vol. 103, no. July, p. 102490, 2021, doi: 10.1016/j.jag.2021.102490.

[24] Z. Nawaz, X. Li, Y. Chen, N. Nawaz, R. Gull, and A. Elnashar, "Spatio-temporal assessment of global precipitation products over the largest agriculture region in pakistan," Remote Sens., vol. 12, no. 21, pp. 1–24, 2020, doi: 10.3390/rs12213650.

[25] A. P. Dimri et al., "Western Disturbances: A review," Rev. Geophys., vol. 53, no. 2, pp. 225–246, 2015, doi: 10.1002/2014RG000460.

[26] U. Asmat and H. Athar, "Run-based multi-model interannual variability assessment of precipitation and temperature over Pakistan using two IPCC AR4-based AOGCMs," Theor. Appl. Climatol., vol. 127, no. 1–2, pp. 1–16, 2017, doi: 10.1007/s00704-015-1616-6.

[27] S. Ullah, Q. You, W. Ullah, and A. Ali, "Observed changes in precipitation in China-Pakistan economic corridor during 1980–2016," Atmos. Res., vol. 210, no. April, pp. 1–14, 2018, doi: 10.1016/j.atmosres.2018.04.007.

[28] M. Arshad, X. Ma, J. Yin, W. Ullah, M. Liu, and I. Ullah, "Performance evaluation of ERA-5, JRA-55, MERRA-2, and CFS-2 reanalysis datasets, over diverse climate regions of Pakistan," Weather Clim. Extrem., vol. 33, p. 100373, 2021, doi: 10.1016/j.wace.2021.100373.

[29] M. W. Ashiq, C. Zhao, J. Ni, and M. Akhtar, "GIS-based high-resolution spatial interpolation of precipitation in mountain-plain areas of Upper Pakistan for regional climate change impact studies," Theor. Appl. Climatol., vol. 99, no. 3–4, pp. 239–253, 2010, doi: 10.1007/s00704-009-0140-y.

[30] F. Abbas, "Analysis of a historical (1981-2010) temperature record of the Punjab Province of Pakistan," Earth Interact., vol. 17, no. 15, pp. 1–23, 2013, doi: 10.1175/2013EI000528.1.

[31] S. Khan and M.-U.- Hasan, "Climate Classification of Pakistan," Int. J. Econ. Environ. Geol., vol. 10, no. 2, pp. 60–71, 2019, doi: 10.46660/ojs.v10i2.264.

[32] U. Asmat, H. Athar, A. Nabeel, and M. Latif, "An AOGCM based assessment of interseasonal variability in Pakistan," Clim. Dyn., vol. 50, no. 1–2, pp. 349–373, Jan. 2018, doi: 10.1007/s00382-017-3614-0.

[33] Z. Nawaz et al., "Spatiotemporal Assessment of Temperature Data Products for the Detection of Warming Trends and Abrupt Transitions over the Largest Irrigated Area of Pakistan," Adv. Meteorol., vol. 2020, 2020, doi: 10.1155/2020/3584030.

[34] G. Huffman et al., "NASA GPM Integrated Multi-satellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 06," Nasa/Gsfc, no. January, p. 29, 2020, [Online]. Available: https://pmm.nasa.gov/sites/default/files/imce/times_allsat.jpg%0Ahttps://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf

[35] A. Yatagai, K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, "APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges," Bull. Am. Meteorol. Soc., vol. 93, no. 9, pp. 1401–1415, Sep. 2012, doi: 10.1175/BAMS-D-11-00122.1.
[36] H. Hersbach et al., "The ERA5 global reanalysis," Q. J. R. Meteorol. Soc., vol. 146, no. 730, pp. 1999–2049, Jul. 2020, doi: 10.1002/qj.3803.
[37] C. C. Funk et al., "A Quasi-Global Precipitation Time Series for Drought Monitoring," U.S. Geol. Surv. Data Ser., vol. 832, no. January, p. 4, 2014, doi: 10.3133/ds832.
[38] C. Funk et al., "The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes," Sci. Data, vol. 2, no. 1, p. 150066, Dec. 2015, doi: 10.1038/sdata.2015.66.
[39] M. N. Anjum et al., "Assessment of PERSIANN-CCS, PERSIANN-CDR, SM2RAIN-ASCAT, and CHIRPS-2.0 Rainfall Products over a Semi-Arid Subtropical Climatic Region," Water, vol. 14, no. 2, p. 147, Jan. 2022, doi: 10.3390/w14020147.
[40] A. Banerjee, R. Chen, M. E. Meadows, R. B. Singh, S. Mal, and D. Sengupta, "An Analysis of Long-Term Rainfall Trends and Variability in the Uttarakhand Himalaya Using Google Earth Engine," Remote Sens., vol. 12, no. 4, p. 709, Feb. 2020, doi: 10.3390/rs12040709.
[41] G. Artan, H. Gadain, J. L. Smith, K. Asante, C. J. Bandaragoda, and J. P. Verdin, "Adequacy of satellite derived rainfall data for stream flow modeling," Nat. Hazards, vol. 43, no. 2, pp. 167–185, Oct. 2007, doi: 10.1007/s11069-007-9121-6.
[42] M. Dembélé and S. J. Zwart, "Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa," Int. J. Remote Sens., vol. 37, no. 17, pp. 3995–4014, 2016, doi: 10.1080/01431161.2016.1207258.