Comparison and Evaluation of Remotely Sensed Gridded and Observed Precipitation Data

By

Engr. Arsalan Ghani

FALL 2018-MS RS & GIS 00000274785

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Remote Sensing and GIS

Institute of Geographical Information Systems

School of Civil and Environmental Engineering

National University of Sciences & Technology

Islamabad, Pakistan

February, 2022

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Engr. Arsalan Ghani (Registration No.274785), of Session 2018 (Institute of Geographical Information Systems)** has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulation, is free of plagiarism, errors, and mistakes, and is accepted as partial fulfillment for the award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: __________________

Name of Supervisor: Dr Ejaz Hussain

Date:

Signature (HOD): _________________

Date:

Signature (Dean/Principal): _____________

Date:

CERTIFICATE

Certified that the contents and form of thesis entitled **"Comparison and Evaluation of Remotely Sensed Gridded and Observed Precipitation Data"** submitted by **Engr. Arsalan Ghani** has been found satisfactory for the requirement of the degree.

Supervisor: ________________

Dr. Ejaz Hussain
(Associate Dean, IGIS-SCEE, NUST)

Member: __________________

Dr. Muhammad Azmat
(Assistant Professor, IGIS-SCEE, NUST)

Member: ________________

Dr. Hamza Farooq Gabriel
(Professor, NICE-SCEE, NUST)

ACADEMIC THESIS: DECLARATION OF AUTHORSHIP

I, Arsalan Ghani declare that this thesis and the work presented in it is my own and have been generated by me as the result of my own original research.

Comparison and Evaluation of Remotely Sensed Gridded and Observed Precipitation Data I confirm that:

1. This work was done wholly by me in candidature for an MS research degree at the National University of Sciences and Technology, Islamabad.

2. Wherever I have consulted the published work of others, it has been clearly attributed.

3. Wherever I have quoted from the work of others, the source has been always cited. With the exception of such quotations, this thesis is entirely my own.

4. I have acknowledged all main sources of help.

5. Where the work of the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

6. None of this work has been published before submission. This work is not plagiarized under the HEC plagiarism policy.

Signed: ……………

Dated: ……………...

DEDICATION

To

My Sweet & Loving, Family…

To my mother for her priceless love and care,

To my brother for being there in every tough moment and in my hard times,

To my sisters for their prayers and concerns,

To my father for moral support, strength, encouragement, and for believing in me.

Thanks for understanding the stressful moments I had, for the

prayers and support to overcome them and for all the joy you have

brought to me.

ACKNOWLEDGEMENTS

All the praise and thanks to Almighty ALLAH, the most gracious and merciful, who enabled me with the power and means to contribute a drop to the existing ocean of knowledge.

I acknowledge with deep appreciation the inspiring guidance, constant encouragement, valuable suggestions, and kind supervision of Dr. Ejaz Hussain, Dr. Muhammad Azmat, and Dr. Hamza Farooq Gabriel, their kind and generous efforts made it possible for the successful completion of this research work.

In the end,

I am short of words and indebted to my parents, who have always been a beacon and source of inspiration for me in perusing my professional pursuits.

Engr. Arsalan Ghani

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Explanation
WD	Western Distribution
ICT	Islamabad Capital Territory
GPM	Global Precipitation Measurement
IMERG	The Integrated Multi-Satellite Retrievals for GPM
TRMM	The Tropical Rainfall Measuring Mission
TMPA	Multi-satellite Precipitation Analysis
3B42_RT	TRMM Product Real-Time Version
3B42_V7	TRMM Product Version Seven
PMD	Pakistan Meteorological Department
IPCC	The Intergovernmental Panel on Climate Change
WFD	WATCH Forcing data
FBI	Frequency Bias Indicator
SD	Standard Deviation
IDW	Inverse Distance Weighting

ABSTRACT

Spatial variability of precipitation directly affects socio-economic and environmental conditions, therefore accurate and valid measurements of precipitation are needed. Due to the free availability and easy access to gridded precipitation data from multiple sources, questions arise about its credibility like which is a better-gridded dataset and for which precipitation patterns. This study focused on these questions by carrying out the comparison and evaluation of three different gridded datasets and the observed data. These gridded and observed datasets include "The Integrated Multi-Satellite Retrievals for GPM (IMERG), The Tropical Rainfall Measuring Mission (TRMM) products 3B42-V7 & 3B42-RT and Pakistan Meteorological Department (PMD)''. These datasets were analyzed for three different seasons (monsoon, pre-monsoon, winter), and annually for multiple weather stations. The weather stations were classified into ''Monsoon, Western Distribution (WD), and Hybrid'' according to the precipitation patterns. Seven comprehensive statistical parameters (mean, Standard deviation, 5-days max rainfall, 95% percentile, percentage wet days, max dry spell length, and max wet spell length,) were calculated for each season and for each precipitation distribution pattern. Pearson correlation of daily data of 20-years at each station was also calculated as an additional parameter. Furthermore, the IDW interpolation technique was used for further analysis of the results of the above-mentioned parameters. Results suggested that in monsoon and WD stations GPM IMERG has the highest performance among the three ($r = 0.882$ & $r = 0.543$). While at Hybrid stations 3B42_V7 performance is better ($r = 0.840$). TRMM 3B42_RT was the least accurate in precipitations estimation in monsoon, WD and as well as hybrid stations ($r = 0.523$, $r = 0.350$, $r = 0.708$). The results suggest that these datasets can be used as an alternative source for various hydro-meteorological and hydro-climatological applications after accuracy assessment.

Chapter 1

INTRODUCTION

Precipitation is one of the most crucial input variables of the water cycle, as the change in its spatial variability directly affects socio-economic conditions. For accurate precipitation calculations, spatially well distributed and a sufficient number of ground stations are required. However, their low numbers, accessibility due to spatial locations, and the huge budgetary requirements are the obstacles. To overcome such obstacles', credible Gridded data is the alternative, The credibility of such data needs to be evaluated on a seasonal and as well as spatial basis. This study aims to focus on and address these aspects for the three different types of selected datasets.

1.1 Background

Precipitation is increasingly important for earth-system studies. Lack of adequate short-term spatial distribution precipitation data has always been a source of errors in hydro-climatic studies. (Anjum et al., 2018). for tactical planning, management, predominantly the high-elevation regions with mostly scarcely distributed stations. The streamflow forecasts are fundamental foundations of the future for understanding the water cycle. Consequently, quantification of the impact of climate change on major hydrological components is of high significance and remains a challenge (Azmat et al., 2018). Precipitation is a vital hydrological element that administers the renewable water resources influencing hydropower generation, ecological integrity, and agro-economic development. So, an accurate assessment of precipitation and valid data is needed (Javanmard et al., 2010).

The spatial variation of mean annual and seasonal precipitation in TRMM display two main precipitation patterns at the Caspian Sea and over the Zagros Mountains. Statistical analyses were used to compare both datasets and the results of the study showed that TRMM underestimates mean annual precipitation. (Guoqiang Tang., 2016).

Considerable advances in quantitative estimation of distributed rainfall have been made but still large uncertainty in the estimated parameters resulting from the spatial variability of rainfall remains. (Davidson and Mike., 2013).

To the best of our knowledge, not substantial research so far has been carried out on this topic to explore the potential. There is still plenty of potential areas yet to be identified in context to the satellite data set (Li et al., 2013). A study was conducted in Southern Africa for the comparison of precipitation datasets. The main objectives of the study were to systematically compare TRMM and WFD (Watch Force Data) and to evaluate their performance. The water resources were equated by hydrological models using regionalization with TRMM-TMPA and WFD data. Calibration was made independently for each basin from the parameter set at each discharge station used. The results reveal that the special variation patterns of mean annual precipitation from TRMM and WFD datasets are similar.

Dahri, et al., 2016 examined the distribution of rainfall in the high altitudes of the Indus basin. The main objectives of the study were to analyze the dependency of precipitation on altitude and to observe the spatiotemporal distribution. Results on comparison of datasets revealed that gridded datasets were prone to significant errors and because of these errors it is necessary to validate the gridded data for accurate precipitation calculations.

The influence of climate variation is recognized as a notable current global alarm (IPCC., 2014). Upsurges in surface temperatures, variability of precipitation patterns both spatially and temporal are susceptible to climatic variations, exclusively temperature and precipitation. (Kharin et al., 2013). Understanding the mechanism of complicated hydrological processes is important for sustainable management of water resources and for that accurate measurement of rainfall is necessary. (Xue L et al.,2018)

For the sustainability of our ecosystem proper management of water resources is important and for this reason, analysis of climate change impacts at a sharper temporal and spatial scale is necessary. (Narayan Kumar and Xinzhong., 2017)

The hydrological cycle of Himalayan rivers is subjugated by monsoon precipitation and snowmelt, but their comparative impact is not well studied because this remote region has scared weather stations. This study used a mixture of validated remotely sensed parameters to distinguish the spatiotemporal variation of precipitation, snowfall, and evapotranspiration to enumerate their comparative influence to mean river discharge. (Bodo Bookhagen., 2010).

During the last two historical eras, abundant datasets have been established for local hydrological valuation, but these datasets frequently show changes in their spatial and temporal distributions of rainfall. which is one of the most critical input variables in global hydrological casting. This study is designed to discover the rainfall characteristics of the Watch force data (WFD) data and compare these with the succeeding characteristics consequential from gridded data (TRMM 3B42 and GPCP 1DD) and observed in-situ data. It equated the consistency and difference between the WFD and gridded data in north India and examined seasonal weather patterns. Results show that Both WFD and gridded datasets underrate rainfall compared to the measured data but the rainfall from WFD is better assessed than that from the gridded dataset. (Smedman et al., 2020)

Statistical models and their relationship with rainfall and topography are essential for precipitation measurements in mountainous regions. The study examines numerous extensions of the classical rainfall in straight comparison, explicitly linear regression, and kriging with external drift (KED). The benefits of modeling components are scrutinized for the interpolation of seasonal and daily mean precipitation using cross-authentication having densely placed weather stations measurements. accuracy can be improved with KED as compared to ordinary kriging. With the results acquired it is concluded that the topography affects rainfall in high-elevation regions. (D. Masson and C. Frei., 2013).

Gridded datasets which capture the spatial variability of rainfall are critical for hydrological modeling. The objective of the study was to develop a high-quality gridded dataset for rainfall by joining gridded data and climate observations. (A.F. Lutz., 2013).

1.2 Artifacts of the dataset used

The Global Precipitation Measurement (GPM) provides the global observations of rainfall. GPM (IMERG) products are still making developments in areas such as spatial and temporal resolutions. This study uses the latest GPM IMERG product is associated with the TRMM TMPA product (3B43) in the boreal summer of 2014 and the boreal winter of 2015 on a worldwide scale. The result shows that the IMERG product can detect heavy precipitation regions in the Northern and Southern Hemispheres rationally well. Differences between IMERG and 3B43 vary with topography and spatial variation of precipitation in both seasons. Positive relative variances are primarily detected at low rainfall rates and negative differences at high rainfall rates. (Zhong Liu., 2016)

The accurate valuation of extreme rainfall is vital for predicting hydrologic extremes and developing a flood risk management system. Recent gridded-based precipitation products provide vital alternative sources of rainfall data for such applications, yet their quality and applicability with respect to extreme rainfall have not been studied sufficiently. The performances of the TRMM-3B42 and GPM IMERG data in extreme rainfall estimation were assessed for this study over China for the periods of 2000–2017 and 2014–2017. The result shows that gridded data can capture the spatial pattern of extreme rainfall reasonably well over China with an overall underestimation for extreme rainfall rate and an overreckoning for annual total extreme rainfall. GPM IMERG performance was high compared to TRMM 3B42 for nearly all evaluation metrics when equated over the same time. Furthermore, the performances were well in south and east China with humid monsoon climatic conditions, then in arid west China with high elevation, demonstrating a significant impact of topography and climate. More studies are still required to validate data in regions with a complex topography and dry climate, and further advance the retrieval algorithm to better support disaster hazard reduction and other hydrological applications, exclusively in areas with a sparse rain gauge network. (JianFanga and WentaoYan.,2019).

The performance of IMERG and its predecessor, the TRMM 3B42 Version, was cross assessed using data from the hourly rain gauge network over the Tibetan Plateau (TP). Analyses of precipitation estimates in the warm season of 2014 reveal that GPM-IMERG shows better correlations and lower errors than 3B42V7. GPM-IMERG also appears to detect light precipitation better than 3B42V7. GPM IMERG shows the potential of detecting solid precipitation, which cannot be estimated from the 3B42V7 products. (Tian et al., 2016)

The Canadian rainfall platform produces 6-hourly rainfall over a 10-km grid across Canada by joining in-situ observations with a setting field provided by the Global Environmental Multiscale (GEM) estimation model. While rainfall data from the GPM are additionally included and are compared. The frequency bias indicator and the equitable threat score are used as performance criteria. Analysis was done on four climatic regions. Results implied that GPM-IMERG advances the ETS and FBI for all regions. For assessment of the value of IMERG that are vital for water capital management in Canada, a fifth zone that has an inferior weather stations density was considered. It is believed that joining satellite data with other remotely sensed products will provide a significant increase in the quality of data. (Alaba Boluwade., 2017).

To obtain the high-resolution rainfall data for different temporal resolutions, the spatial downscaling technique may offer a better description of the spatial variability of rainfall. GPM multitemporal precipitation analysis at 0.05° resolution is developed and applied in the humid region of China. While assessment of the results reveals, that rainfall outperformed all precipitation variables in relation to the coefficient of determination value, whereas the outperformance was for the annual precipitation variables but it underperformed for seasonal and the monthly variables in terms of the calculated root mean square error value. The downscaling technique developing from the suggested methodology captured the spatial patterns with high accuracy at higher spatial resolution. (Zheng and Kang Zuo., 2020).

Evaluation of precipitation datasets is of high significance. A comprehensive assessment of GPM-IMERG product and valuations over south India at a daily precipitation time scale for the monsoon season is done in this study. The GPM precipitation is compared with widely used TRMM-TMPA and in-situ observations. The result implied that the GPM-IMERG estimates corresponding to the mean monsoon precipitation are more realistically estimated than the adjusted TRMM-TMPA. The three gridded-based estimates show high false rainfall which is a rain-shadow region. These initial results need to be confirmed in the coming seasons of monsoon in future studies. (Satya Prakash., 2016)

Three gridded rainfall products 3B42V7, GPM IMERG-V05, GPM-IMERG-V04, and China half-hourly rainfall products are assessed using measurements from compactly placed weather stations in Guangdong Province, China. The products are equated with in-situ data on annual, monthly, daily, and hourly bases. Overall, the CMPA estimate performance is best in comparison with the in-situ data. The improvement of GPM-IMERG-V5 over 3B42_V7 is noteworthy, especially in dropping the hit bias and missed precipitation, resulting in better recognition of the low precipitation and high precipitation. The products have dissimilar error characteristics and display high spatial variations. TRMM-3B42V7 and GPM-IMERG have high areas of overestimation in the mountainous areas and underestimation in the coastal areas, while CMPA is characterized by an alternate distribution of small positive and negative values. (Dashan Wang and Xianwei Wang., 2018).

The assessment of gridded rainfall products at local scales is vital for enhancing satellite algorithms and sensors, it can also offer valuable guidance when selecting alternative precipitation. It is essential to evaluate GPM IMERG products and make associations with TRMM TMPA products in various regions to attain a global view of the application of GPM IMERG products. The study aims to evaluate the potential of the latest Integrated Multi-satellite Retrievals for GPM (IMERG) and TRMM products in estimating rainfall. Singapore is a tropical region. The evaluation was performed at daily, monthly, seasonal, and annual scales from 1 April 2014 to 31 January 2016. The findings showed that GPM IMERG had high performance in the representation of spatial rainfall variability and rainfall detection capability compared to the TMPA products. This study is one of the earliest assessments of GPM IMERG and evaluation of it with TMPA products in Singapore. (Mou Leong Tan., 2017).

1.3 Justification of the research

Gridded evaluated data can be used as an alternative source of precipitation data, especially for the regions that have sparsely rain gauge placement or where it is physically very challenging to get in-situ precipitation data. This research can help in addressing some of these problems. Due to topographical terrain differences in earth surface and effects i.e., orographic, the gridded data accuracy differs from region to region. So, evaluation of these datasets at different terrain is required to find out their accuracy at different spatial locations. According to my knowledge, no extensive research has been done so far. Daily data of gridded-based and in-situ precipitation data of 20-years analyzed by different comprehensive statistical parameters for different seasons as well as annually and monthly.

Chapter 2

MATERIAL AND METHODS

The study aims to understand and compare the performance of three high-resolution remotely sensed gridded precipitation datasets namely 'The Integrated Multi-Satellite Retrievals for GPM (IMERG)', 'The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA)' Products (3B42_V7 & 3B42_RT) for different seasons at daily time scales over Indus basin for the period of 2000 to 2020.

2.1 Datasets

The GPM IMERG product was produced by NASA to estimate surface precipitation over the globe. The data are derived from the half-hourly late expedited estimate of the daily collected rainfall. Rainfall data are produced by totaling the valid rainfall retrievals for the day and represented as (mm/day). While TRMM is designed to monitor rainfall and energy exchange in tropical regions of the world. TRMM data are important for understanding, identifying, and estimating the global water cycle. The first product of TRMM used in this study is, post real-time 3B42_V7 with a Root Mean Square precipitation-error (RMS) estimate and TRMM-adjusted, unified infrared precipitation in (mm/hr.), on a half-hourly temporal and a 0.25-degree spatial resolution. The 2nd TRMM 3B42_RT used is near real-time late precipitation at the same resolution. The main difference between these products is calibration, retrieval algorithms, and the time is taken to do so. Furthermore, to evaluate and validate these gridded datasets, ground data (observed) were accordingly acquired from Pakistan metrological department (PMD) an autonomous body tasked with weather forecasts and public warnings regarding weather for safety and fortification. PMD Data required were for years 2000-2020 for each station. As shown below

Table 1 Datasets used and their respective details.

Sr. No	Categories	Data Source	Resolution	Raw Format/ Time	Swath Coverage
1	**GPM IMERGE** (The Global Precipitation Measurement/ The Integrated Multi-Satellite Retrievals)	NASA / GES DISC	**0.1° X 0.1°** (11.1 Km)	**Daily.NC4** 2000-2020	885 Km
2	**TRMM-TMPA 3B42** (The Tropical Rainfall Measuring Mission)) (Post Real Time)	NASA / GES DISC	**0.25°X 0.25°** (27.75 Km)	**Daily.nc4.sub.nc4/** 2000-2020	780 km
3	**3B42_RT TRMM-TMPA** (The Tropical Rainfall Measuring Mission) (Near Real-Time)	NASA / GES DISC	**0.25°X 0.25°** (27.75 Km)	**Daily.nc4.sub.nc4/** 2000-2020	780 km
4	**Daily Rainfall Data**	Pakistan Meteorological Department	Ground data	**Daily mm/day** 2000-2020	

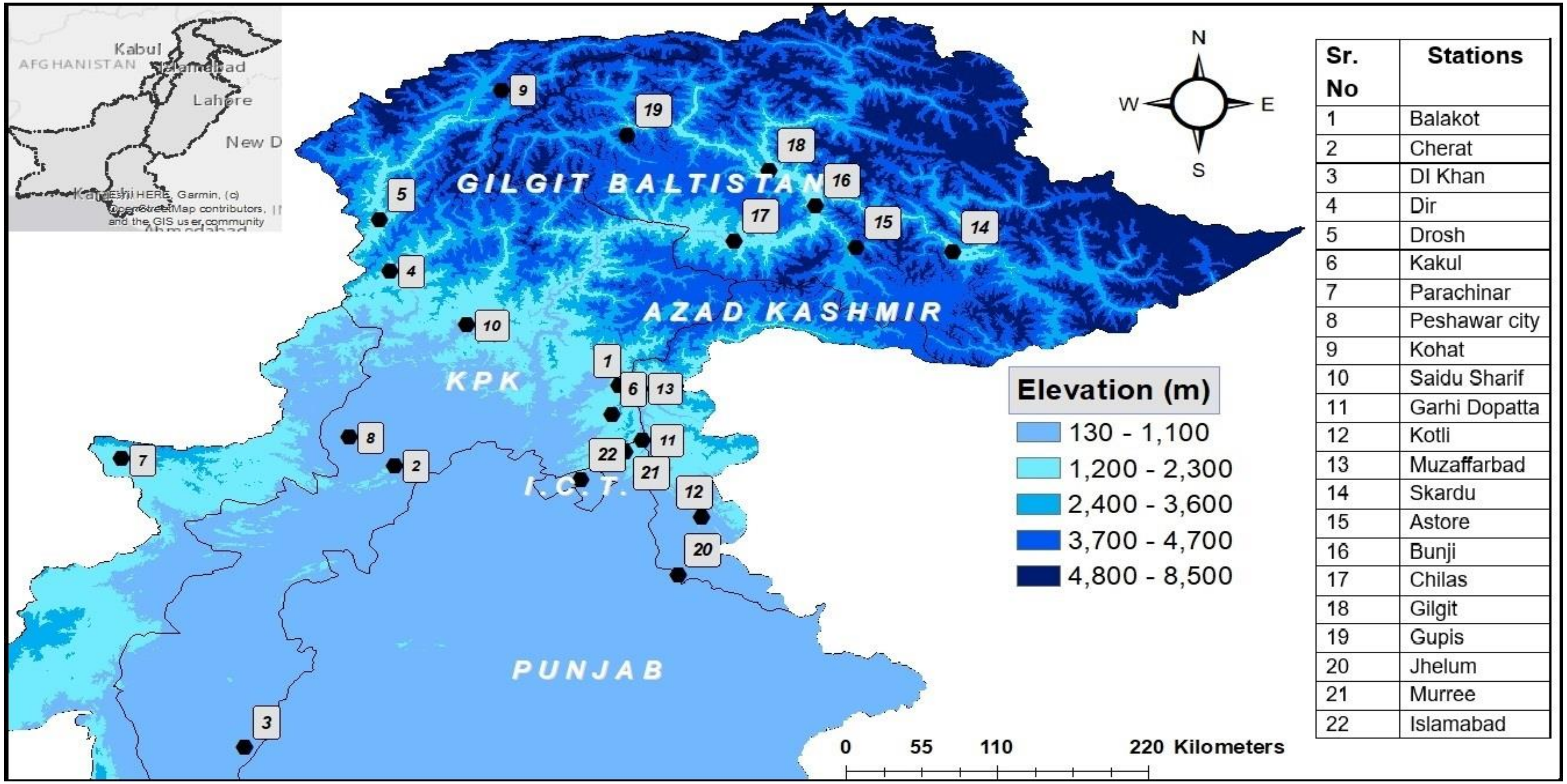

Figure 1 Study area map showing elevation ranges of PMD weather stations.

Table 2 Details of the selected PMD ground stations.

Sr. No	**Weather Stations**	**Long (dd)**	**Lat (dd)**	**Elevation (m)**
Region	**K. P. K**			
1	Balakot	73.35	34.383	995
2	Cherat	71.883	33.817	1372
3	DI Khan	70.9	31.81	164
4	Dir	71.85	35.2	1425
5	Drosh	71.783	35.567	1464
6	Kakul	73.3	34.183	1308
7	Parachinar	70.094	33.867	1726
8	Peshawar city	71.583	34.017	329
9	Kohat	72.583	36.49	3505
10	Saidu Sharif	72.35	34.817	970
Region	**AZAD KASHMIR**			
11	Garhi Dopatta	73.5	34	814
12	Kotli	73.888	33.452	610
13	Muzaffarbad	73.483	34.367	2303
Region	**GILGIT BALTISTAN**			
14	Skardu	75.536	35.336	2316
15	Astore	74.9	35.367	2394
16	Bunji	74.633	35.667	1372
17	Chilas	74.1	35.417	1251
18	Gilgit	74.334	35.919	1460
19	Gupis	73.4	36.167	2156
Region	**PUNJAB**			
20	Jhelum	73.74	33.04	287
21	Murree	73.383	33.917	2127
Region	**ICT**			
22	Islamabad	73.1	33.717	525

2.2. Study Area

The study area covers the high-altitude parts of Pakistan. Within the study area, 22, PMD stations were selected, and gridded data was downloaded covering these stations. The details of the stations along with their georeferenced coordinates and elevations are given in Table 2, the study area map is shown in Figure.1

2.3. Data Acquisition and Assimilation

Four types of datasets were required for this study, three of them were gridded data while the fourth was the in-situ precipitation data. Gridded data of TRMM TMPA products (3B42_RT & 3B42_V7) and GPM (IMERG) are available on NASA/GES DISC in raw format with extensions nc4.sub.nc4 & NC4 format. These precipitation data were required for all 22-stations for the period from 2000-2020 in an (mm/day). The data in a daily format was huge having 21,900 files and for downloading the process needs to be automated. For the automated downloading of both TRMM TMPA products ‘GNU Wget’ script was used. Wget is basically a command prompt command-line tool that can be used to download files from the internet. The sample script is available online, can be edited to the required format according to the need. Moreover, this script doesn’t work directly on windows built-in administrative command-line interpreter (CMD) for Microsoft Windows. To use this tool, an additional GNU Wget package needs to be installed. After the successful implementation of this script for the TRMM TMPA product, the data is automatically saved in the pre-defined folder mentioned in the script. The drawback of this script is that it doesn’t work for GPM (IMERG) datasets. The script and its implementation are displayed in Figure.3. To automate the downloading process for GPM daily precipitation data, ‘Google Mass downloader’ was used. It is basically a batch download manager with flexible filtering and mass renaming options that makes downloading files a lot easier & productive. It does this by extracting from the bulk links of web pages (advanced filtering system).

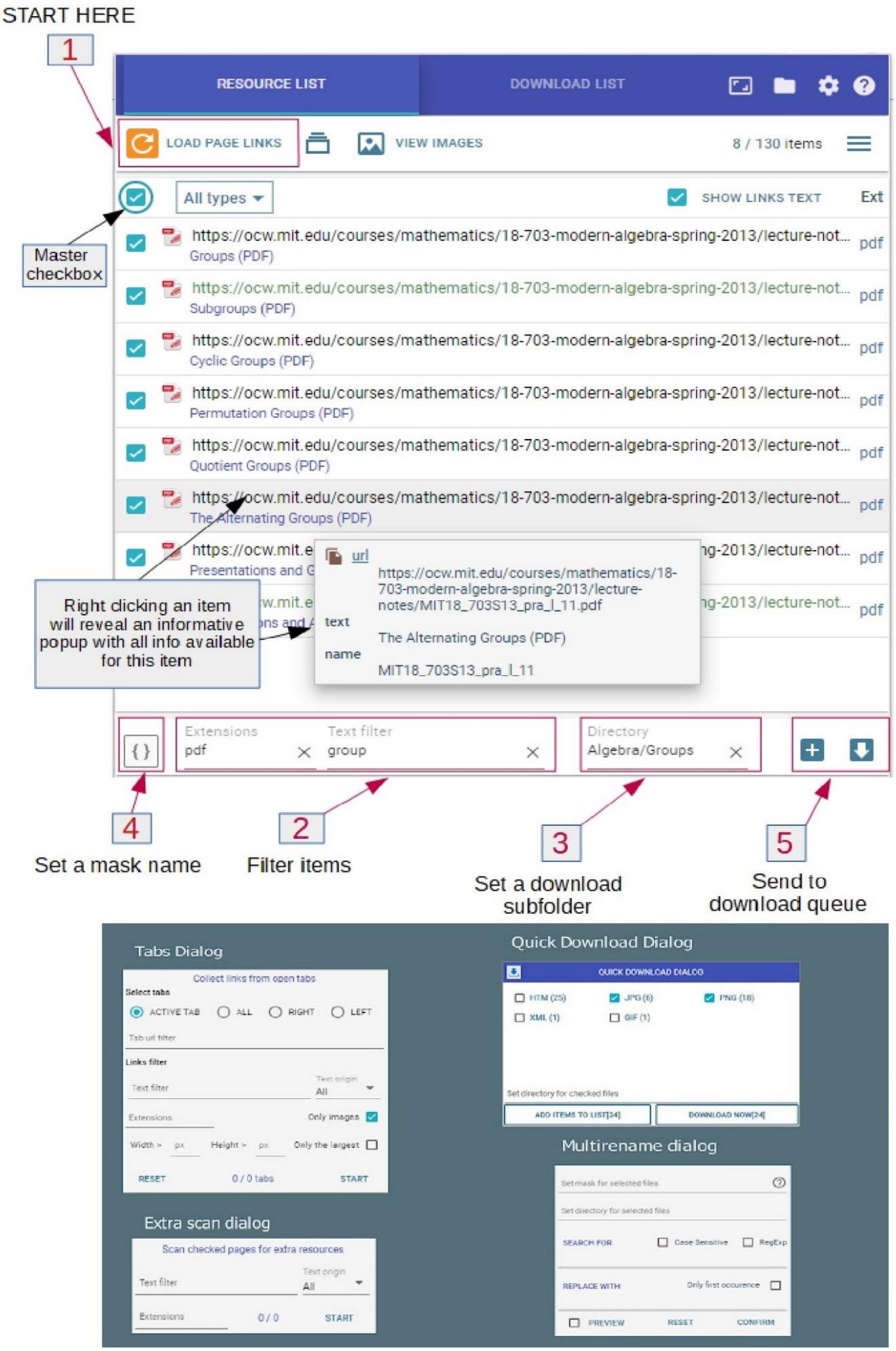

Figure 2 Displaying the process of Google Mass Downloader.

2.4 Data Processing

The most crucial and difficult part of this whole research was data pre-processing, and the very first step was to extract all the raw NC4 and nc4.sub.nc4 files of the gridded data sets. After that, all the extracted data must be converted into excel format. This too cannot be done manually as the raw NETCDF flies of TRMM products and GPM were collectively more than 23000.

As the extensions of the GPM (IMERG) and TRMM TMPA Products are not similar, so different models were needed to extract the data and then convert it to the required excel format. To automate this process two models were created on ArcMap. The usefulness of these models is that they can automatically pick a single raw NETCDF file from the given location extract it via prescribed tools, convert it to excel format, and then save it to a pre-defined desired destination and finally repeat the process for the next file in the specified folder.

The main difference between both models is that the model used for TRMM TMPA products basically makes raster layer-based files and then extract values as point data, further convert these into tabular form, save them as DBF format, and finally converts into excel readable format, While for GPM the model works with some major changes such as instead of making raster layer, the models directly convert the NETCDF files from the given folder to NETCDF table view format. Afterward, those values are exported to excel and saved as a CSV file in the desired location. Figures. 3 and 4 show the overall process flow of these models respectively. The in-situ precipitation data acquired from the Pakistan Metrological department also needed manual pre-processing. For extraction of data for selected stations.

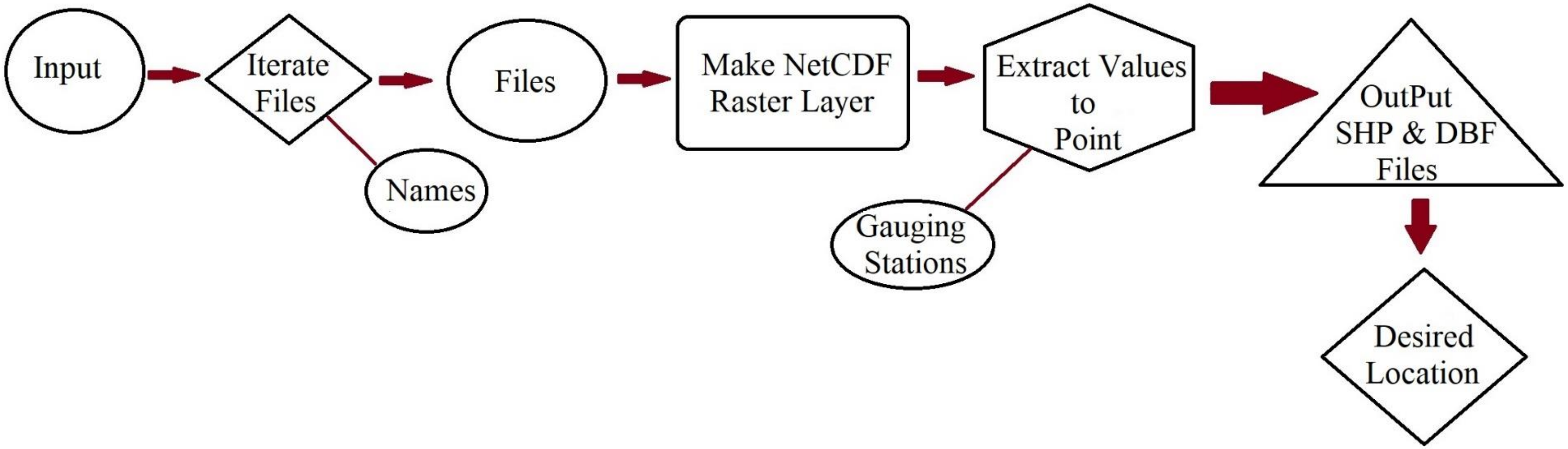

Figure 3 The model used for the extraction of TRMM TMPA products (3b42_V&7 and 3b52-RT). There were '15284' Files in total in nc4.sub.nc4 format.

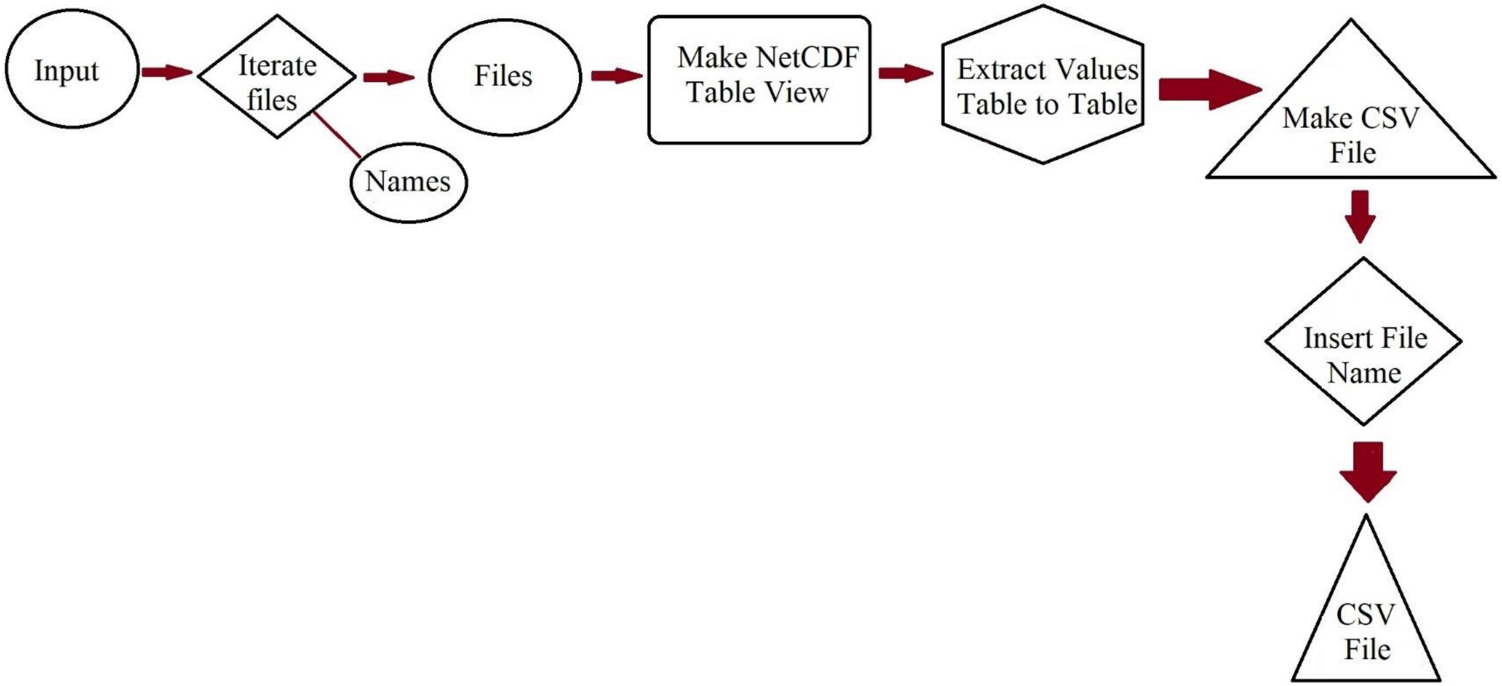

Figure 4 The model used for the extraction of GPM (IMERG). There were ‘7518’ Files in total in NC4 format.

From the whole observed precipitation data in raw CSV format, precipitation values against only 22 selected stations were extracted. The gridded data is basically in tile format for each day, and one tile covers very large and the data for further evaluation is needed in point format, there were more than 23000 files and each file contained more than four thousand entries against different locations. This too cannot be done manually, therefor excel tool POWER QUERY was used.

Power Query is a data transformation and data preparation engine. It comes with a graphical interface for getting data from sources and a Power Query Editor for applying transformations. The tool can extract, transform, load (ETL), and processing of data and saves a lot of data preparation time.

For the initiation of the power query, it must link the folder in the power query transformation window which contains the mixed CSV files. After the successful establishment of the link, it needs to follow all the transformations steps such as entering corresponding coordinates of the station and applying advance filtering in the power query editor manually for each station one at a time. The final step is to load the transformed data into power query ETL and save the transformed file. The whole process is repeated for each dataset, and for every station. The visual display of the power query process is given in figure.7.

After the successful implementation of all the above pre-processes, the files were further sub-divided into seasons (Monson, Pre-Monson, Winter) and annual format. It was done using defined criteria for advance filtering for above mentioned seasons and annual.

This process was repeated for the selected stations, for each gridded dataset, and on observed precipitation data of 20 years from 2000-2020. Figure.8 shows the process of advanced filtering.

2.5 Analytical Framework

The research methodology has been subdivided into four phases and the flowchart is shown in Figures 9 and 10.

2.6 Objectives

The objective of the study was ‘‘To compare and evaluate GPM IMERG & TRMM-TMPA products 3B42 version-7 & 3B42 Real-Time for precipitation estimation in terms of accuracy using comprehensive statistical parameters for different precipitation patterns (Monsoon, Western Distribution, and hybrid) with respect to in-situ precipitation data for Indus Basin’’.

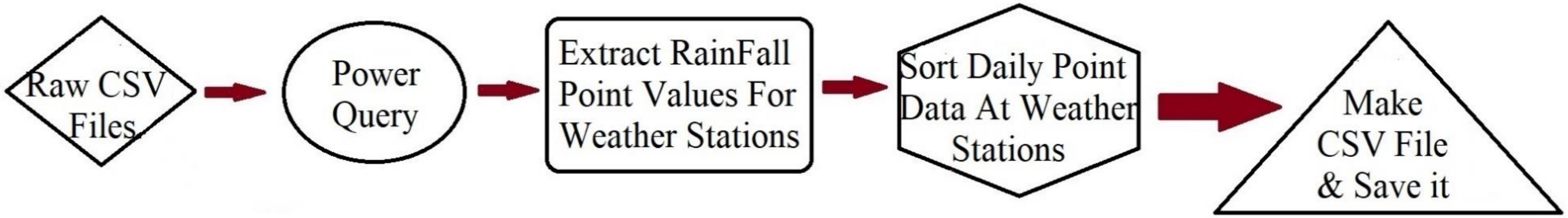

Figure 5 Visual display of Power Query process.

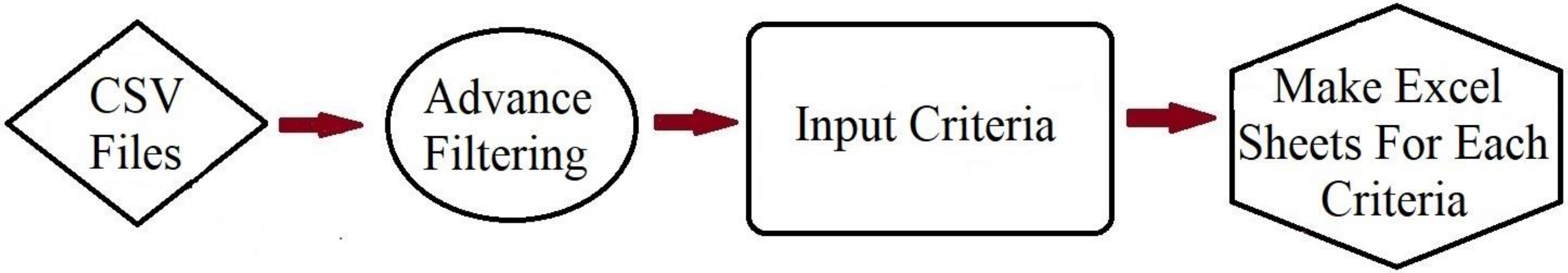

Figure 6 Displaying the Advanced Filtering process used for converting all the data into respective seasons.

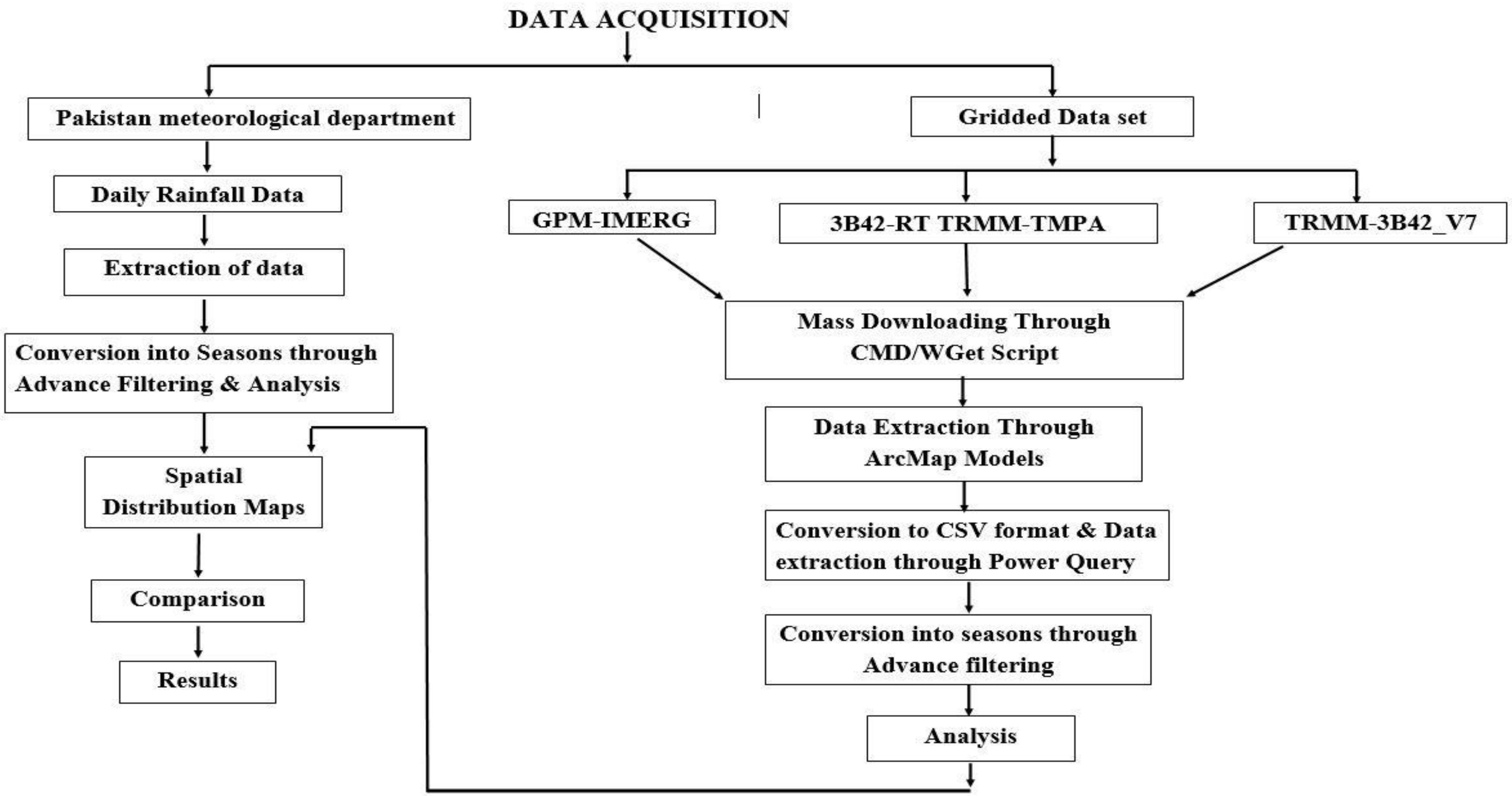

Figure 7 Methodology flowchart (Phase-1, 2 and 3)

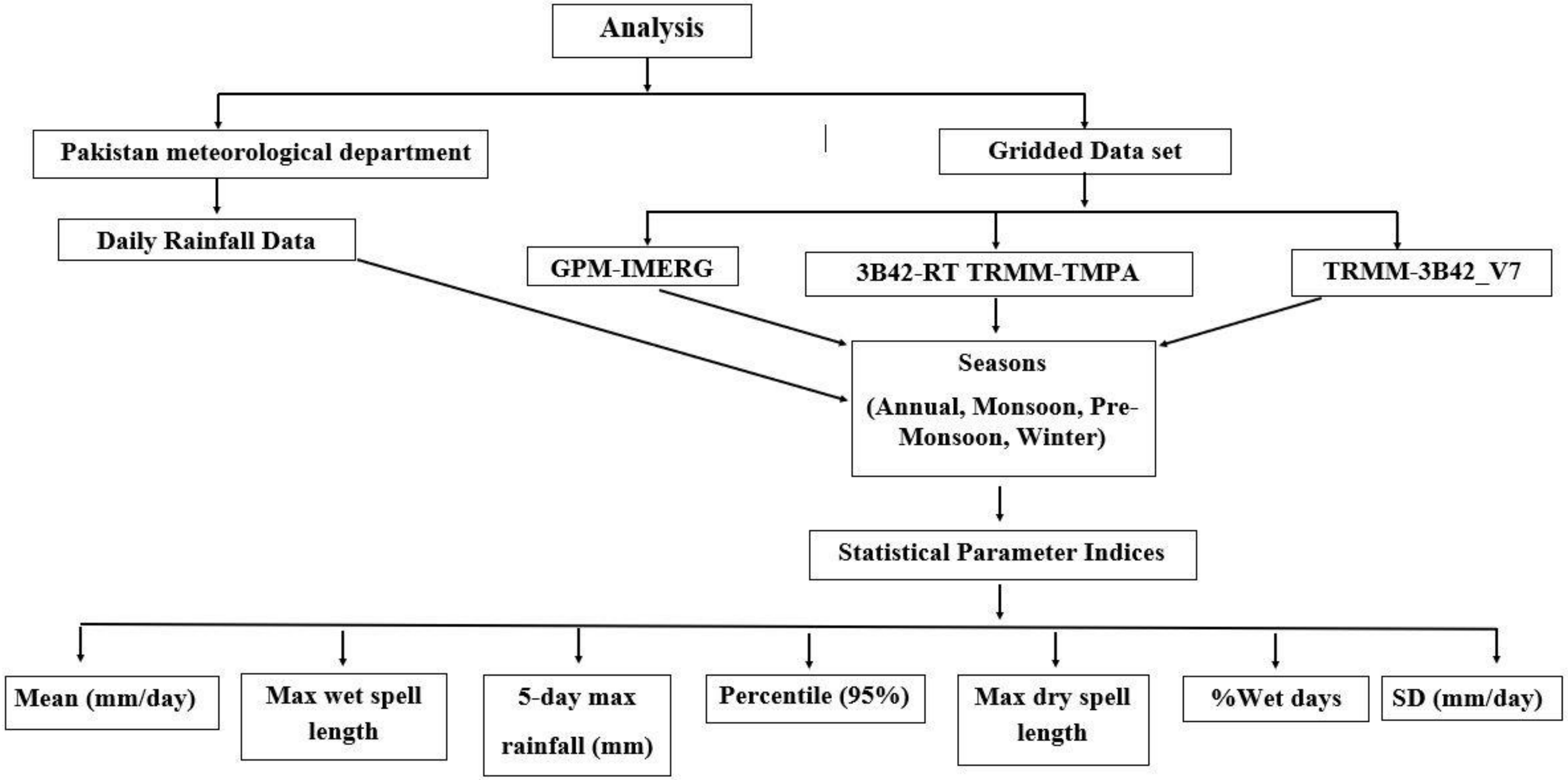

Figure 8 Methodology flowchart (Analysis)(Phase-4).

2.7 Precipitation Monthly Average Dataset

After the above-mentioned pre-processing, the next phase was to study and find out which out of the selected stations were affected by which type of precipitation pattern. To define the stations influenced by the respective seasonal categories (Monsoon, Western Distribution, and Hybrid) the daily data were used to calculate the monthly average of all 20-years at each station.

This process was applied for both gridded and observed data. To find out the precipitation distribution patterns on selected weather stations, the observed monthly averaged data were plotted for 11-stations per graph.

2.8 Selected Stations Categorization

With the help of these plotted graph, the stations with the comparatively highest peaks during June to September were considered as monsoon stations. Those which have high peaks during January to May were considered as Western Distribution stations while stations that had peak values in both seasons were marked as hybrid stations.

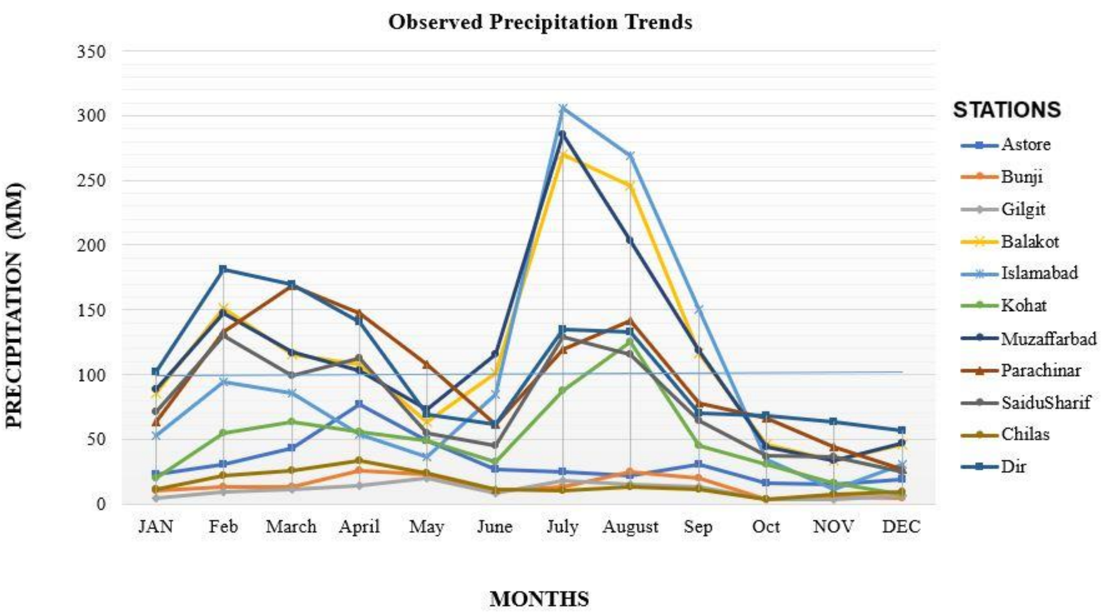

Figure 9 Graph displaying monthly averaged observed data of first 11- stations.

.

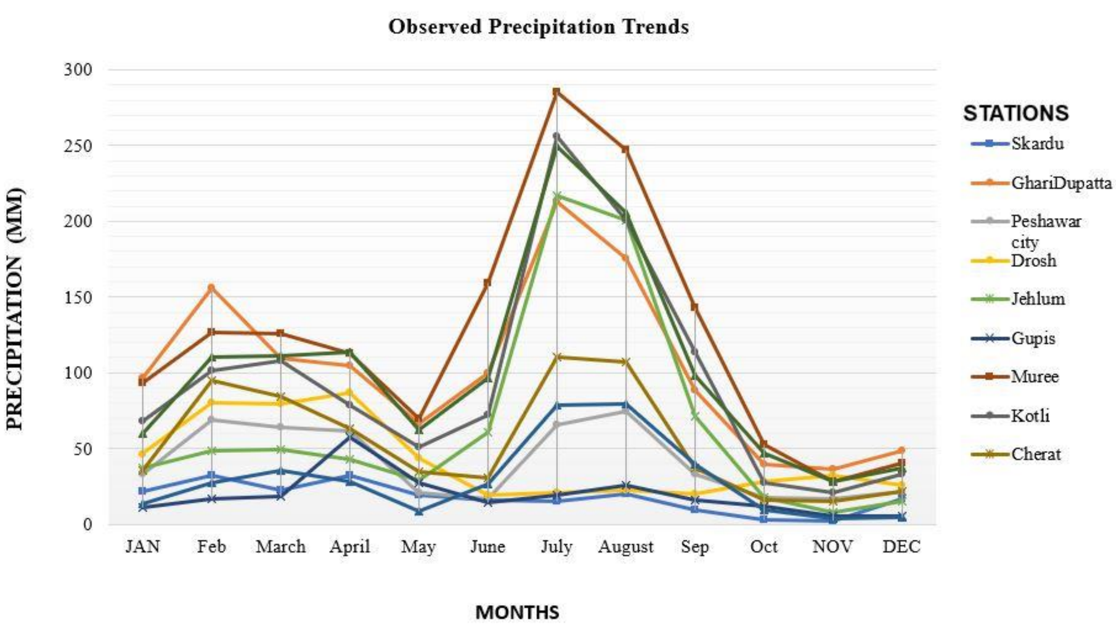

Figure 10 Graph displaying monthly averaged observed data of second 11- stations.

2.9 Monsoon (M) Affected Stations

Monsoon is a seasonal reversing wind accompanied by corresponding changes in rainfall. The seasonal changes in atmospheric circulation and precipitation are associated with annual latitudinal oscillation of the Intertropical Convergence Zone between its limits to the north and south of the equator. This climatic change happens from June to September. shown in Fig.12.

2.10 Western Distribution (WD) Affected Stations

Western disturbance is an extratropical storm originating in the Mediterranean region that brings sudden winter rain to the northern parts of the Indian subcontinent, it is a non-monsoonal precipitation pattern driven by the westerlies. The moisture in these storms usually originates over the Mediterranean Sea, the Caspian Sea, and the Black Sea. This climatic change happens from January to May. shown in Fig.13.

2.11 Hybrid (H) Distribution

The stations in the Indus basin which are affected by both WD and Monsoon precipitation patterns are considered hybrid stations. Mainly these regions get precipitations from both mentioned precipitation patterns.

2.12 Categorized map of the Study area

After the successful categorization of the selected stations, the map of the study area was remade to visualize the categorization process. As shown in Figure 14.

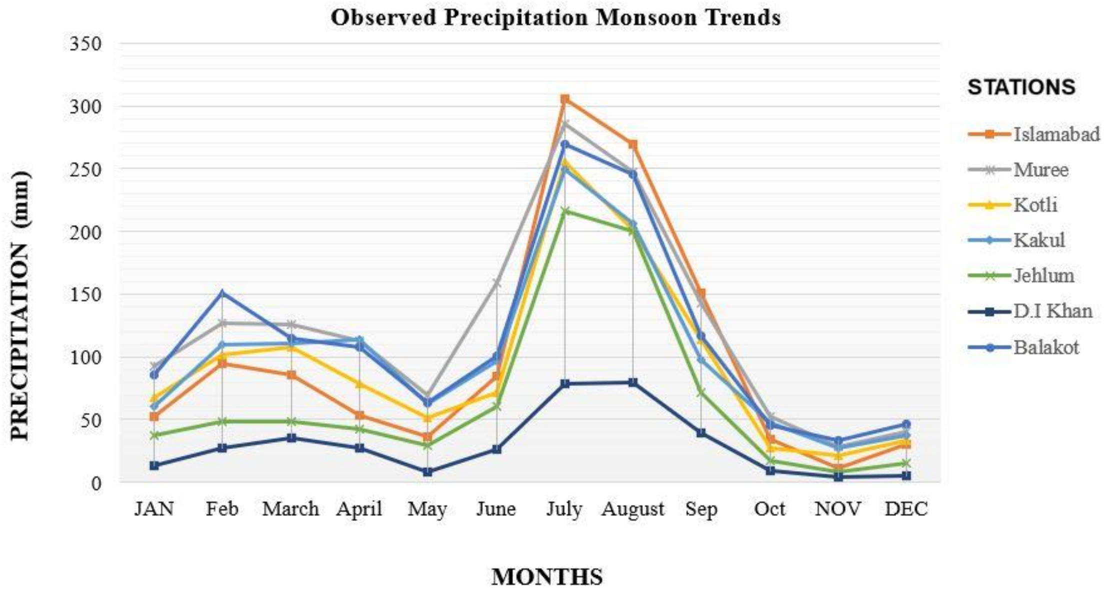

Figure 12 The graph of Monsoon affected stations.

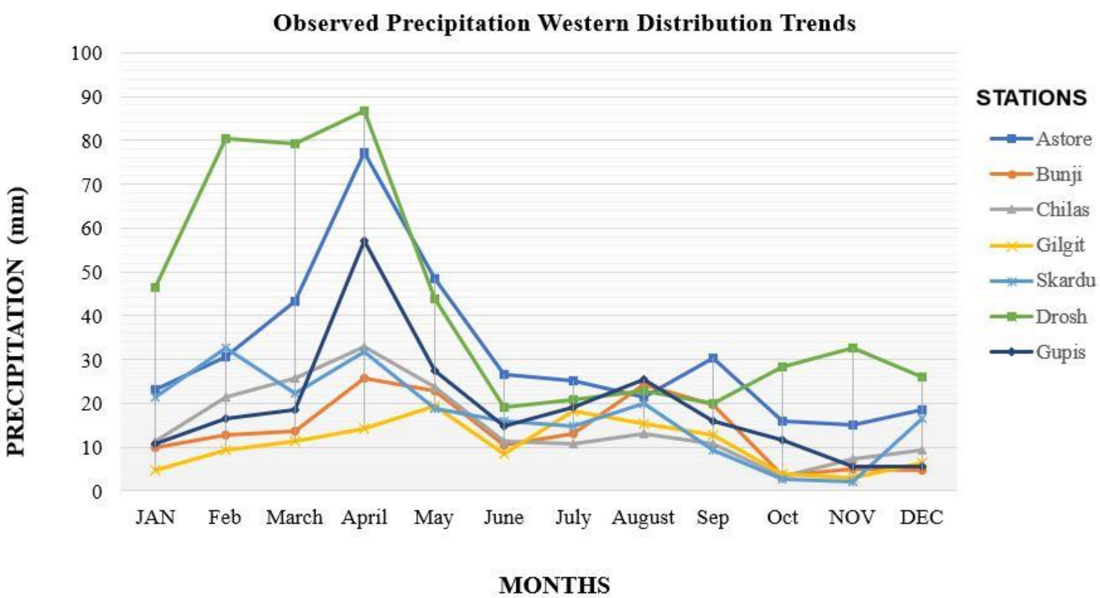

Figure 11 The graph of Western Distributed affected stations.

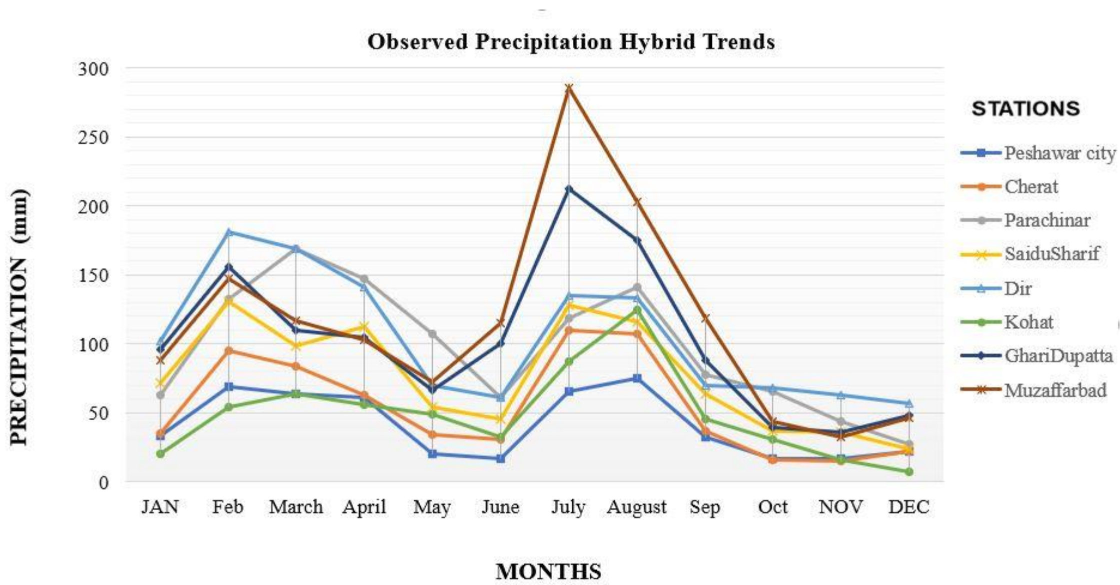

Figure 13 The graph of Hybrid stations.

Table 3 Ground Stations Categorization

Monsoon	**Western Distribution**	**Hybrid**
Islamabad	Astore	Parachinar
Balakot	Chilas	Dir
D.I Khan	Bunji	SaiduSharif
Muree	Gilgit	Kohat
Kotli	Skardu	Cherat
Kakul	Drosh	Peshawar city
Jehlum	Gupis	GhariDupatta
		Muzaffarbad

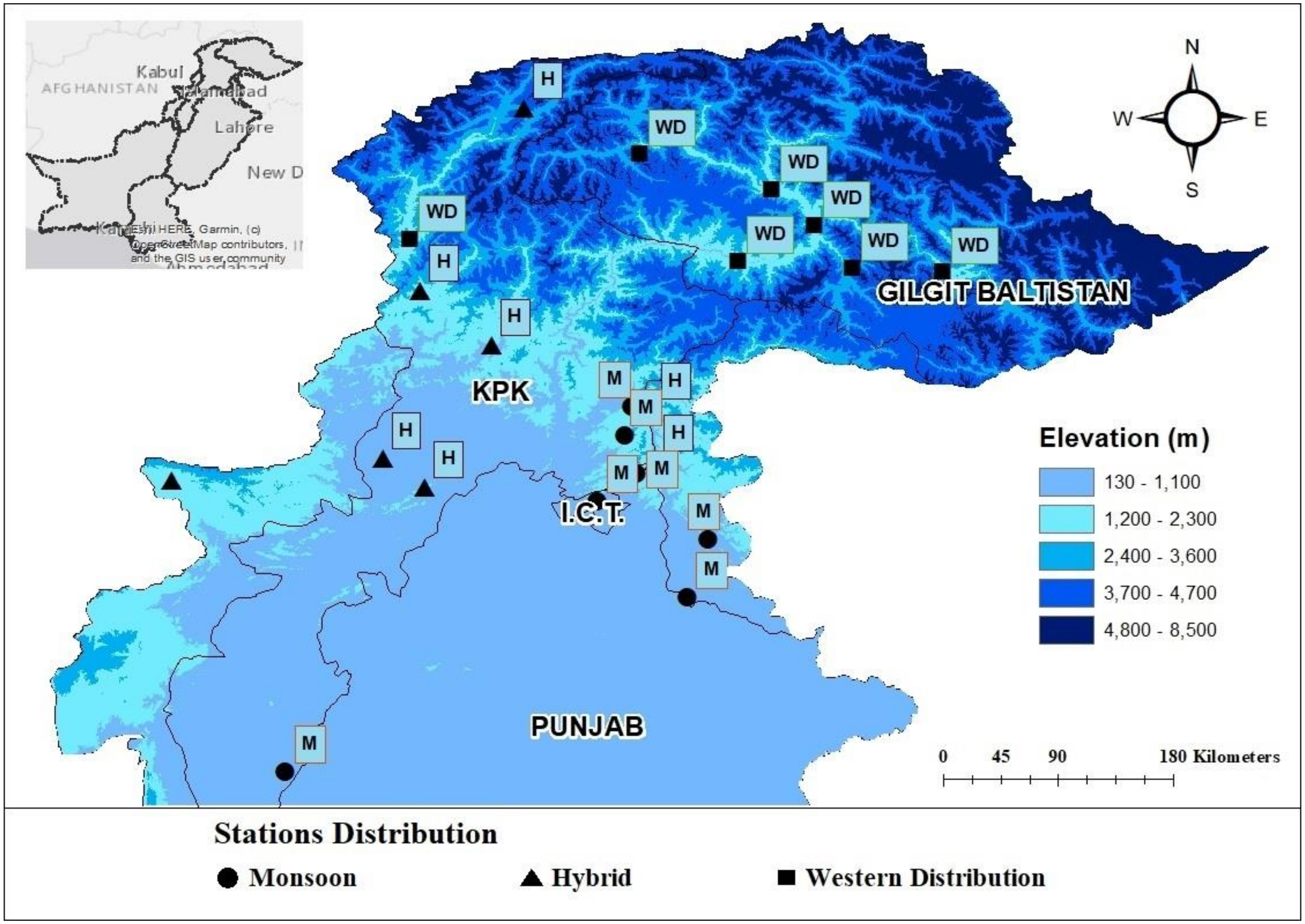

Figure 14 Map of the study area displaying Stations categorization

Chapter 3

RESULTS AND DISCUSSIONS

3.1 Calculation of Statistical Parameter

To achieve the objective of the study, the following seven comprehensive statistical parameters must be calculated. As shown below.

Table 4 Statistical parameters.

Sr. No	Statistical parameter	Formula
1	**Mean (mm/day)**	m = Sum of terms / No of terms Where a term is the sum of precipitation & number of terms is number of days of precipitation
2	**SD (mm/day)**	$\sigma = \surd\sum(Xi-\mu)2 / N$ where σ= rainfall SD, N= rainfall values, Xi= each value of rainfall, & μ= Rainfall mean
3	**Percentile (95%)**	n = (P/100) x N where N=number of values & P= percentile
4	**5-day max rainfall (mm)**	Code is written in Visual basic shown in Annex
5	**%Wet days**	Max value of precipitation in all of the days of rain (threshold=<1mm)
6	**Max dry spell length**	Total values of rain days with (threshold=>1mm)
7	**Max wet spell length**	Total values of rainfall (Threshold=<1mm)

3.2 Mean Precipitation (mm/day)

For the selected stations arithmetic mean was calculated individually in each season for 20-years. This process was repeated for the gridded and observed datasets. For the final mean values at each station, the in-situ observed mean was subtracted from the gridded dataset mean. This meant the values at stations that are closer to zero are more accurate, and the gridded data at such stations recorded accurately. Similarly, the negative and positive values between the mean indicate the over and under estimation of precipitation.

3.3 Precipitation Standard Deviation (SD) (mm/day)

The standard deviation is a measure of the amount of variation of a set of values. A low standard deviation indicates that the values tend to be close to the mean of the set, while a high standard deviation indicates that the values are spread out over a wider range. SD calculations were made for each station for 20-years for both datasets, like the mean above, i.e., the difference in SD values between gridded and observed data.

3.4 Precipitation Percentile (95%)

A percentile is a contrast of precipitation between particular precipitation and the rest of the precipitation, here the 95% percentile is a score below which a given percentage of scores in its frequency distribution falls at or below. After the successful calculation of the percentile of 20-years data at each station for each dataset, percentile calculations were made for each station for 20-years for both datasets, i.e., the difference in percentile values between gridded and observed data.

3.5 5-Day Max Rainfall (mm)

To calculate 5-days max rainfall values, the process was divided into three steps. The first step was to set the threshold value of 1mm/day precipitation. This simply meant to extract rainy days with precipitation measurement values. i.e., Days on which the rain greater than or equal to 1 mm/day occurred at each station. The second step was to find out 5 consecutive days on which the precipitation has occurred. Obviously, in 20 years timeline, there were many days where the stations had consecutive 5 days of rain which was greater than the 1mm threshold. The final step was to find out the highest value among these sets of consecutive 5 days of rainfall. This step was repeated for both in-situ and gridded datasets. This process was performed using visual basic-based script attached in Annex. After calculation of 5-days max rainfall at each station, the observed values were subtracted from the gridded values to find out the final values of the parameter.

3.6 Precipitation Percentage Wet Days

Percentage wet days is the percentage of days in 20 years period that received precipitation with a threshold value equal to or greater than 1mm/day. This step was repeated for every data set at each station and the final values were calculated. i.e., a difference of gridded and observed precipitation percentage wet days.

3.7 Precipitation Max-Dry Spell Length

Maximum dry spell length is simply the 20-years precipitation average of days at each station on which precipitation does not occur at all or occurs below the threshold value of 1mm/day. This too was done for each dataset and at each station. To find the difference, the values of the observed dataset were subtracted from the gridded dataset values.

3.8 Precipitation Max-Wet Spell Length

Maximum wet spell length is simply the 20-years precipitation average of days at each station on which precipitation does occur greater than or equal to the threshold value of 1mm/day. This too was done for each dataset and at each station. To find the difference, the values of the observed dataset were subtracted from the gridded dataset values. The results of all above mentioned statistical parameters are shown in Tables 5, 6, and 7.

Table 5 Results for Monsoon affected stations. The negative (−) and positive (+) values show over and under estimation respectively. Each identifier has been ranked (e.g., dark, light Gray, underlines and bold Italic ranked 1, 2, 3 respectively, with respect to observed data). 1 shows the best performing gridded data among all seasons.

Category	Dataset	Seasons	Mean (mm/day)	SD (mm/day)	Percentile (95%)	5-day max rainfall (mm)	%Wet days	Max dry spell length	Max wet spell length	Total of each identity	Overall Ranking
Monsoon	GPM	Annual	0.32	1.27	3.86	**-64.27**	**-24.02**	<u>8.78</u>	<u>-13.44</u>	9	**1**
		Winter	<u>1.56</u>	3.40	<u>4.09</u>	<u>11.71</u>	**29.74**	2.85	-3.92	8	
		Monsoon	0.91	1.27	3.92	**-70.70**	<u>-15.27</u>	3.05	-4.73	<u>7</u>	
		Pre-Monsoon	<u>1.78</u>	1.25	3.99	-7.63	-2.43	-2.72	-3.66	**4**	
	TRMM-3B42RT	Annual	1.28	3.08	6.85	**45.49**	**75.33**	-4.64	<u>18.52</u>	4	**3**
		Winter	<u>5.43</u>	<u>3.59</u>	<u>7.08</u>	**53.92**	<u>31.45</u>	-4.02	7.44	9	
		Monsoon	0.72	3.09	6.87	**54.31**	**84.29**	<u>5.11</u>	-0.56	<u>10</u>	
		Pre-Monsoon	<u>4.37</u>	3.02	6.95	<u>35.41</u>	9.12	<u>-5.77</u>	<u>8.92</u>	**5**	
	3B42_V7 TRMM-TMPA	Annual	0.50	1.26	2.39	**28.97**	<u>30.55</u>	<u>8.26</u>	<u>5.65</u>	9	**2**
		Winter	1.96	1.40	2.51	2.51	**44.73**	-1.08	<u>4.69</u>	7	
		Monsoon	1.39	1.25	-1.18	**30.83**	**41.79**	<u>8.22</u>	-3.70	<u>7</u>	
		Pre-Monsoon	<u>2.20</u>	1.22	-1.14	<u>15.30</u>	**-45.01**	0.31	2.84	**5**	

Table 6 Results for WD affected stations. The negative (−) and positive (+) values shows over and under estimation respectively. Each identifier has been ranked (e.g., dark, light Gray, underlines and bold Italic ranked 1, 2, 3 respectively, with respect to observed data). 1 shows the best performing gridded data among all seasons.

Category	Dataset	Seasons	Mean (mm/day)	SD (mm/day)	Percentile (95%)	5-day max rainfall (mm)	%Wet days	Max dry spell length	Max wet spell length	Total of each identity	Overall Ranking
WD	GPM	Annual	-0.30	0.24	-1.26	-12.86	-1.17	16.27	**-25.25**	11	**1**
		Winter	-1.18	0.26	-1.24	-3.08	-0.54	7.41	-9.53	7	
		Monsoon	-0.87	0.25	-1.22	-13.45	-12.24	1.11	-4.55	8	
		Pre-Monsoon	-0.61	0.25	-1.23	-5.33	**24.08**	5.54	-8	**2**	
	TRMM-3B42RT	Annual	-0.50	-1.78	-2.71	-11.75	-26.43	19.14	-9.41	9	**3**
		Winter	-1.99	-1.77	-2.63	4.80	**-61.22**	1.26	1.26	7	
		Monsoon	-1.39	-1.72	-2.57	-18.43	-19.34	14.10	-11.37	11	
		Pre-Monsoon	-1.01	-1.91	-2.61	9.57	-4.79	-1.71	3.74	**1**	
	3B42_V7 TRMM-TMPA	Annual	-0.19	-0.53	-0.40	12.92	-2.94	14.12	-4.43	10	**2**
		Winter	-0.78	-0.51	-0.83	4.83	**-37.00**	-0.60	3.18	6	
		Monsoon	-0.54	-0.52	-0.81	-1.56	-3.89	9.33	-6.64	11	
		Pre-Monsoon	-0.34	-0.53	-0.76	8.73	-2.05	3.22	-1.27	**1**	

Table 7 Results for Hybrid affected stations. The negative (−) and positive (+) values shows over and under estimation respectively. Each identifier has been ranked (e.g., dark, light Gray, underlines and bold Italic ranked 1, 2, 3 respectively, with respect to observed data). 1 shows the best performing gridded data among all seasons.

Category	**Dataset**	**Seasons**	**Mean (mm/day)**	**SD (mm/day)**	**Percentile (95%)**	**5-day max rainfall (mm)**	**%Wet days**	**Max dry spell length**	**Max wet spell length**	**Total of each identity**	**Overall Ranking**
Hybrid	GPM	Annual	-0.11	0.68	0.39	9.38	**30.75**	**13.01**	**-18.71**	9	**2**
		Winter	-0.38	0.67	0.40	3.17	11.21	3.55	-4.64	7	
		Monsoon	-0.38	0.65	0.32	-7.66	**38.13**	5.58	-5.97	8	
		Pre-Monsoon	-0.45	0.66	2.28	12.07	-18.37	4.31	-6.73	**4**	
	TRMM-3B42RT	Annual	0.52	1.75	2.70	24.90	**80.23**	10.57	4.01	9	**3**
		Winter	2.21	1.75	2.82	**35.70**	-11.09	-0.12	4.08	6	
		Monsoon	1.48	1.76	2.83	-6.81	**104.26**	11.18	-5.90	9	
		Pre-Monsoon	2.49	1.79	3.02	**34.95**	-3.17	-2.24	4.93	**4**	
	3B42_V7 TRMM-TMPA	Annual	0.37	1.15	2.06	**40.72**	**45.57**	13.39	1.22	10	**1**
		Winter	1.49	1.16	2.13	**44.14**	-2.17	-0.09	3.33	7	
		Monsoon	1.00	1.51	2.05	-3.53	**54.94**	10.41	-5.61	7	
		Pre-Monsoon	0.97	1.41	2.10	25.26	-27.04	1.35	1.34	**4**	

3.9 Monsoon Results

The results of Monsoon influenced station for each season were combined to find out the overall accuracy of the gridded datasets. In the case of monsoon, GPM IMERG outperforms the other two datasets. As GPM estimated better than TRMM 3B42_RT and 3B42_V7 at Monsoon influenced, accordingly it is ranked 1^{st}. After GPM, TRMM-3B42_V7 performs better than TRMM 3B42_RT so ranked as 2^{nd}. while TRMM-3B42_RT performance is the lowest rank as 3^{rd}.

3.10 Hybrid Results

All the results of the Hybrid influenced station for each season were combined to find out the overall accuracy of the gridded datasets. In the case of Hybrid TRMM_V7 outperforms the other two datasets. In terms of ranking the 3B42_V7 predicts better than TRMM 3B42_RT and GPM IMERG at hybrid stations, so 3B42_V7 is given rank 1^{st}. After 3B42_V7, GPM IMERG performs better than TRMM 3B42_RT so it’s been given rank 2^{nd}. While TRMM-3B42_RT performs slightly less at hybrid stations comparatively, it has been assigned the lowest rank 3^{rd}.

3.11 Western Distribution Results

All the results of the Western Distribution influenced station for each season were combined to find out the overall accuracy of the gridded datasets. In the case of Western Distribution, GPM IMERG outperforms the other two datasets.

In terms of ranking the GPM predicts better than TRMM 3B42_RT and 3B42_V7 at Western Distribution influenced stations, So GPM is given rank 1st. After GPM, TRMM-3B42_V7 performs better than TRMM 3B42_RT so it’s been given rank

2nd. While TRMM-3B42_RT also performs the lowest in Western Distribution comparatively, it has been assigned the lowest rank 3rd.

3.12 Spatial Distribution Maps

Maps are an integral part of the process of spatial data handling. They are used to visualize spatial data, to reveal and understand spatial distributions and relations. Spatial distribution is the arrangement of a phenomenon across the Earth's surface. Such graphical display arrangement is an important tool for geographical and environmental statistics. Maps were made by using the IDW interpolation technique. As the data used in this study are continues, bilinear interpolation was implemented to visualize and further analyze the results of estimated statistical parameters.

For this study, precipitation spatial distribution maps of all the calculated parameters at each station for each season of gridded datasets were made to visualize the variance and their performance. These maps can help find what change occurred in which season and for what dataset. By comparing the maps simultaneously of the same season and same statistical parameter for all the gridded datasets used in this study, their performance for a particular parameter and season can be easily visualized.

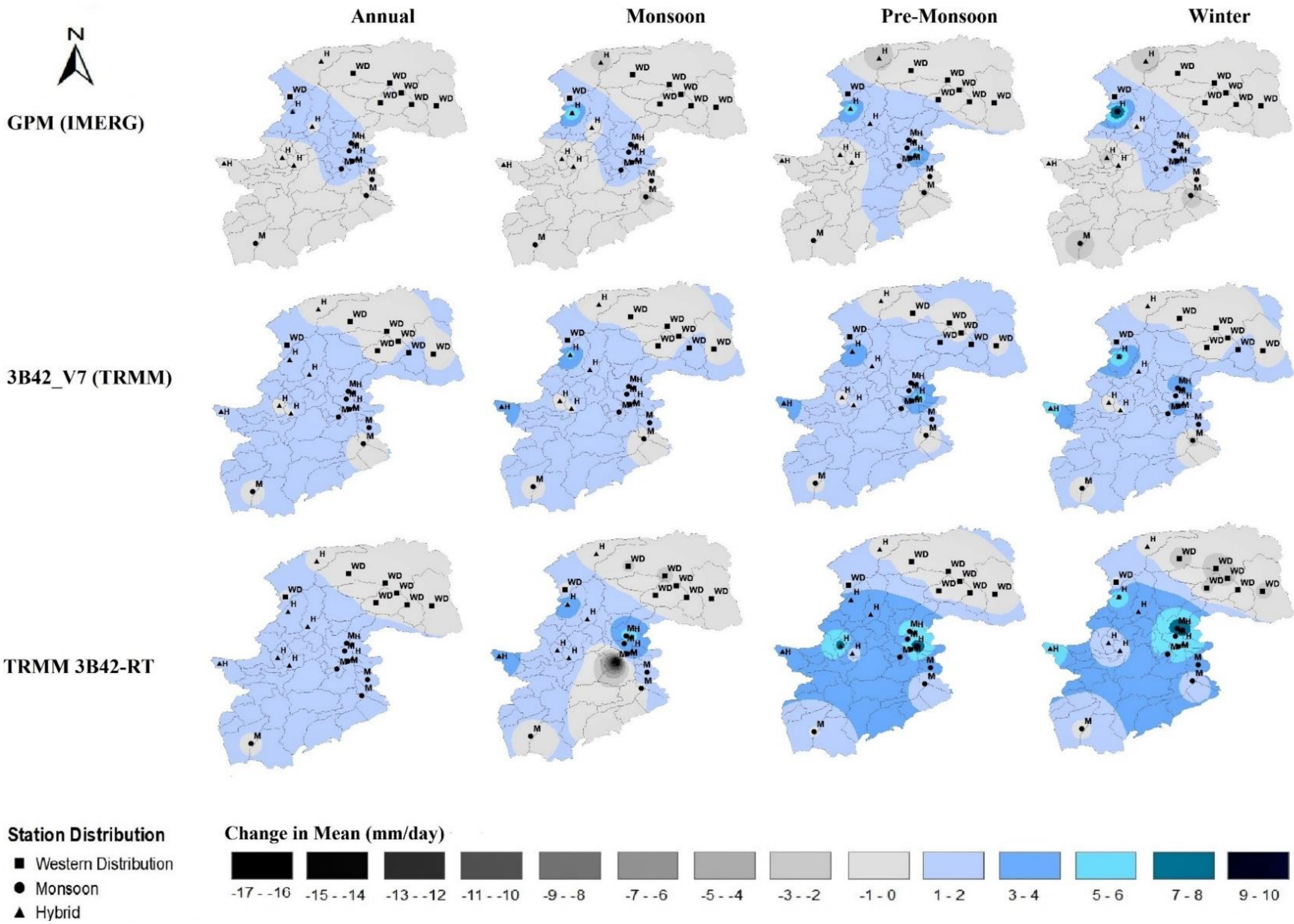

Figure 15 Spatial and seasonal change in Mean precipitation (mm/day), relative to the reference period (2000–2020)

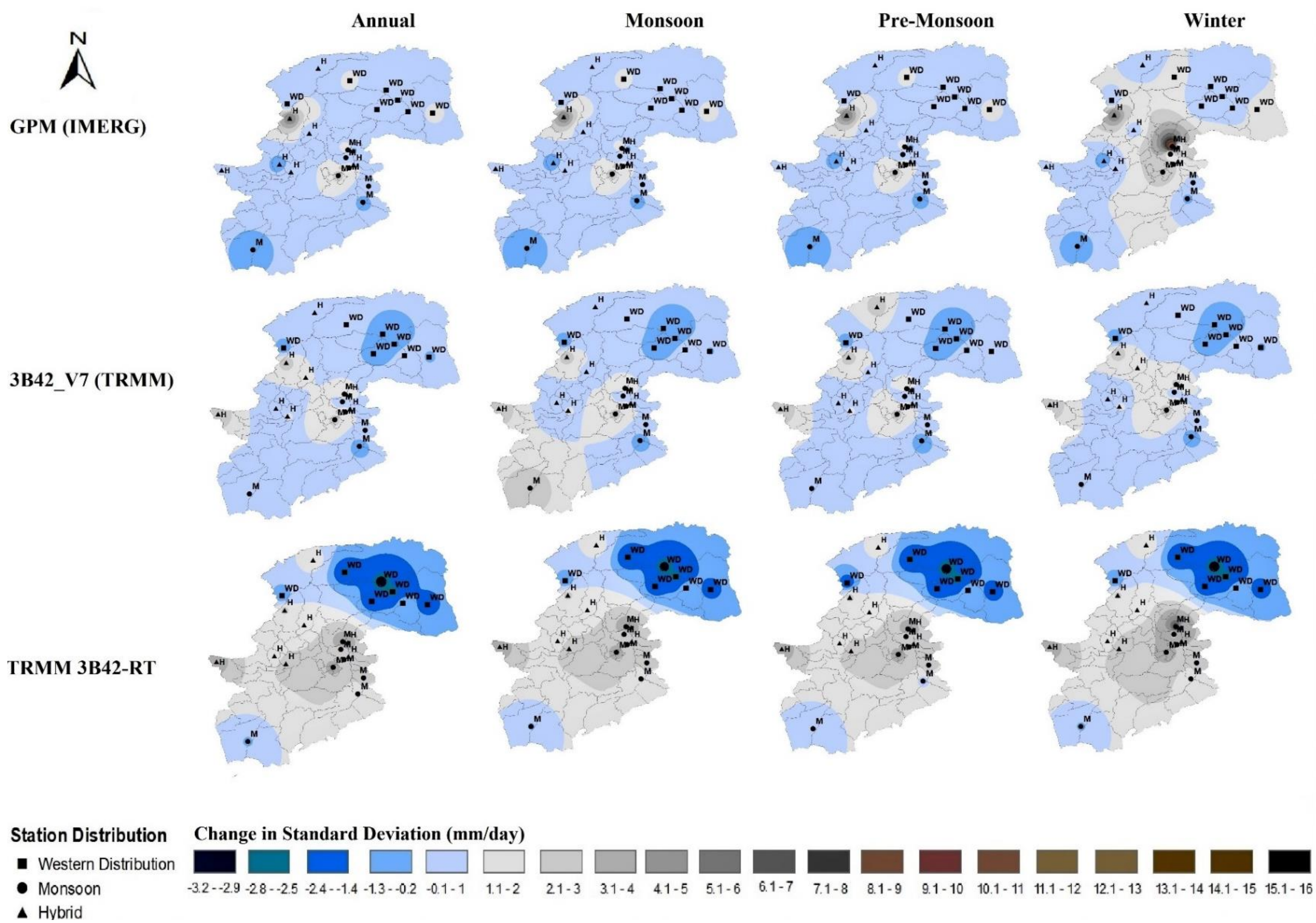

Figure 16 Spatial and seasonal change in SD precipitation (mm/day), relative to the reference period (2000–2020)

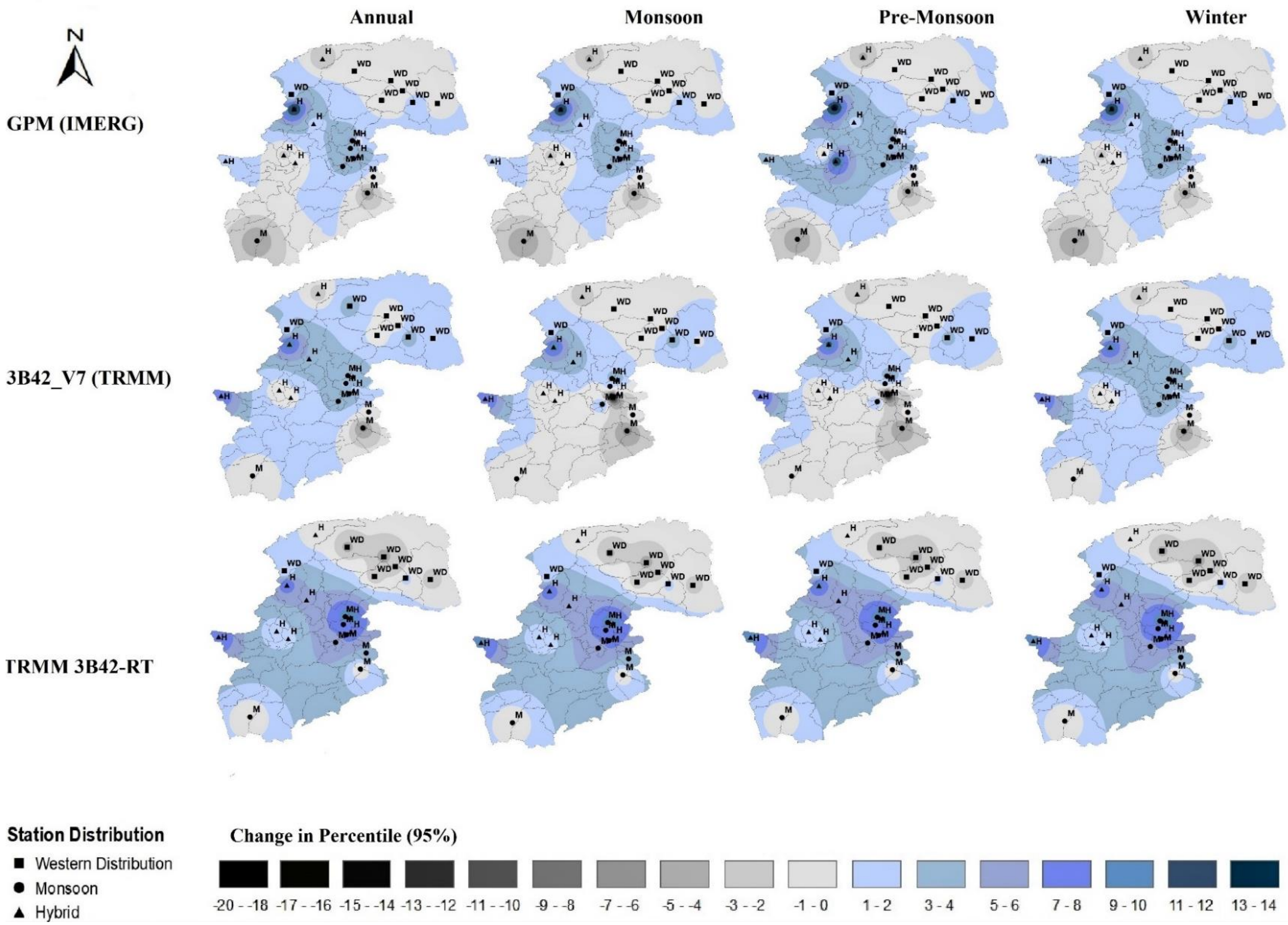

Figure 17 Spatial and seasonal change in 95% Percentile rain (mm/day), relative to the reference period (2000–2020)

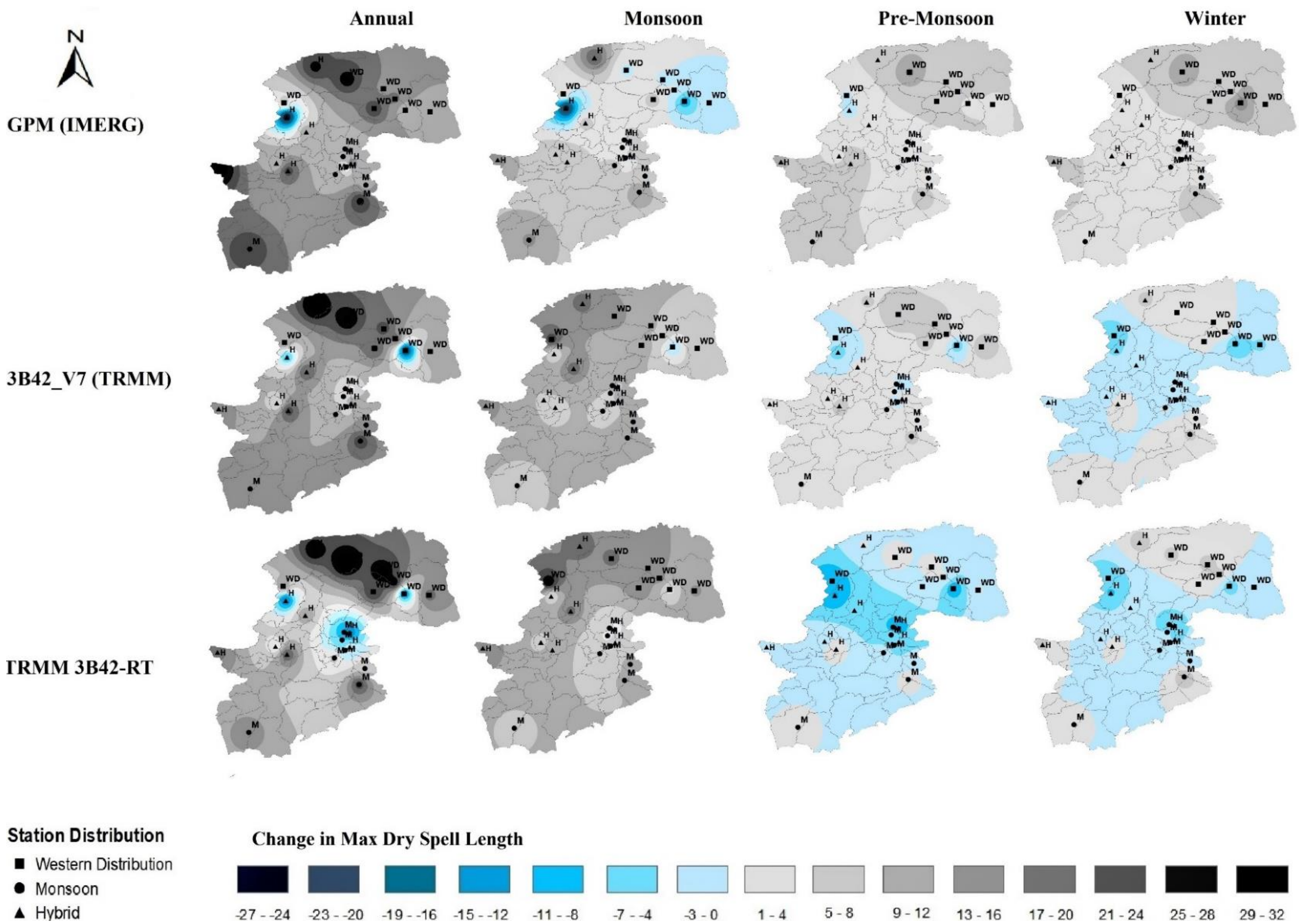

Figure 18 Spatial and seasonal change in Max Dry Spell length, relative to the reference period (2000–2020)

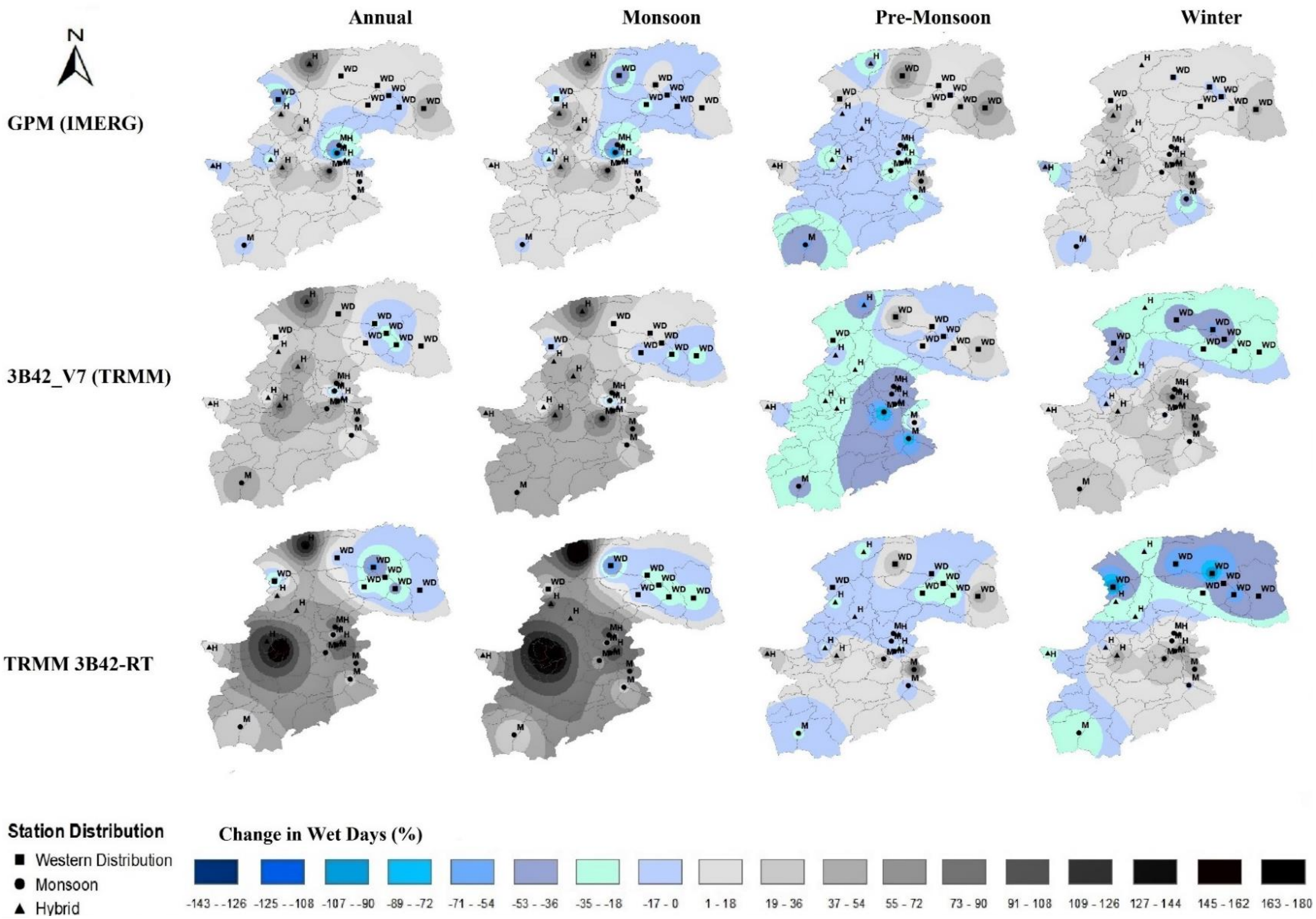

Figure 19 Spatial and seasonal change in Percentage Wet Days, relative to the reference period (2000–2020)

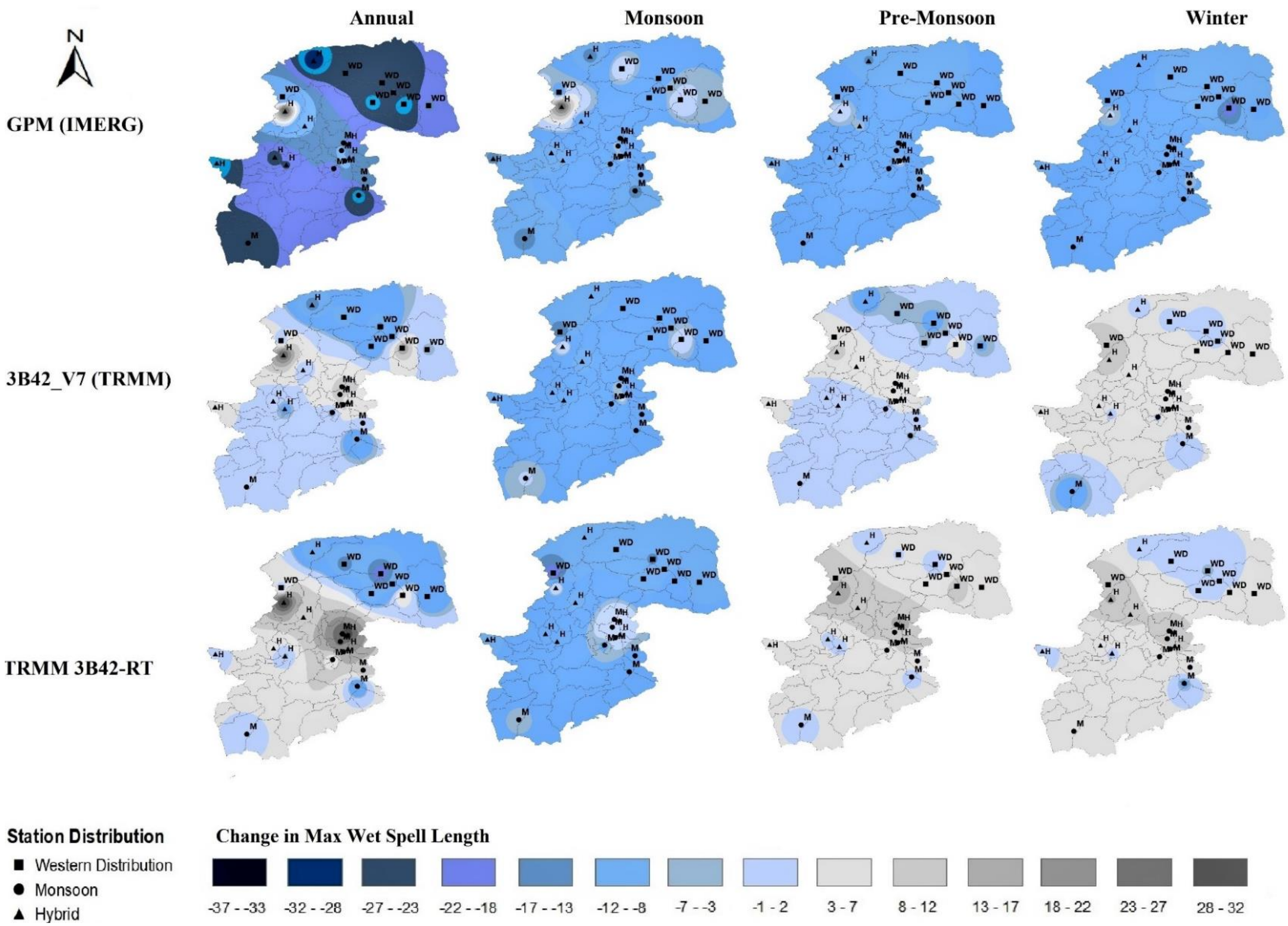

Figure 20 Spatial and seasonal change in Max Wet Spell length, relative to the reference period (2000–2020)

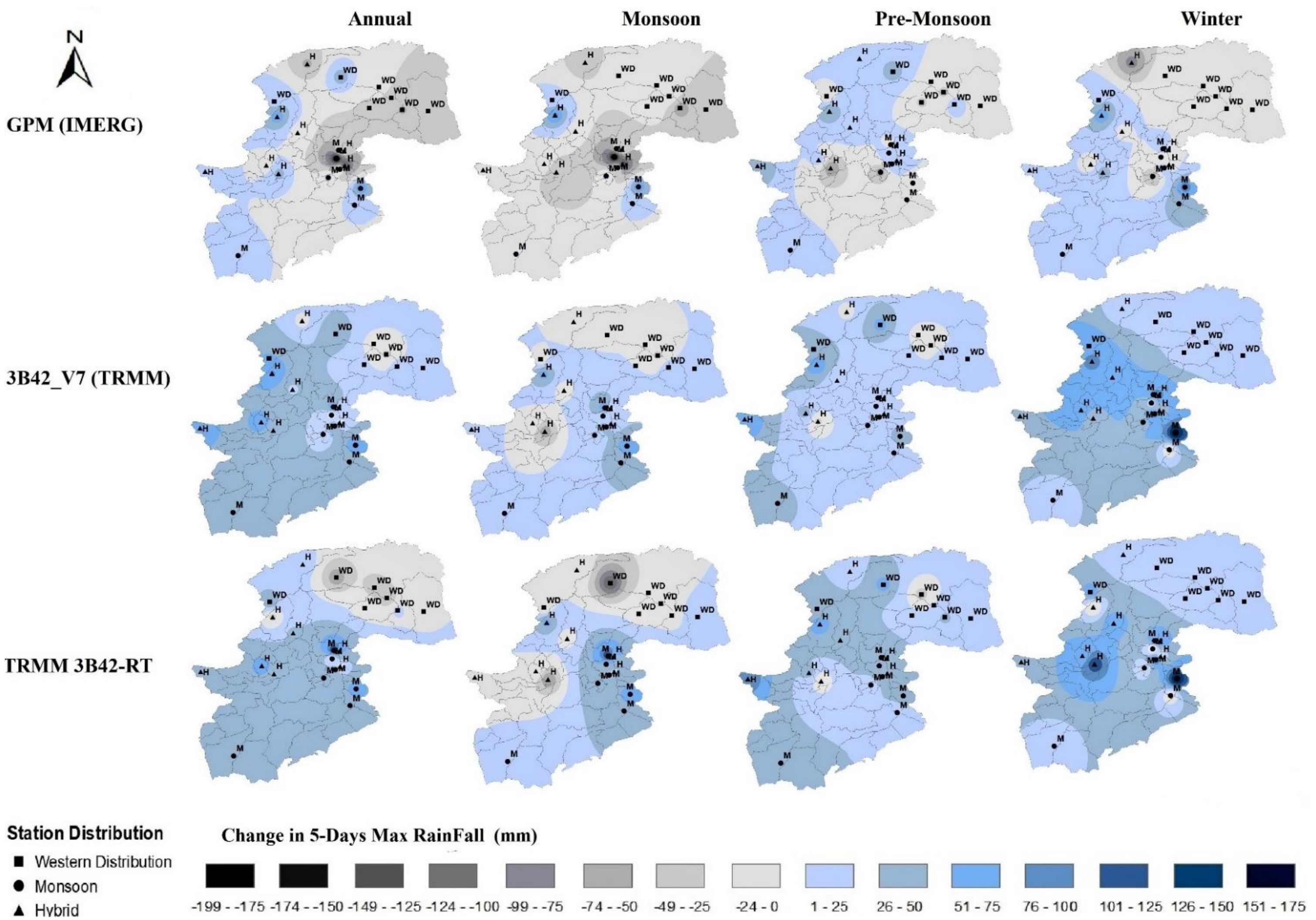

Figure 21 Spatial and seasonal change in 5-Days Max rainfall (mm/day), relative to the reference period (2000–2020)

3.13 Results of Spatial Distribution Maps

Figure 15 to 21 shows the results of the parameter at each season for gridded datasets. The maps show the spatial variations of precipitation at each station and display the precipitation patterns it was affected by. The lighter the color in the above-mentioned figures, the better the performance of the gridded dataset in that region and vice versa. The variation in the blue color tone shows positive values displaying under predicted regions while the gray color tone represents negative values that display the overestimation zones in the map. Statistical parameter maps of each season were further compared with each other to find the better performing gridded dataset and assignment of rank accordingly.

3.14 Summary of the Results

The above-mentioned comprehensive statistical parameters can capture the small spatial variability of precipitation. These parameters were calculated individually at each station for gridded datasets and were further averaged in accordance with the precipitation patterns they were affected by (Monsoon, Western Distribution & Hybrid). The results reveal that gridded data at each station record over or underestimated precipitations, which further needs to be adjusted by area-wise bias correctness factors. The statistical parameters map analysis concluded that the gridded dataset performance is dependent on topographical conditions. The performance of gridded datasets was significantly reduced at WD respective stations due to orographic effects and altitude dependency.

3.15 Pearson Correlation

The Pearson correlation measures the strength of the linear relationship between two variables. This relationship is considered strong when their r value is larger than 0.7 and vice versa. It was calculated for the different variables used in the study and shown below in Table 8.

Pearson correlations results reveal that the GPM IMERG performance on the monsoon and western disturbance affected stations was comparatively highest while for the station which was affected by hybrid precipitation patterns TRMM 3B42_V7 performance was recorded better among the selected gridded datasets.

Table 8 Precipitation Pearson Correlation analysis on Daily time Scale (20-years)

Datasets	**Seasons**	**Pearson Correlation**
GPM IMERG	Monsoon	0.882
	Western Distribution	0.543
	Hybrid	0.747
TRMM 3Bb42_V7	Monsoon	0.755
	Western Distribution	0.424
	Hybrid	0.840
3B42_RT	Monsoon	0.523
	Western Distribution	0.350
	Hybrid	0.708

Chapter 4

CONCLUSION AND RECOMMENDATIONS

4.1 Conclusion

The study was carried out to compare and identify the performance or the quality of gridded precipitation datasets for different seasons and variable locations spread over a large area. Traditionally, these are measured at weather stations on an hourly and daily basis, and later totaled and averaged on a daily, monthly, and yearly basis. Basically, the weather station is a combined system of integrated components that can measure, record, and sometimes transmit weather data.

The major advantage of using a weather station is the accuracy of measurements. A demerit of using traditional weather station data is reliance on a person responsible for these measurements. This problem can be solved by the use of an automatic weather station which can eliminate the risk of human error. In Pakistan most of the stations are manually operated maybe because of the budgetary constraints as well as the non-availability of required facilities in far-flung areas of the country.

Gridded data are the future, as there are no physical barriers for obtaining such data through remote sensing satellite-based technology, that can help acquire optical, precipitation, and thermal data of any place at a very high spatial grid and temporal frequency to offset the vastly spread ground station data. Gridded data are still in developing stages, such as precipitation retrieval algorithms as well as the satellite resolutions which are continuously improving. It may take some more time but sooner or later will be supplemented by the old traditional ground station-based precipitation measurement techniques.

The primary objective of this research was to compare and evaluate “The Integrated Multi-Satellite Retrievals for GPM (IMERG), The Tropical Rainfall Measuring Mission (TRMM)(TMPA) products 3B42_Version-7 & 3B42_Real-time” through observed data acquired from the Pakistan Meteorological Department (PMD) ground stations. These three datasets on daily basis are analyzed for three different seasons (Monsoon, Pre-Monsoon, Winter) as well as annual and monthly basis. The comparison, validation, and accuracy assessment of these datasets were based on seven comprehensive statistical parameters using the 20 years of daily precipitation data for 22-stations. These stations were divided in terms of precipitation patterns as (WD, Monsoon, Hybrid). Evaluation is necessary due to the topographical variability even at the local scale.

The result of the study shows that in the case of Western distribution and Monsoon influenced stations overall GPM_IMERG outperforms the other datasets and the results of 3B42_V7 at hybrid stations outperform other data sets. While the 3B42_RT performance was the lowest of the three at each station’s distribution.

Above in view, it can be concluded that each gridded dataset requires some correction and improvements. As GPM IMERG overall estimate of precipitation in case of Monsoon season is lower than actual, and for Pre-monsoon and winter, it is overestimating precipitation, so it requires both negative and positive correctness factors, respectively. Similarly, TRMM 3B42_V7 overestimated precipitation in all seasons thus requires a negative correctness factor. TRMM_RT overestimated precipitation in a monsoon while underestimating it in all other seasons but as this study result suggests that due to the lowest accuracy of this dataset among other analyzed dataset, it is highly unsuitable for estimating precipitation.

The study provides a better evaluation of gridded estimated precipitation, which is comparable and consistent with the corresponding observed PMD ground data. The estimated precipitation distribution can effectively serve as a basis for bias correction of any gridded precipitation products for the study area, however, it will be different for different topographical regions.

4.2 Recommendations for Further Research

To arrive at a more holistic picture of the accuracy and the accurate area-wise bias correctness factors for the gridded datasets, there is a need of having densely placed ground rain gauge stations to give more accurate area-wise bias correctness factors. However, huge budgetary requirements may be a constraint to building more ground stations.

REFERENCES

1. Anjum, M. N., Ding, Y., Shangguan, D., Ahmad, I., Ijaz, M. W., Farid, H. U., … Adnan, M., (2018) Performance evaluation of latest integrated multi-satellite retrievals for Global Precipitation Measurement (IMERG) over the northern highlands of Pakistan.
2. Bumke, K.; König-Langlo, G.; Kinzel, J.; Schröder, M. HOAPS., (2016) ERA-Interim precipitation over sea validation against shipboard in situ measurements. Atmos. Meas. Tech. 2016, 9, 2409–2423.
3. Muhammad Azmat, Muhammad Uzair Qamar, Christian Huggel a, Ejaz Hussain., (2018) Future climate and cryosphere impacts on the hydrology of a scarcely gauged catchment on the Jhelum River basin, Northern Pakistan. 2018,961-976.
4. Jingjing Yu, Xiao-Feng Li, Elizabeth Lewis, Stephen Blenkinsop & Hayley J. Fowler., (2020) UK high-resolution gauge–radar–satellite merged hourly precipitation analysis dataset. 2919–2940 (2020).
5. Beck HE, Wood EF, Pan M, Fisher CK, Miralles DG, van Dijk AIJM, McVicar TR, Adler RF., (2019) MSWEP V2 global 30-hourly 0.1° precipitation: methodology and quantitative assessment. Bull Am Meteorol Soc 100:473–500. https://doi.org/10.1175/BAMS-D-17-0138.1
6. Guoqiang Tang, Yixin Wen, YaoYao Zheng, Di Long, and Yang Hong., (2016) hydrology remote sensing, Capacity Building for Sustainability and Resilience.
7. Dahri, Z. H., Ludwig, F., Moors, E., Ahmad, B., Khan, A., & Kabat, P., (2016). An appraisal of precipitation distribution in the high-altitude catchments of the Indus basin. Science of The Total Environment.
8. Beck HE, van Dijk AIJM, Levizzani V, Schellekens J, Miralles DG, Martens B., (2017) MSWEP: 3-hourly 0.25 degrees global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. Hydrol Earth Syst Sci 21:589–615. https://doi.org/10.5194/hess-21-589-2017
9. Li, L., Ngongondo, C. S., Xu, C. Y., & Gong, L., (2012). Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa. Hydrology Research, 44(5), 770-788.
10. Javanmard, S., Yatagai, A., Nodzu, M. I., BodaghJamali, J., & Kawamoto, H., (2010) Comparing high-resolution gridded precipitation data with satellite rainfall estimates of TRMM_3B42 over Iran. Advances in Geosciences, 25, 119-125
11. Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., (2014) The WFDEI meteorological forcing data set: WATCH forcing data methodology applied to ERAinterim reanalysis data. Water Resour. Res. 50, 7505–7514.
12. Ashouri et al., (2015) H. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies Bull. Am. Meteorol. Soc., 96 (1) (2015), pp. 69-83, 10.1175/BAMS-D-13-00068.1
13. Teutschbein, C., Seibert, J., (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J. Hydrol. 456, 12–29.

14. Froidurot, S.; Zin, I.; Hingray, B., (2016) Gautheron, A. Sensitivity of precipitation phase over the Swiss Alps to different meteorological variables. J. Hydrometeorol. 2014, 15, 685–696.
15. Liaqat, U.W., Choi, M., Awan, U.K., (2015) Spatio-temporal distribution of actual evapotranspiration in the Indus Basin irrigation system. Hydrol. Process.
16. Jiang, Y., Liu, C., Li, X., Liu, L., Wang, H., 2015. Rainfall-runoff modelling, parameter estimation and sensitivity analysis in a semiarid catchment. Environ. Model. Softw. 67, 72–88.
17. li, A., Amani A., Diedhiou A., and Lebel T., (2005) Rainfall estimation in the Sahel. Part II: Evaluation of rain gauge networks in the CILSS countries and objective intercomparison of rainfall products. J. Appl. Meteor., 44, 1707–1722.
18. Yin, Z.-Y., Zhang, X., Liu, X., Colella, M., Chen, X., (2008) An assessment of the biases of satellite rainfall estimates over the Tibetan plateau and correction methods based on topographic analysis. J. Hydrometeorol. 9, 301–32.
19. Huffman, G. J., Adler R. F., Bolvin D. T., and Nelkin E., (2010) The TRMM Multi-Satellite Precipitation Analysis. Satellite Rainfall Applications for Surface Hydrology, F. Hossian and M. Gebremichael, Eds., Springer, 3–22.
20. Yatagai, A., Kawamoto, H., (2008) Quantitative Estimation of Orographic Precipitation Over the Himalayas by Using TRMM/PR and a Dense Network of Rain Gauges. Vertical Heating Profiles from Satellites and Models. In: Krishnamurti, T.N., Goswami, B.N., Yatagai, A. (Eds.), SPIE Proceedings vol. 7148, 71480C. International Society for Optical Engineering. http://dx.doi.org/10.1117/12.811943
21. Masson, D., Frei, C., (2014) Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types. Hydrol. Earth Syst. Sci. 18, 4543–4563. www.hydrol-earth-syst-sci.net/18/4543/ 2014/doi:10.5194/Hess-18-4543-2014
22. Lutz, A.F., Immerzeel, W.W., Kraaijenbrink, P.D.A., (2014) Gridded Meteorological Datasets and Hydrological Modelling in the Upper Indus Basin, Final Report, for International Centre for Integrated Mountain Development (ICIMOD), FutureWater, Costerweg 1V, 6702 AA Wageningen, The Netherlands.
23. Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A.P., Black, J., Theriault, J.M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., Gutmann, E., (2012) How well are we measuring snow: the NOAA/FAA/NCAR winter precipitation testbed. Bull. Am. Meteorol. Soc. 93, 811–829. http://dx.doi.org/10.1175/ BAMS-D-11-00052.1
24. Li, J., Heap, A.D., (2011) A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. Ecol. Inform. 6 (3–4), 228–241. http://dx.doi.org/10.1016/j.ecoinf.2010.12.003
25. Dinku, T., Ruiz F., Connor S. J., and Ceccato P., (2010) Validation and intercomparison of satellite rainfall estimates over Colombia. J. Appl. Meteor. Climatol., 49, 1004–1014.
26. Kaab, A., Treichler, D., Nuth, C., Berthier, E., (2015) Brief Communication: contending estimates of 2003–2008 glacier mass balance over the Pamir–

Karakoram–Himalaya. Cryosphere 9 (557–564), 2015. http://dx.doi.org/10.5194/tc-9-557-2015

27. Jobard, I., Chopin F., Berges J. C., and Roca R., (2011) An intercomparison of 10-day satellite precipitation products during West African monsoon. Int. J. Remote Sens., 32, 2353–2376.

28. Bumke, K., (2015), Validation of ERA-Interim precipitation estimates over the Baltic Sea. Geophys. Res. Abstr. 17, EGU2015–EGU3035.

29. Dinku, T., Ceccato P., Cressman K., and Connor S. J., (2010) Evaluating detection skills of satellite rainfall estimates over desert locust recession regions. J. Appl. Meteor. Climatol., 49, 1322–1332.

30. Bumke, K., (2015), Validation of ERA-Interim precipitation estimates over the Baltic Sea. Geophys. Res. Abstr. 17, EGU2015–EGU3035.

31. Goovaerts, P., (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydrol. 228, 113–129.

32. Smedman, A.-S.; Gryning, S.-E.; Bumke, K.; Högström, U.; Rutgersson, A.; Andersson, T.; Batchvarova, E., (2015) Precipitation and evaporation budgets over the Baltic Proper: Observations and modelling. J. Atmos. Ocean Sci. 2015, 10, 163–191.

33. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P, Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al, (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 2011, 137, 553–597

34. Elo-Pereira, M.; Dutra, E.; Viterbo, P., (2011) Evaluation of global precipitation data sets over the Iberian Peninsula. J. Geophys.

35. Adler, R. F., Kidd C., Petty G., Morissey M., and Goodman H. M., (2001) Intercomparison of global precipitation products: The third Precipitation Intercomparison Project (PIP-3). Bull. Amer. Meteor. Soc., 82, 1377–1396.

36. Adler, R. F., and Coauthors., (2003) The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–present). J. Hydrometeor., 4, 1147–1167.

37. ohen Liechti, T., Matos J. P., Boillat J. L., and Schleiss A. J., (2011) Comparison and evaluation of satellite-derived precipitation products for hydrological modelling of the Zambezi River Basin. Hydrol. Earth Syst. Sci. Discuss., 8, 8173–8201.

38. Dinku, T., Ceccato P., Grover-Kopec E., Lemma M., Connor S. J., and Ropelewski C. F., (2007) Validation of satellite rainfall products over East Africa's complex topography. Int. J. Remote Sens., 28, 1503–1526.

39. Kidd, C., Levizzani V., Turk J., and Ferraro R., (2009) Satellite precipitation measurements for water resource monitoring. J. Amer. Water Resour. Assoc., 45, 567–579.y

40. Diro, G. T., Grimes D. I. F., Black E., O'Neill A., and Pardo-Iguzquiza E., (2009) Evaluation of reanalysis rainfall estimates over Ethiopia. Int. J. Climatol., 29, 67–78.

41. Ebert, E. E., Janowiak J. E., and Kidd C., (2007) Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bull. Amer. Meteor. Soc., 88, 47–64.
42. Hughes, D. A., (2006) Comparison of satellite rainfall data with observations from gauging station networks. J. Hydrol., 327, 399–410.
43. Ly, S., Charles C., and Degré A., (2011) Geostatistical interpolation of daily rainfall at catchment scale: The use of several variogram models in the Ourthe and Ambleve catchments, Belgium. Hydrol. Earth Syst. Sci., 15, 2259–2274.
44. Nicholson, S. E., and Coauthors., (2013) Validation of TRMM and other rainfall estimates with a high-density gauge dataset for West Africa. Part I: Validation of GPCC rainfall product and pre-TRMM satellite and blended products. J. Appl. Meteor., 42, 1337–1354.
45. Romilly, T. G., and Gebremichael M., (2011) Evaluation of satellite rainfall estimates over Ethiopian river basins. Hydrol. Earth Syst. Sci., 15, 1505–1514.
46. Archer DR, Fowler HJ., (2015) Characterizing flash flood response to intense rainfall and impacts using historical information and gauged data in Britain. J. Flood Risk Manage 11:5121–5133. https://doi.org/10.1111/jfr3.12187
47. Blenkinsop S, Lewis E, Chan SC, Fowler HJ., (2017) Quality-control of an hourly precipitation dataset and climatology of extremes for the UK. Int J Climate 37:722–740. https://doi.org/10.1002/joc.4735
48. Shankar Sharma, Nitesh Khadka, Kalpana Hamal., (2020), How Accurately Can Satellite Products (TMPA and IMERG) Detect Precipitation Patterns, Extremities, and Drought Across the Nepalese Himalaya? https://doi.org/10.1029/2020EA001315
49. Cai Y, Jin C, Wang A, Guan D, Wu J, Yuan F, Xu L., (2015) Spatio-temporal analysis of the accuracy of tropical multi-satellite precipitation analysis 3B42 precipitation data in mid-high latitudes of China. PLoS One 10: e0120026. https://doi.org/10.1371/journal.pone.0120026
50. ai, P., & Liu, X. M., (2018). Evaluation of five satellite-based precipitation products in two gauge-scarce basins on the Tibetan Plateau. Remote Sensing, 10, 1316.
51. Dembélé, M., & Zwart, S. J., (2016). Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. International Journal of Remote Sensing, 37, 3995– 4014. https://doi.org/10.1080/01431161.2016.1207258
52. ang, J., Yang, W., Luan, Y., Du, J., Lin, A., & Zhao, L., (2019). Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. Atmospheric Research, 223, 24– 38. https://doi.org/10.1016/j.atmosres.2019.03.001