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 the newly developed gridded rainfall data set is important for its effective applications. Over the past two decades, the availability of gridded rainfall measurements has increased, however, it remains challenging to identify suitable proxies for conventional station-based measurements. Therefore, this study performed a comparative assessment of rainfall estimates from IMERG, CHIRPS, ERA-5, and APHRODITE, with meteorological station data of five cities in Pakistan including Lahore, Faisalabad, Multan, Islamabad, and Murree. The assessment was performed at multiple temporal scales (daily, monthly, and yearly) using the daily data recorded between 2001 to 2022. The analytical metrics applied in this study included Bias, Mean Error (ME), Root Mean Square Error (RMSE), Correlation Coefficient (CC), and Coefficient of determination (R^2). The results showed interesting spatial and temporal patterns of agreement among the datasets. Correlation for daily data was weak for all the gridded data sets, and among all of these, APHRODITE performed better. The monthly aggregates showed that the IMERG has the highest association with the ground data, followed by the CHRIPS. Likewise, yearly accumulated rainfall records indicated IMERG with the highest correlation, followed by CHIRPS. Overall, IMERG has shown higher consistency across the stations at monthly and yearly temporal scales. CHIRPS excelled with low errors (RMSE and bias) for most locations especially Lahore, but higher errors were found in Murree at the monthly time scale. This study concluded that a single satellite alone cannot show significant results over large areas, rather hybrid of products may be required for better estimation.

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

Introduction:

The global water cycle is strongly influenced by precipitation, which makes it a crucial element with significant 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 key role in the study of regional and global climate change[4][5]. It is essential for agricultural growth, food security, water management, and hydropower generation. It also plays a vital role in identifying weather and climatic conditions [6][7][8]. In addition, it is a key variable for assessing extreme climate changes such as drought, which can have adverse effects on crop production [9][10][11]. The main problem we often face when conducting such studies is the lack of climate data over a long period. Despite the availability of some data sources, observed data from ground stations in Pakistan, including Punjab province, is very limited [12][13]. The quality of available data is often insufficient for hydro-climatological assessments. However, advances in technology have enabled hydroclimatic scientists to estimate climate variables and produce reliable data over longer periods, especially in regions with limited data availability [14][15]. Relying on high-resolution remote sensing data – detailed and precise data collected from satellites or aircraft - appears to be the optimal solution to address the lack of ground station data [12][13] supplementing gauge-based measurement observations, which are direct measurements of precipitation using instruments at specific locations and integrating different datasets. The use of satellite measurements of rainfall has been instrumental in overcoming the challenges posed by the unlimited or limited availability of ground data [16][17][18]. These methods have demonstrated significant effectiveness and cost-effectiveness in recent years [19][20][21].

Recent studies suggested a declining annual trend over central Asia in the last few decades [22][23] reported a slightly increasing trend in annual, seasonal, and monthly rainfall over the northwest Himalayan region from 2000 to 2022. Several current studies in Pakistan found uneven rainfall patterns from 1980 to 2020, where [24] noticed an increasing rainfall trend in the Punjab province using numerous remotely sensed gridded precipitation datasets. Various other studies have been performed to estimate rainfall distribution and trends in Pakistan and noticed an overall increased trend at annual and seasonal scales in the last four decades [22][25], [26]. A declining trend in winter and post-monsoon periods with an increasing trend during monsoon and pre-monsoon periods has been observed from 1980 to 2016 [27].[28] conducted a study for the estimation of the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) dataset against the Surface Precipitation Gauge (SPG) dataset for the past 38 years (1981-2018) in Pakistan. Statistical systems of measurement 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 on daily, monthly, annual, and seasonal time scales. On average, CHIRPS showed an underestimation of rainfall by -0.80 to -0.17 mm/day for the period 1981-2018 at 46 stations. The highest ME, MAE, and RMSE estimated was 2.0 mm/day, 1.8 mm/day, and 1.8 mm/day at the yearly time scale. A high associate was observed at monthly and seasonal timescales. [13] also found CHIRPS as a useful alternative to station data to understand the spatial patterns and temporal trends in the precipitation station-scarce regions of Pakistan.

Objectives:

This study compares the performance of the four rainfall products including 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) with climate station data to overcome limitation of availability of spatially contiguous rainfall data in Pakistan. These evaluations were carried out using gridded rainfall datasets validated over five cities in Punjab including Lahore, Faisalabad, Multan, Islamabad, and Murree. Analytical metrics used 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 understanding of the accuracy and reliability of these rainfall estimates, the analysis covered three timescales (daily, monthly, and annual) using the daily data recorded between 2001 to 2022.

Novelty Statement:

This study performed a comparative assessment of rainfall estimates from IMERG, CHRIPS, ERA-5, and APHRODITE with meteorological station data of five cities in Pakistan, providing a solution to identify proxies for conventional station – based measurements

Material and Methods:

A brief description of the methodology is illustrated in the flow diagram (**Figure 1)**

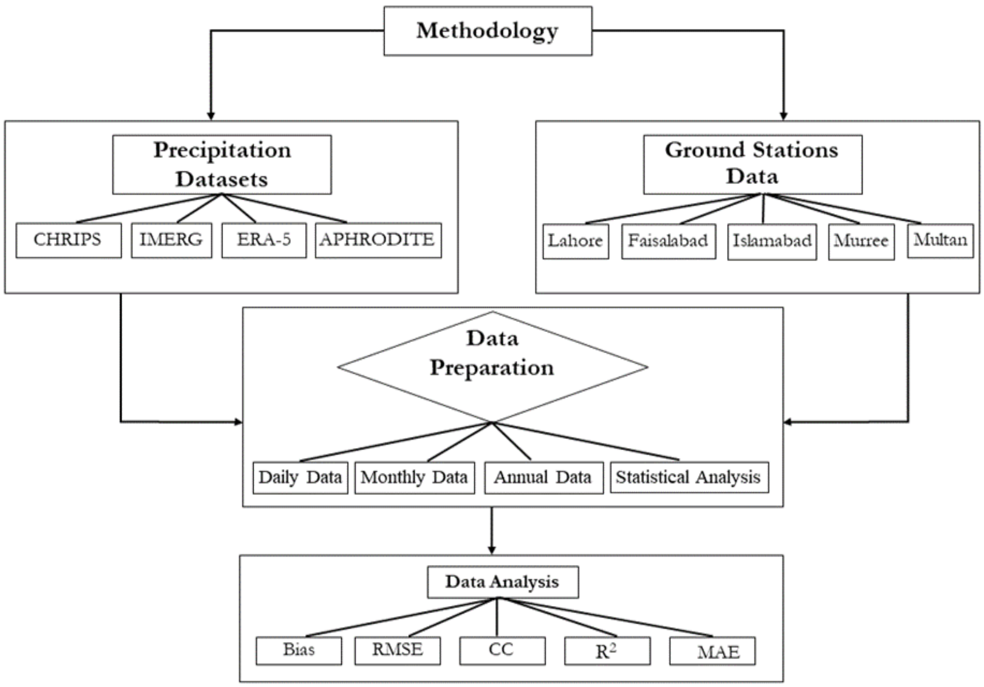

Figure 1: An overview of the methodology applied in this study

Description of Study Area and Location of Ground Stations:

Pakistan is geographically located in Southwest Asia with an area of 881,913 km^2 between the latitude and longitude range of 24◦ to 37◦ N and 60◦ to 75◦ E **(Figure 2)**. The country has a diverse topography that ranges from the Karakoram and Himalayan mountains in the north and northwest to agricultural plains in the Centre and south of the Indus basin along the southern coast of the Arabian Sea [29]. Pakistan is an agrarian country and Punjab is the second largest province. The province has the largest population and produces more than 50% of the country's agricultural commodities [30]. The regional average temperature ranges from 23 to 26 ◦C, with minimum and maximum temperatures ranging from 16 to 19 ◦C, and 29 to 33 ◦C. The northern part of the province receives more rainfall than the southern part, with total annual average rainfall ranging from less than 300 mm in the southern part to greater than 800 mm in the northern part [31]. Two main seasons dominate the overall hydrology of the region, i.e., summer monsoon (June–September) and winter monsoon season (December–April). Maximum rainfall occurs during the summer monsoon due to the monsoon system that originates from the Bay of Bengal and enters the country from the east and northeast. Winter precipitation originates from the Mediterranean Sea due to western disturbances and enters the country from southwest and northeast directions[32]. Large variations and extreme events occur during the summer and winter monsoon rainfall seasons, making the region highly vulnerable to climate change [33].

The selection of ground stations for our study was based on data provided by the Pakistan Meteorological Department (PMD), which ensured representation of key urban Centre and one hill station in Punjab, and the Federal Capital Territory. The selected cities included 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°E, 73.3943°E) reflect the geographical and climatic diversity of the region, each of which has significance in the socio-economic fabric of the province **(Table 1)**. According to the PMD, Lahore has moderate to increased rainfall, especially during the monsoon season, with an average annual rainfall above 617 mm and fluctuating between 332 and 903 mm per year for the past 22 years. Faisalabad, with an average annual rainfall of 615 mm, and Multan experienced an annual rainfall range of 60 to 395 mm per year from 2001 to 2022. With estimated annual rainfall between 580 and 1796 mm per year over the last 22 years, Islamabad shows considerable variability in rainfall patterns. Known for its scenic beauty, Murree experienced an annual rainfall range of 1136 to 2420 mm per year 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

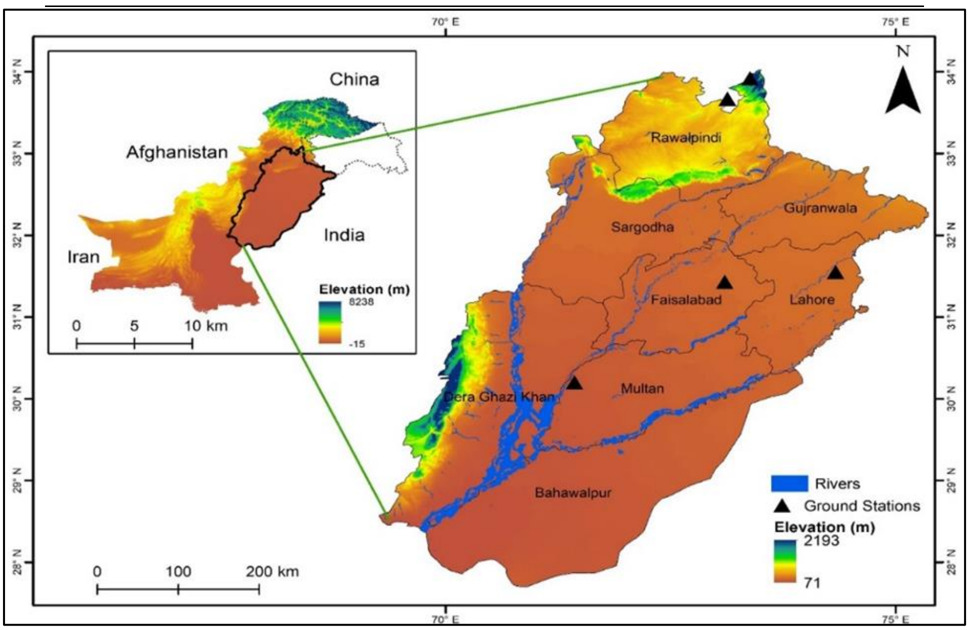

Figure 2: Topographic settings and location of the ground stations

Gridded Rainfall Measurements:

The gridded rainfall measurements were acquired from four sources including CHIRPS [9], IMERGE[34], APHRODITE [35], and ERA-5 [36], spanning from 2001 to 2022. A brief description of these datasets is given below in Table 2 and subsequent sub-sections.

CHIRPS:

The CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) gridded rainfall product is available at high spatial resolution (0.05°) and covers an extended time series from 1983 to the near present. The temporal resolution of CHIRPS includes daily, monthly, pentad, and annual composites [20][37][38] available at https://www.chc.ucsb.edu/data/chirps. The daily, monthly, and annual gridded were used for the period 2001 to 2022 in this study.

IMERG:

The 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 three different temporal aggregates (30 minutes, 1 day, and 1 month). In this study, we obtained the IMERG-Final v6 data from https://pmm.nasa.gov/data-access/downloads/gpm.

ERA-5:

The ERA-5 (Fifth generation ECMWF atmospheric reanalysis of the global climate covering) is the fifth global atmospheric reanalysis product from ECMW. In this study, daily data with a spatial resolution (0.25°×0.25°) was downloaded from https://climatedataguide.ucar.edu/climate-data/era5-atmospheric-reanalysis and then converted to monthly total and annual composites for subsequent analysis.

APHRODITE:

The APHRODITE (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources) project developed state-of-the-art high-resolution (0.25°) gridded daily rainfall datasets for the Asian region. The APHRODITE Water Resources project has been implemented by the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of the Japan Meteorological Agency (MRI/JMA) since 2006 [35]. This data is available from 1961 and can be obtained from http://aphrodite.st.hirosaki-u.ac.jp/japanese/products.html and http://aphrodite.st.hirosaki-u.ac.jp/product/.

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	ECMW F	1950 - Present	0.125° × 0.125°

Data Analysis:

A rigorous analysis was performed to analyze the associations between CHIRPS, IMERG, ERA-5, and APHRODITE datasets and the rainfall data recorded at the stations. The statistics indicators used for the 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)**, denotes the average disparity between the two datasets, which could be positive or negative. A negative bias signifies underestimation, whereas a positive bias implies overestimation.

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

Mean Absolute Error (MAE):

The Mean Absolute Error (MAE) indicates the magnitude of the errors, **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), emphasizes larger errors compared to MAE and serves as a metric for assessing the average magnitude of errors, **Eq (3).**
It's calculated as:

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

Correlation Coefficient (CC):

The correlation coefficient (CC) is a measure used to evaluate the consistency between two observations **(Eq 4)**. Its range falls between -1 and +1 indicating high negative and high positive associations, whereas the values near zero show a lack of strong relation.
The formula for CC is:

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

Results:

In this study daily, monthly, and annually aggregated precipitation datasets derived from CHRIPS, ERA-5, IMERG, and APHRODITE were evaluated against rain-gauge observations spanning 22 years (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 obtained from the rain gauge stations was compared with all four gridded rainfall estimates (CHRIPS, IMERG, ERA-5, and APHRODITE). The associate was determined and evaluated using the metrics including Bias, MAE, RMSE, R^2, and p **(Table 3)**. Overall, the highest correlations were observed with APHRODITE data, and the maximum correlation was observed at Lahore station (0.63), followed by Muree (0.56). A high bias was observed with the IMERG data across all stations **(Table 3, Figure 3)**. In the case of Lahore, CHIRPS, ERA-5, and IMERG exhibited very weak correlations for daily precipitation datasets, with values ranging between (0. 2 to 0.35). The results for Murree, which is geographically closer to Islamabad, APHRODITE once again showed a slightly better relationship, with correlations slightly above 0.5. However, all other precipitation datasets exhibit weaker correlations, ranging from 0.1 to 0.37. On the other hand, CHIRPS indicated a very slight overestimation from the mean of ground station data (MAE=0.00147), while APHRODITE and ERA-5 both underestimated (MAE= -0.28 and MAE= -0.139). At the Multan station, CHIRPS was underestimated but it was closer to the mean of ground station daily data, while ERA-5 and APHRODITE overestimated with values of (0. 088 and 0. 44). The mean error value of IMERG was higher than all other datasets. In Lahore, CHIRPS exhibited a low RMSE, suggesting higher accuracy, while ERA-5 and APHRODITE followed with values exceeding 12.5 and 20.9. For Islamabad, only CHIRPS showed a small RMSE, indicating high accuracy, consistent with its good coefficient correlation. All precipitation datasets exhibited considerable dispersion with observation at the Murree station, however, ERA-5 and APHRODITE show relatively lower accuracy.

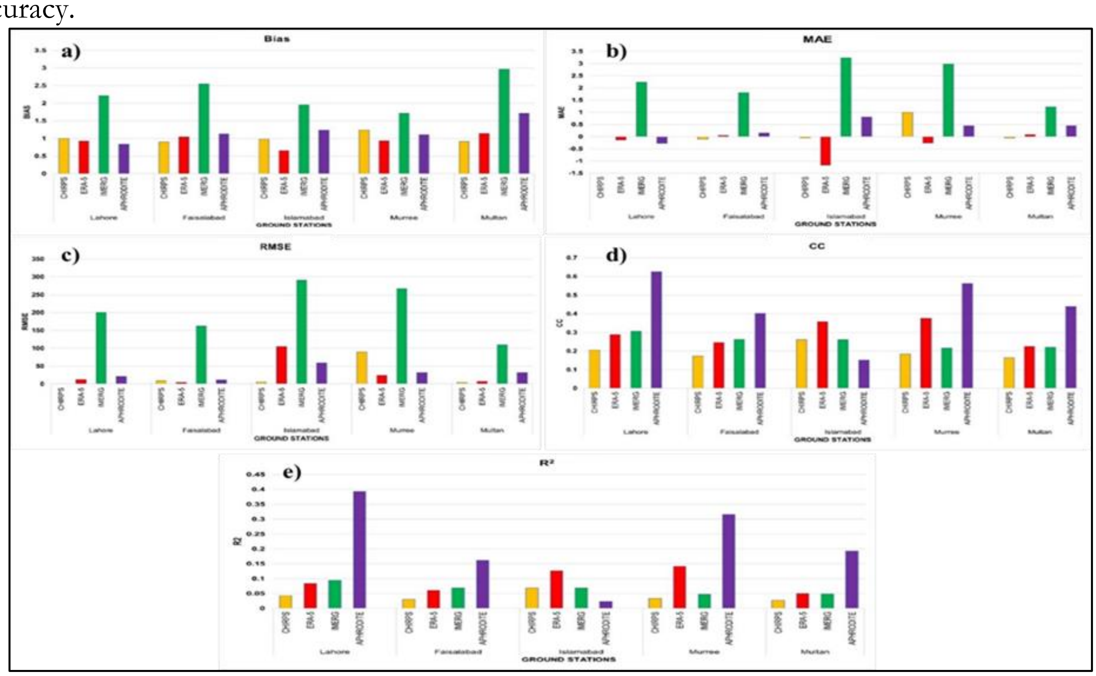

Figure 3: Statistical indicators showing the relationship between the ground station data and daily gridded precipitation datasets a) Bias, b) MAE, c) RMSE, d) CC, and e) R2 with correspondence observatories from 2001–2022.

Association of the Data at a Monthly Timescale:

All the datasets indicated a significantly strong correlation at the monthly timescale **(Table 4, Figure 4)**. The maximum correlation was observed between the IMERG monthly aggregates. For the Lahore station, the correlation was highest with IMERG (0.93), followed by CHIRPS (0.9), APHRODITE (0.89), and ERA-5 (0.73). At Faisalabad station also, IMERG showed the highest correlation with ground station data for monthly aggregates (0. 90), followed by CHIRPS (0. 8). Similarly at Islamabad station, IMERG secured the highest correlation (0. 8). In terms of the MAE, it was noted that CHIRPS measured lower than the mean of ground station data for Multan, Islamabad, and Faisalabad, ranging from -1 to -3, indicating closeness to the ground station's monthly aggregates. Interestingly, a very low MAE of overestimation was observed for Lahore, while a relatively higher overestimation was found for Murree. ERA-5 underestimated Lahore, Islamabad, and Murree, with a relatively high amount of mean underestimated error in Islamabad, exceeding -35. CHIRPS also exhibited very low RMSE values, less than 1 for Lahore, indicating high accuracy in estimating precipitation. Conversely, it showed very high dispersion for Murree, indicating the lowest accuracy among all studied cities. Despite providing the strongest relationship for monthly precipitation datasets, IMERG's high RMSE can be attributed to the presence of outliers and significant differences between sub-datasets of ground stations and precipitation datasets at specific dates. APHRODITE, despite having a low correlation, yielded high RMSE values for all ground stations, ranging from 115.3 to 178.8, suggesting lower accuracy compared to CHIRPS and ERA-5. It's noteworthy that despite APHRODITE having a lower correlation compared to IMERG, it showed lower RMSE values due to less variance between sub-datasets of APHRODITE 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

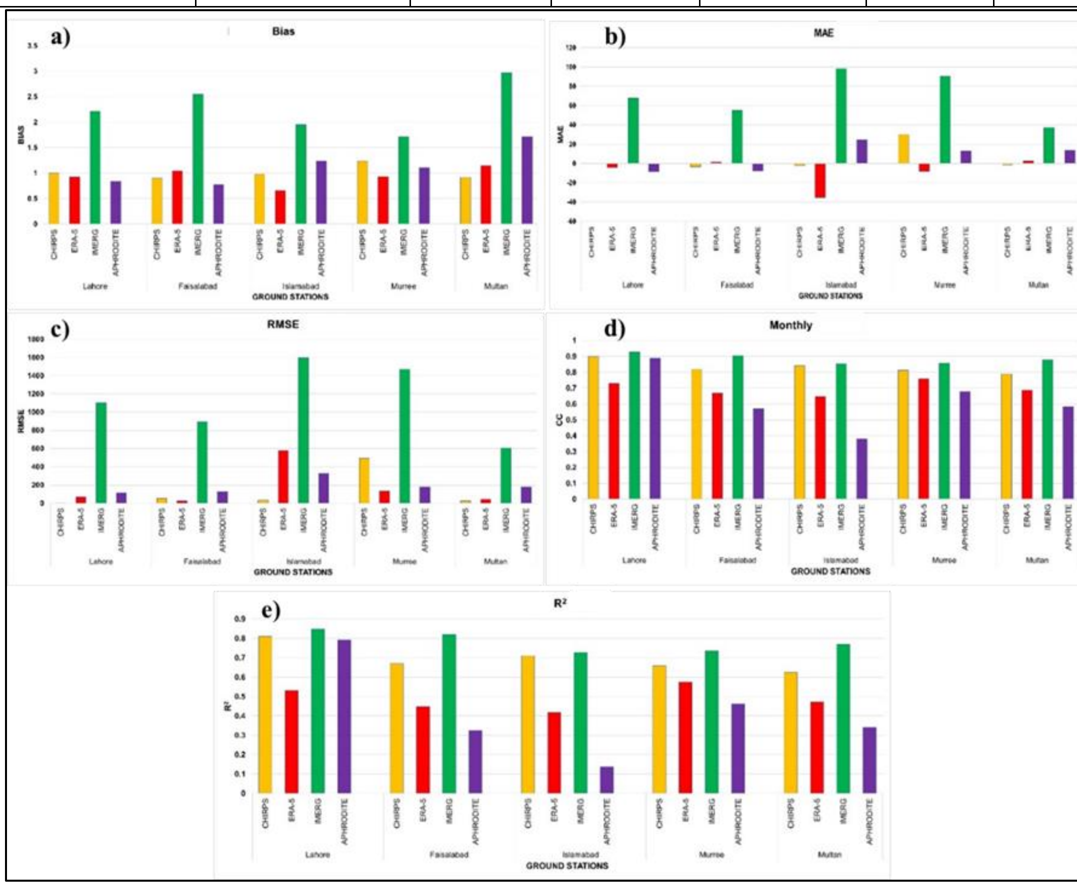

Figure 4: Statistical indicators showing the relationship between the ground station data and daily gridded precipitation datasets at a monthly timescale, a) Bias, b) MAE, c) RMSE, d) CC, and e) R2 with correspondence observatories from 2001–2022.

Association of the Data at an Annual Timescale:

In terms of yearly correlation analysis between CHIRPS and ground station data, a good to strong correlation was evident in all ground stations, ranging from 0.61 to 0.78 **(Table 5, Figure 5)**. Among them, Lahore exhibited the strongest correlation, followed by Islamabad, Multan, Faisalabad, and Murree. IMERG also indicated a good to strong relation in all ground stations, ranging from 0.55 to 0.87. Notably, IMERG showed higher correlations for Lahore and Multan compared to CHIRPS, with values exceeding 0.8 for both ground stations. However, APHRODITE showed slightly better correlations for Faisalabad and Islamabad. Conversely, for the yearly precipitation datasets of Murree, APHRODITE demonstrates a higher correlation (0. 74) and it also showed a moderate relation for Multan (0.62). CHIRPS exhibited a minimal mean error for Lahore, while for Multan and Faisalabad, it was underestimated (-17.8 and -38.5). Conversely, for Murree and Islamabad, it was overestimated (348.87 and 18.07). These statistics indicated the deviation of rainfall datasets from the yearly mean of ground station data.

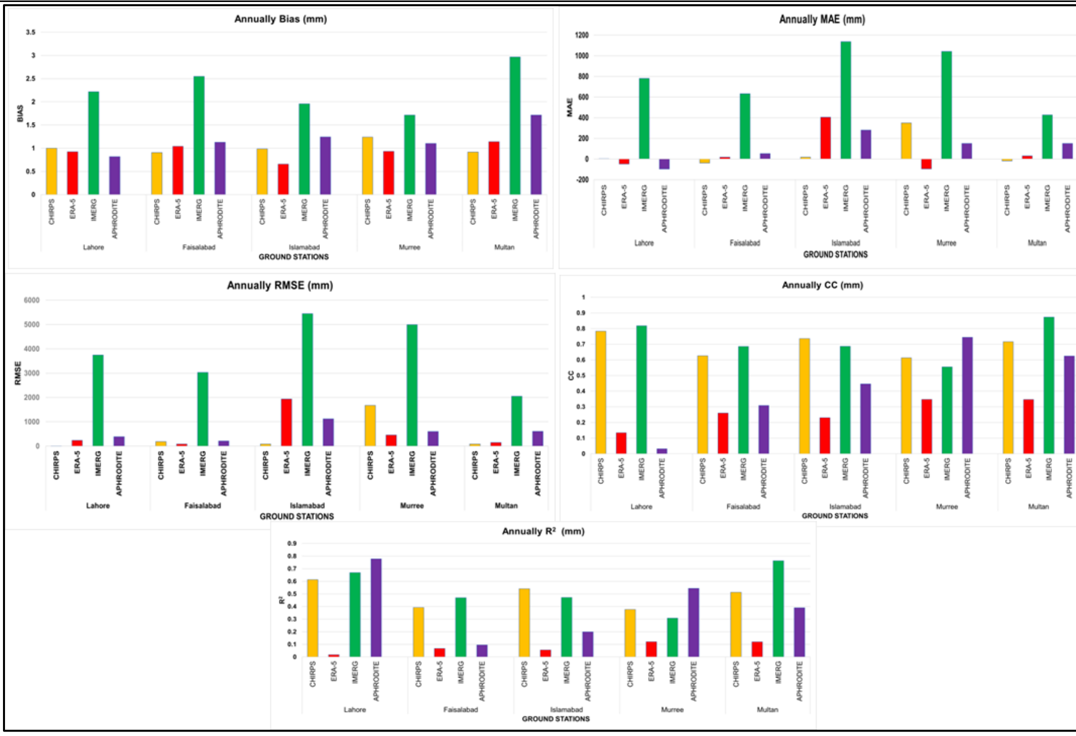

Figure 5: Statistical indicators showing the relationship between the ground station data and daily gridded precipitation datasets at an annual timescale, a) Bias, b) MAE, c) RMSE, d) CC, and e) R2 with correspondence observatories from 2001–2022.

Discussion:

The four gridded precipitation datasets in this study—ERA5, IMERG, CHIRPS, and APHRODITE—are thoroughly evaluated using data from five ground stations in Punjab Province, Pakistan, covering the period from 2001 to 2022. Utilizing statistical techniques including bias, mean error (ME), root mean square error (RMSE), correlation coefficient (CC), and coefficient of determination (R^2), the evaluation is conducted on a daily, monthly, and annual basis. The results highlight notable variations in the dataset's performance throughout different periods and geographical regions in Punjab. Daily, IMERG showed flaws, especially in Faisalabad, whereas CHIRPS showed the best correlation with ground data in Lahore. In contrast, IMERG exhibited greater dispersion from ground data than CHIRPS, which showed lower Mean Absolute Error (MAE) values in Multan, Islamabad, and Faisalabad. However, IMERG led in correlation across cities monthly. Both IMERG and CHIRPS showed strong cross-station correlations in the yearly evaluations, with IMERG exhibiting stronger correlations and CHIRPS having lower mean errors. The study's merits are found in the thorough assessment it conducted over several time scales and climatic zones in Punjab, which improves the validity of its conclusions. The study's regional emphasis and the fluctuations in the quality of ground station data are its drawbacks, though. This study adds to the body of knowledge regarding the performance of gridded precipitation datasets. It is consistent with earlier research that highlights the dependability of monthly forecasts [39] and the relative accuracy of IMERG and CHIRPS in various parts of the world [18], [40]. The research indicates that, with the exception of Faisalabad, most cities showed greater daily correlations for CHIRPS, although IMERG outperformed it on a monthly and annual basis. These results support past studies that found monthly projections to be more reliable. Furthermore, as demonstrated by studies conducted in China and India, CHIRPS consistently showed smaller errors than IMERG across a range of timeframes, highlighting its relative accuracy. Nonetheless, disparities were observed: in some areas, IMERG tended to overstate precipitation as previously documented, but ERA-5 regularly underestimated rainfall in comparison to other datasets. This is probably because Punjab has particular topography and climatic circumstances. These findings highlight how crucial it is to take regional and spatiotemporal variables into account when using gridded precipitation datasets. Future research should investigate hybrid models that integrate multiple datasets, expand the geographic scope of the study outside Punjab, integrate recent data to evaluate the impacts of climate change, and employ sophisticated statistical techniques to thoroughly assess the performance of the datasets, especially in recording extreme weather events. These developments will improve the choice and application of gridded rainfall datasets for improving water resource management, disaster planning, and climate monitoring.

Conclusion:

In countries like Pakistan, where the distribution of rain gauges is sparse, effective alternatives are required to address this problem. Research suggests that satellite datasets can be used to assess precipitation, but they must be efficient and provide near-real-time data after error correction. This study assesses the four gridded rainfall measurements against the rain gauge data recorded at the five ground stations. This research concludes that the gridded datasets have higher errors in providing daily estimates compared to ground station data. However, these datasets effectively estimated monthly and annual precipitation. Among the datasets investigated in this study, CHIRPS proved to be the most suitable for both monthly and annual cases due to its strong correlation and small RMSE [41]. Biases differ significantly when comparing daily satellite remote sensing estimates with daily rain observatory data[42]. However, when analyzing monthly or longer timescales, a reduction in bias is observed due to the temporal accumulation of data. For Murree and Faisalabad, it is appropriate to use ERA-5 for monthly evaluation because it showed a small RMSE with high accuracy. APHRODITE is recommended for annual estimation in Murree because of its lowest variance and highest correlation. However, CHIRPS proved to be highly effective in measuring monthly and annual rainfall with minimal RMSE and bias across all stations

Acknowledgement:

We would like to express our gratitude to the Pakistan Meteorological Department (PMD) provided long-term station observation rainfall data from Lahore, Faisalabad, Multan, Islamabad, and Murree, which was essential for this study. We are also thankful for the satellite-derived rainfall datasets used in our analysis: NASA's Goddard Space Flight Center for the Integrated Multi-satellite Retrievals for GPM (IMERG) data, the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS), the ECMWF Fifth Generation Atmospheric Reanalysis (ERA-5), and the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of the Japan Meteorological Agency for the Asian Precipitation - Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE). Their efforts in making these valuable datasets available have significantly contributed 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.