*Desalination and Water Treatment*
www.deswater.com
doi:10.5004/dwt.2017.20859

79 (2017) 73–86
June

# Evaluation of the performance of gridded precipitation products over Balochistan Province, Pakistan

Kamal Ahmed[a,b], Shamsuddin Shahid[a,c,*], Rawshan Othman Ali[d], Sobri Bin Harun[a], Xiao-jun Wang[e,f]

[a]*Faculty of Civil Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Malaysia, Tel. +60 7 5531624, Fax +60 7 5566157, email: sshahid@utm.my*
[b]*Faculty of Water Resources Management, Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Balochistan, Pakistan*
[c]*Centre for Environmental Sustainability and Water Security, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Malaysia*
[d]*Petrolium Department, Koya Technical Institute, Erbil Polytechnic University, Kurdistan, Iraq*
[e]*State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China*
[f]*Research Center for Climate Change, Ministry of Water Resources, Nanjing 210029, China*

Received 2 September 2015; Accepted 15 January 2017

**ABSTRACT**

Gauge-based gridded precipitation estimates are emerged as a supplementary source of precipitation data where in-situ precipitation data are not readily available. In this study, four widely used gauge-based gridded precipitation products, namely Global Precipitation Climatology Centre (GPCC), Climatic Research Unit (CRU), Asian Precipitation Highly Resolved Observational Data Integration towards Evaluation (APHRODITE), and Center for Climatic Research – University of Delaware (UDel) are compared with in-situ precipitation at three stations located in semi-arid, arid, and hyper-arid regions of Balochistan province, Pakistan. The assessment is carried out at monthly scale and at 0.5° resolutions during 1961–2007. The performance of the data products is evaluated using various statistical approaches including root mean square error (RMSE), bias, non-parametric Kendall rank correlation, and the Mann-Kendall's trend test. The results reveal that the performance of different products varies at different stations. However, GPCC is found considerably better than other products showing high agreement with annual and seasonal precipitation. GPCC also showed lower errors and higher correlations than other products. Albeit with the lack of spatially dense precipitation data in the study domain, this study suggests GPCC precipitation estimates as the most suitable product for the climatic and hydrological studies in a predominantly arid region like Balochistan.

*Keywords:* Precipitation; Gauge-based precipitation analysis; Arid region; Statistical assessments; Pakistan

## 1. Introduction

Precipitation data are essential to identify and understand the variations and changes in regional and global climate [1]. The major problem often faced in conducting such studies in most parts of the world is the availability of long-term data. Despite availability, uneven distribution and the quality of available data often make it inadequate for the hydrological applications [2]. To overcome these difficulties, numbers of multi-source climate data products have been developed which allow climatic and hydrological studies, particularly in the area where long-term reliable observation data are not available or gauge data are sparse [3–5]. These climate data products are mainly classified into gauge-based precipitation, satellite-based precipitation, reanalysis precipitation, and mixtures of different products. Gauge-based gridded precipitation analysis data are often

*Corresponding author.

1944-3994 / 1944-3986 

used by scientific community because of their spatial and temporal continuity and availability for longer period [6]. Number of studies revealed that the gauge-based gridded precipitation analysis data can be used for climate studies such as, trend analysis [6–9], climate downscaling [10–12], as well as hydrological studies like drought characterization [13–17], flood studies [18–20], etc.

Balochistan, the arid province of Pakistan is considered as one of the most vulnerable regions of Pakistan to hydrological hazards, particularly droughts. Therefore, hydro-climatic studies, particularly seasonal rainfall variability, droughts, and their trends are very important for the agriculture and water resources management as well as development and planning activities in the region. Availability of long-term quality rainfall data is the major obstacle for hydro-climatic studies in Balochistan. Other major problems with climatic data in the region are the missing data, inhomogeneity, and sparse distribution of rain gauges [14]. It is anticipated that droughts and water scarcity will continuously increase in the region throughout the 21st century due to climate change [21]. Assessment of on-going and future changes in climate and droughts in the region is therefore very urgent for adaptation and mitigation planning. This emphasizes the need for long-term quality rainfall data for the region.

The objective of present study is to assess the suitability of different gauge-based precipitation data for replicating observed monthly rainfall of Balochistan. The performance of four gridded precipitation datasets is assessed in the present study namely, datasets of global precipitation climatology centre (GPCC), version V.6 (GPCC) [22]; datasets of climatic research unit (CRU TS 3.22), University of East Angelia [23]; Asian precipitation – highly-resolved observational data integration towards evaluation (APHRODITE) [24] and; the data produced by the Center for climatic research, University of Delaware (UDel) [25]. The above mentioned datasets were selected because of their availability and popularity in climatic and hydrological applications [1–8]. Various statistical approaches including root mean square error (RMSE), bias, non-parametric Kendall rank correlation, and the trend test were used to evaluate the performance of those datasets. It is expected that identification of reliable gauge-based gridded precipitation data will help to conduct hydrological and climatic studies in Balochistan, which in turn will help in impact assessment and mitigation planning in this vulnerable region of the country.

## 2. Study area and datasets

### 2.1. Study area

Balochistan is a mountainous, desert and an arid province, located between 25° to 32°N and 61° to 70.5°E. The location of Balochistan on the map of Pakistan is shown in Fig. 1. Physiographically, it is plateau of rough terrain divided into basins by ranges of various heights and ruggedness. The topography of the study area extracted from the Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 2 (GDEM V2) data, shown in Fig. 1, reveals a large variation of topography over a short distance. The high variability in topography strongly influences the climate in the region [26].

The climatic map of province for the period 1961–2004 [26] is shown in Fig. 2(a). The map shows that the climate of the province lies in the region of hyper-arid, arid and semi-arid based on UNESCO (1979) aridity model. The rainfall in the area is scanty and unevenly distributed. Spatial distribution of the annual rainfall in the area during 1961–2004 [26] is shown in Fig. 2(b). The figure shows that the rainfall in the area varies from <100 mm in the southwest desert to >300 mm in the northeast.

Monsoon winds and the western depressions are the main sources of rainfall during summer (June to September) and winter (December to March), respectively in the study area. Monsoon is a seasonal prevailing wind in the region of South and Southeast Asia that bring moist air from Bay of Bengal during summer. The winds enter the area from southeast corner of the province and therefore, southeastern part of the province receives more rainfall during summer. As the monsoon progresses through the land, air moisture content reduces, and the amount of monsoon rainfall gradually decreases from the east to the west. On the other side,

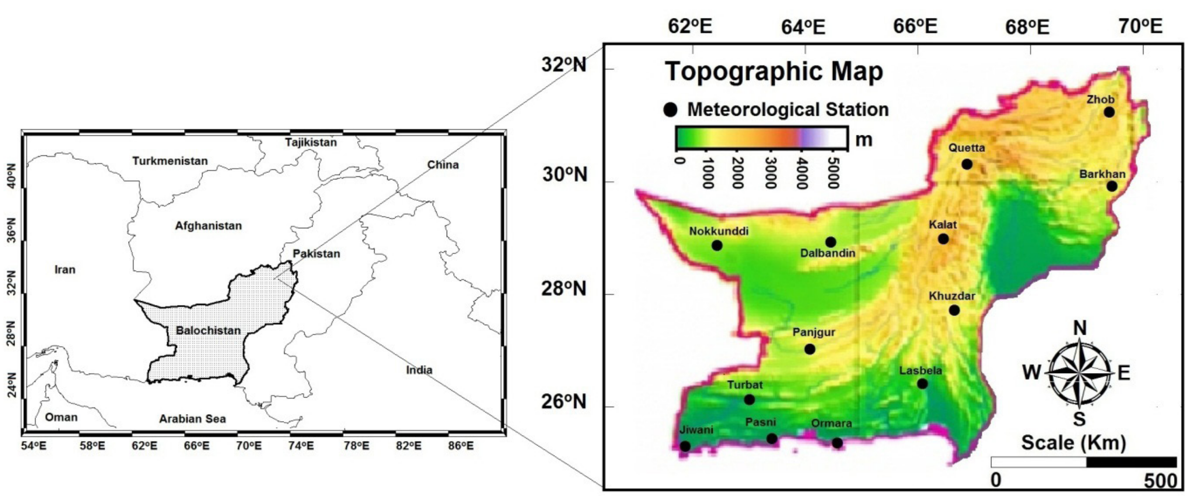


Fig. 1. Geographical location, topography and location of meteorological stations of Balochistan province, Pakistan.

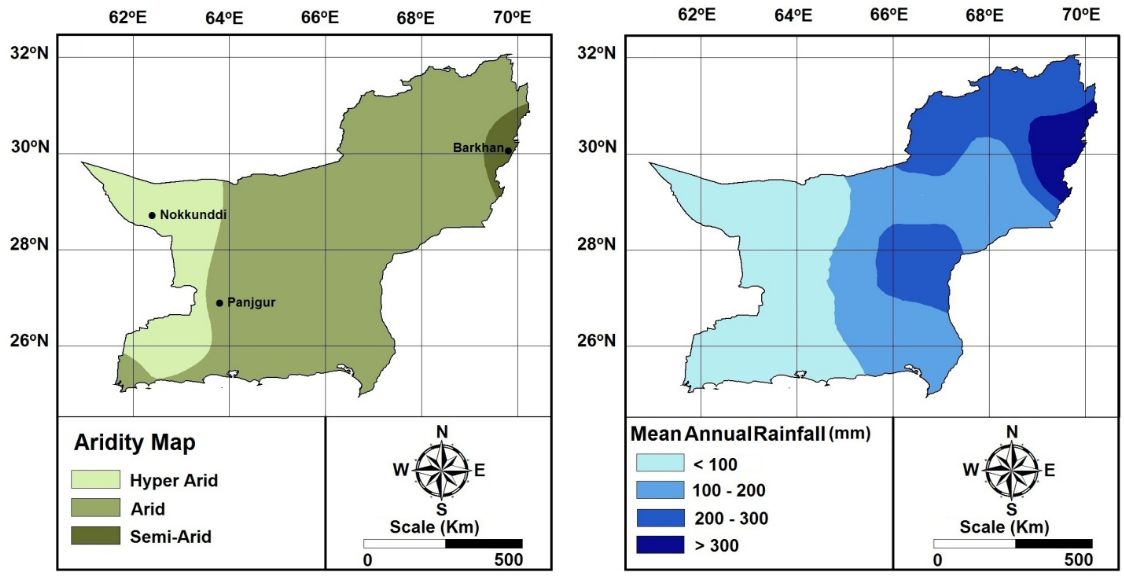


Fig. 2. Map of (a) aridity; and (b) mean annual rainfall of the study area.

western depressions that originate in the Mediterranean Sea travel eastward in higher latitudes (within 30° to 60° N) and the regions of Pakistan lie in 34°N to 36°N receives maximum rainfall due to the topography of Himalayan Mountains. These disturbances induce secondary rainfall in the lower latitudes of 25° to 30°N, where most parts of Balochistan are located [27]. The secondary disturbances extract moisture from the Arabian Sea and bring a small amount of rainfall to most parts of the province. The area receives 58% of annual rainfall during winter, while 31% during monsoon.

Seasonal distribution of rainfall at three stations in the study area is shown in Fig. 3. It can be seen that there are high variations in seasonal rainfall at different stations located in different climatic zones in the study area. The highest rainfall at semi-arid and arid station is obtained in the month of July due monsoon winds while at hyper-arid station highest rainfall is recorded in the month of February due to western depressions.

Majority of the Balochistan's population (85%) live in the rural areas and rely on agriculture as their main source of income [28]. Agriculture and livestock industries employ about 67% of the labor force and account for 50% of the gross domestic product (GDP) of the province [29]. Rangelands provide about 90% of livestock feed requirements [30]. Any variation in climate severely affects the agriculture and people's livelihood in the region.

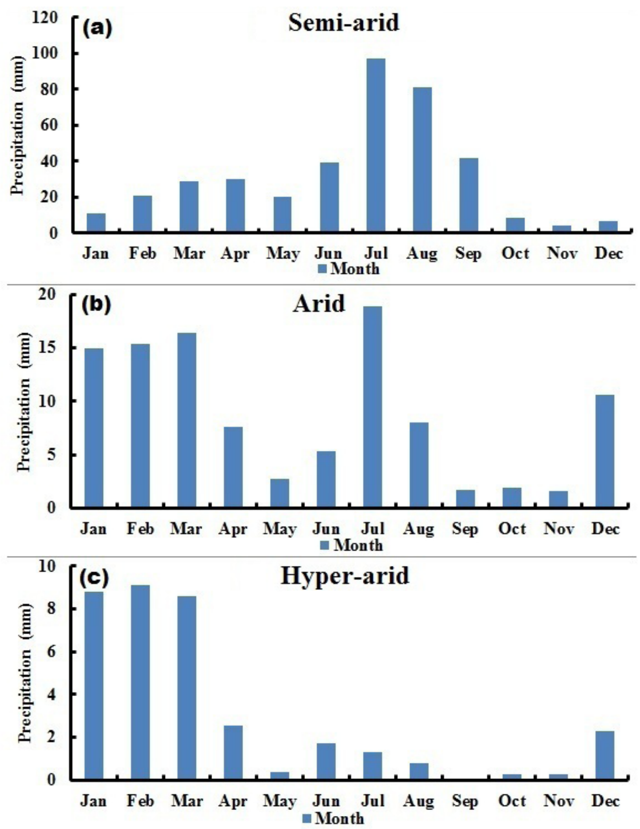


Fig. 3. Seasonal distribution of rainfall in different climatic zones in Balochistan: (a) semi-arid, (b) arid, and (c) hyper-arid.

### 2.2. Datasets

The accuracy of hydro-climatological studies often requires long-term reliable and continuous precipitation data. It is also expected that the data should be homogeneous [31]. Several methods have been developed and applied for the assessment of homogeneity in time series data [32]. These methods are mainly divided into groups namely, relative methods and absolute methods. Relative methods are more reliable and recommended as the test is conducted by correlating the test series with the homogeneous data series of a neighboring station. On the other hand, when homogeneous data series is not available or if the correlation between test series and reference series does not have a good association, absolute methods are preferred [33].

Rainfall in arid regions is very erratic and sparse, and therefore, varies widely over a short distance. Furthermore, rainfall stations are usually sparsely located in most of the arid regions in the world due to less human settlements. Therefore, it is often hard to find homogeneous rainfall time series in the neighboring area. In the present study, absolute methods are used to assess homogeneity of rainfall time series.

There are fifteen meteorological stations in Balochistan operated by Pakistan Meteorological Department (PMD). The homogeneity of rainfall data was assessed using absolute homogeneity tests i.e. Pettitt's test, standard normal homogeneity test (SNHT) and von Neumann ratio test. The tests were selected based on their wide and recent application in rainfall homogeneity assessments [33]. These tests were conducted over rainfall data (1961–2007) of each month at each station to detect the inhomogeneity at a confidence level of 95% with null hypothesis ($H_0$): rainfall data are homogeneous; and alternative hypothesis ($H_a$): rainfall data are not homogeneous. The results revealed that the null hypothesis could not be rejected for rainfall at five stations. Monthly rainfall data recorded at three stations out of those five stations were used in the study. The three meteorological stations were chosen based on their location in semi-arid, arid, and hyper-arid regions. The normality of data was assessed at a confidence level of 95% with null hypothesis ($H_0$), rainfall data follows a normal distribution; and alternative hypothesis ($H_a$), rainfall data does not follow a normal distribution.The results based on Kolmogorov-Smirnov and chi-square goodness-of-fit test showed that null hypothesis is rejected at all stations. The location of the stations is shown in Fig. 2(a). The basic statistics of rainfall at each station are provided in Table 1. Table 1 shows a big difference in precipitation between semi-arid and hyper-arid regions. Semi-arid station receives most rainfall during monsoon while arid and hyper-arid regions receive major part of annual rainfall during winter.

Gauge-based gridded precipitation data with a spatial resolution of 0.5° latitude × 0.5° longitude for the time period 1961–2007 were extracted from the websites of the data products. GPCC data were downloaded from the website: http://gpcc.dwd.de/; UDel from: http://climate.geog.udel.edu/~climate/; APHRODITE from: http://www.chikyu.ac.jp/precip/; and CRU from the website: https://crudata.uea.ac.uk/cru/data/. The time period 1961–2007 was selected for the assessment of the performance of gauge-based gridded data as the observed precipitation data are available only for that time period. A brief description of each precipitation product is given below.

#### *2.2.1.GPCC precipitation*

GPCC global gridded precipitation dataset is developed by combining the precipitation data obtained through telecommunication system (GTS), synoptic weather reports (SYNOP) and monthly climate reports (CLIMAT). Proper quality control and harmonization of the station meta-information are carefully done during merging of precipitation datasets from different sources. GPCC uses smart interpolation technique for gridding, which has the ability to consider the systematic relationship between elevation and station observations and consequently, the ability to enhance the estimation accuracies [34]. The data coverage of GPCC monthly precipitation database varies from ~10,000 stations at the beginning of the 20th century to more than 45,000 stations in 1986–1987. At present, GPCC uses rainfall data from more than 85,000 rain gauges for generating global gridded precipitation, which is the highest among all gridded precipitation databases in the world [35].

#### *2.2.2.CRU precipitation*

CRU of East Anglia (UK) University developed number of 0.5° resolution gridded datasets for monthly mean precipitation. In the present analysis, CRU TS 3.22 data covering the period 1901–2013 is used. The CRU database is developed from rain gauge measurements at about 4,000 weather stations distributed around the world. All the collected data are passed through two-stage extensive manual and semi-automatic quality control measures. In the first stage, the data are checked for the consistency and in the second stage, the stations or months are removed which gives large errors during interpolation process. CRU uses angular distance-weighted (ADW) interpolation method to develop the gridded monthly precipitation anomalies [23].

#### *2.2.3.APHRODITE precipitation*

The APHRODITE project was established in 2006 to develop high resolution daily precipitation datasets for the whole Asia [24]. The APHRODITE database is developed using rainfall data collected from 5000 to 12000 rain gauges distributed over Asia. A brief screening based on geographic location of the station is conducted before inclusion of data in APHRODITE database. APHRODITE has different gridded rainfall products with spatial resolutions ranging from 0.05° to 0.25°. The spatial resolution and available time span of APHRODITE dataset differ in different domains of Asia. In this study, daily gridded precipitation data of APHRODITE project version V1101R2/APHRODITEO_MA/050 deg are used. Daily precipitation was converted to monthly

Table 1
Statistical summary of rainfall data at three stations used in the study

| Station name | Climatic zone | Altitude (m) | Annual rainfall (mm) | Monsoon (mm) | Winter (mm) | Missing data (%) | Homogeneity | Normality |
|---|---|---|---|---|---|---|---|---|
| Barkhan | Semi-arid | 1097 | 397.10 | 259.21 | 67.17 | 1.22 | $H_0$ | $H_a$ |
| Panjgur | Arid | 968 | 104.93 | 33.86 | 57.38 | 0.20 | $H_0$ | $H_a$ |
| Nokkunddi | Hyper-arid | 682 | 36.10 | 3.71 | 29.13 | 0.81 | $H_0$ | $H_a$ |

totals for comparison with other gridded precipitation products used in the study.

#### 2.2.4. UDel precipitation

University of Delaware developed several precipitation datasets with support from NASA's Innovation in Climate Education (NICE) Program. The UDel precipitation database is compiled from a number of updated sources e.g. Global Historical Climatology Network (GHCN2). The UDel uses precipitation data from 4,100 to 22,000 stations. DEM-assisted and climatologically aided interpolation (CAI) [36] is used to generate the gridded monthly precipitation. Precipitation dataset version V 3.02 of UDel is used in the present study.

## 3. Methodology

The reliability of gauge-based gridded precipitation datasets is assessed by comparing them with observed precipitation. In literature, the gridded data are often validated with observed data either by interpolating or correlating with nearby stations. Several studies, such as [5,37,38,39] suggest validation of gridded-data by correlating with nearby station data. In the present study, observed data was correlated with the gridded data at four grid points around the observed station. For example, observed rainfall data at the station located in semi-arid region (29.88°N, 69.72°E) was correlated with the gridded precipitation data at four grid points surrounding the station. The obtained correlation coefficient for GPCC gridded precipitation data at points (29.5°E, 69.5°N), (30°E, 69.5°N), (29.5°E, 70°N), and (30°E, 70°N) were 0.95, 0.96, 0.94, and 0.97, respectively. As the higher correlation was found with the data at grid point (30°E, 70°N), it was selected for validation for GPCC data. Similar correlation coefficients were found for other gridded precipitation data and therefore, those datasets at grid point (30°E and 70°N) were used for validation at semi-arid region. The best grid at different stations was assessed for replicating the threshold rainfall. Since, there is no definition of extreme rainfall in Pakistan; 50-, 90- and 95-percentile rainfalls were considered as moderate, severe and extreme rainfall, respectively. Therefore, the threshold assessment was carried out by comparing the 50-, 90- and 95$^{th}$ percentiles of observed against the percentiles of gauge-based precipitation. The temporal variability in annual and seasonal precipitation was also assessed by comparing observed data with gauge-based data. Further, different statistical methods were used for the validation of gridded precipitation data. Details of the methods used in the present study are discussed below.

### 3.1. Root mean square error (RMSE)

RMSE summaries the mean difference between observed and predicted values to provide overall measures of model performance [40,41]. Therefore, RMSE is used in this study to assess the performance of different gauge-based precipitation datasets. RMSE can be calculated using following equation:

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}\left(x_{gridded,i} - x_{obs,i}\right)^2\right]^{1/2} \quad (1)$$

where $x_{gridded}$ is the precipitation at $i_{th}$ grid, $x_{obs}$ is observed data, and $n$ is the number of the observations. A dataset is considered more accurate if RMSE is closer to zero [42,43].

### 3.2. Bias

Bias provides long-term information of the model and, therefore, recommended to use along with RMSE in order to provide adequate indication of model performance [44]. Bias can be calculated as,

$$Bias = \left[\frac{1}{N}\sum_{i=1}^{N}\left(x_{gridded,i} - x_{obs,i}\right)\right] \quad (2)$$

where $x_{gridded}$ is the precipitation at $i_{th}$ grid, $x_{obs}$ is observed data, and $n$ is the number of observations. A positive bias means overestimation, while a negative bias means underestimation by the model. Smaller bias value indicates a better performance of precipitation product.

### 3.3. Correlation tests

Correlation analysis estimates the degree of association between two variables and, therefore, it can be used as a measure to show how gauge-based precipitation data is allied to observed data. A number of correlation tests are available in literature which can be broadly classified as parametric and non-parametric. Parametric correlation tests are suitable when the distribution of data is normal, while non-parametric are used when the data does not follow normal distribution. The distribution of rainfall data in Table 1 showed that data is non-normal therefore non-parametric Kendall rank correlation was estimated to assess the reliability of gauge-based precipitation data.

Kendall's rank correlation (or τ test) is commonly used to assess the relationships between two variables when their distributions are not known. It is a rank-based procedure and therefore resilient to extreme values. The Kendall rank correlation coefficient (τ) between two variables $X$ and $Y$ can be estimated as,

$$\tau = 1 - \frac{4Q}{n(n-1)} \quad (3)$$

where $Q$ denotes the number of inversions among the values of $Y$ that are required to obtain the same order as the values of $X$. If all the pairs are in increasing order, then:

$$\tau = 1 - \frac{4*0}{n(n-1)} = 0 \quad (4)$$

If all the pairs are in reverse order, then:

$$\tau = 1 - \frac{2n(n-1)}{n(n-1)} = 1 \quad (5)$$

*3.4. Trend Test*

Mann-Kendall test (MK) is a simple linear nonparametric test widely used to assess changes in climatic variables [45]. Mann-Kendall test has been recommended by the World Meteorological Organization (WMO) for analyzing hydrological and meteorological trends [46]. One of the main advantages of this method is its capability of coping with missing values or outliers. Furthermore, the performance of this method does not depend on the distribution of the data [47]. Therefore, the Mann-Kendall test was used to assess the annual trends in observed as well as in gauge-based gridded precipitation dataset in order to reveal the similarity in trends. Similarity in trends was considered in terms of sign i.e. positive or negative. Mann-Kendall test statistics ($S$) for precipitation series $\{x_1 \dots x_n\}$ is calculated as,

$$S = \sum_{i=2}^{n} \sum_{j=1}^{i-1} Sign\left(x_i - x_j\right) \tag{6}$$

where $n$ is the length of data series; $x_i$ and $x_j$ are sequential data in time series; and

$$sign\left(x_i - x_j\right) = \begin{cases} -1 \;\; for \;\left(x_i - x_j\right) < 0 \\ 0 \;\; for \;\left(x_i - x_j\right) = 0 \\ 1 \;\; for \;\left(x_i - x_j\right) > 0 \end{cases} \tag{7}$$

The variance of the statistic is calculated using the following equation

$$Var(S) = \frac{s(n-1)(2n+5) - \sum_{P-1}^{q} tp\left(t_p - 1\right)\left(2t_p + 5\right)}{18} \tag{8}$$

where $t_p$ is the number of ties of the $p_{th}$ value, and $q$ is the number of tied values. Standardized test statistics for Mann-Kendall test, which is approximately normally distributed, can then be calculated as,

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}} \;\; if \; S > 0 \\ 0 \qquad\qquad if \; S = 0 \\ \dfrac{S+1}{\sqrt{Var(S)}} \;\; if \; S < 0 \end{cases} \tag{9}$$

Positive or negative value of $Z$ indicates the direction of the trend in time series.

## 4. Results

*4.1. Precipitation threshold assessment*

The observed and gauge-based precipitation at different thresholds was first compared to assess the performance of different gauge-based precipitation in replicating the threshold rainfall. The obtained results are presented in Fig. 4. It can be seen that observed precipitation thresholds are well captured by GPCC in semi-arid station (Fig. 4a). GPCC slightly over-estimated moderate precipitation and slightly under-predicted severe precipitation, whereas all other products are found to under-estimate all rainfall thresholds. GPCC is also found to replicate all the rainfall thresholds at arid (Fig. 4b) and hyper-arid (Fig. 4c) stations. The CRU and UDel are found to well capture the moderate and severe precipitation thresholds, but under-estimate the extreme rainfall in arid station (Fig. 4b). In hyper-arid station, all datasets are found to over-estimate the moderate precipitation. The differences among products are found minor. However, severe and extreme precipitation at this station is captured well by GPCC compared to other products. The CRU is found to over-estimate all the precipitation thresholds at hyper-arid station (Fig. 4c).

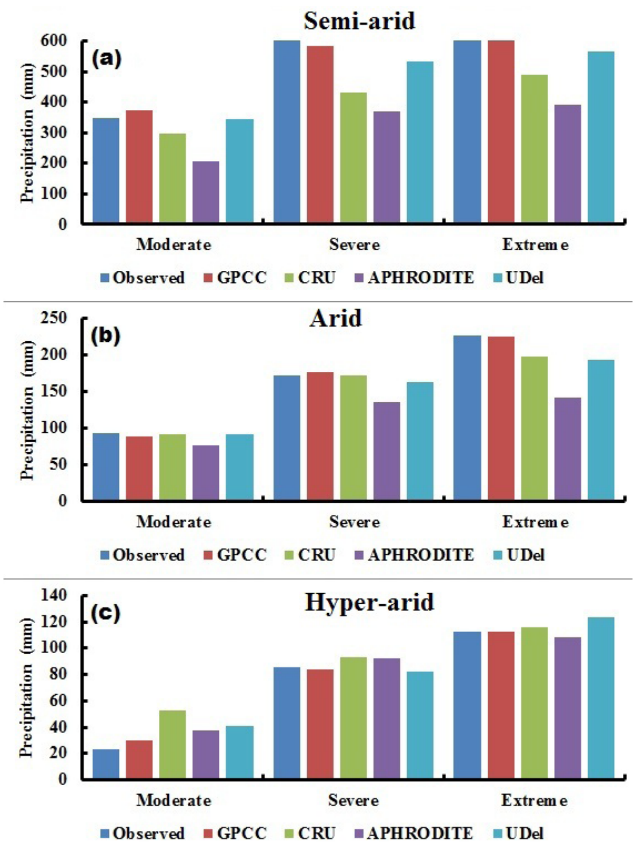


Fig. 4. Comparison of observed and gauge-based precipitation thresholds at (a) semi-arid; (b) arid; and (c) hyper-arid stations for the period 1961–2007.

*4.2. Temporal variability of annual precipitation*

The annual time series of observed and gauge-based precipitation were compared to evaluate the ability of different data products in replicating the observed precipitation. Comparisons were performed individually for each product at each station. The obtained results are shown in Fig. 5. The results at semi-arid station (Figs. 5a–5d) show that GPCC and UDel can replicate the annual precipitation. However, the performance of GPCC (R = 0.83) is found much better compared to UDel (R = 0.46). APHRODITE is found to under-predict the rainfall, but able to replicate the temporal evolution. Similar comparisons at arid station

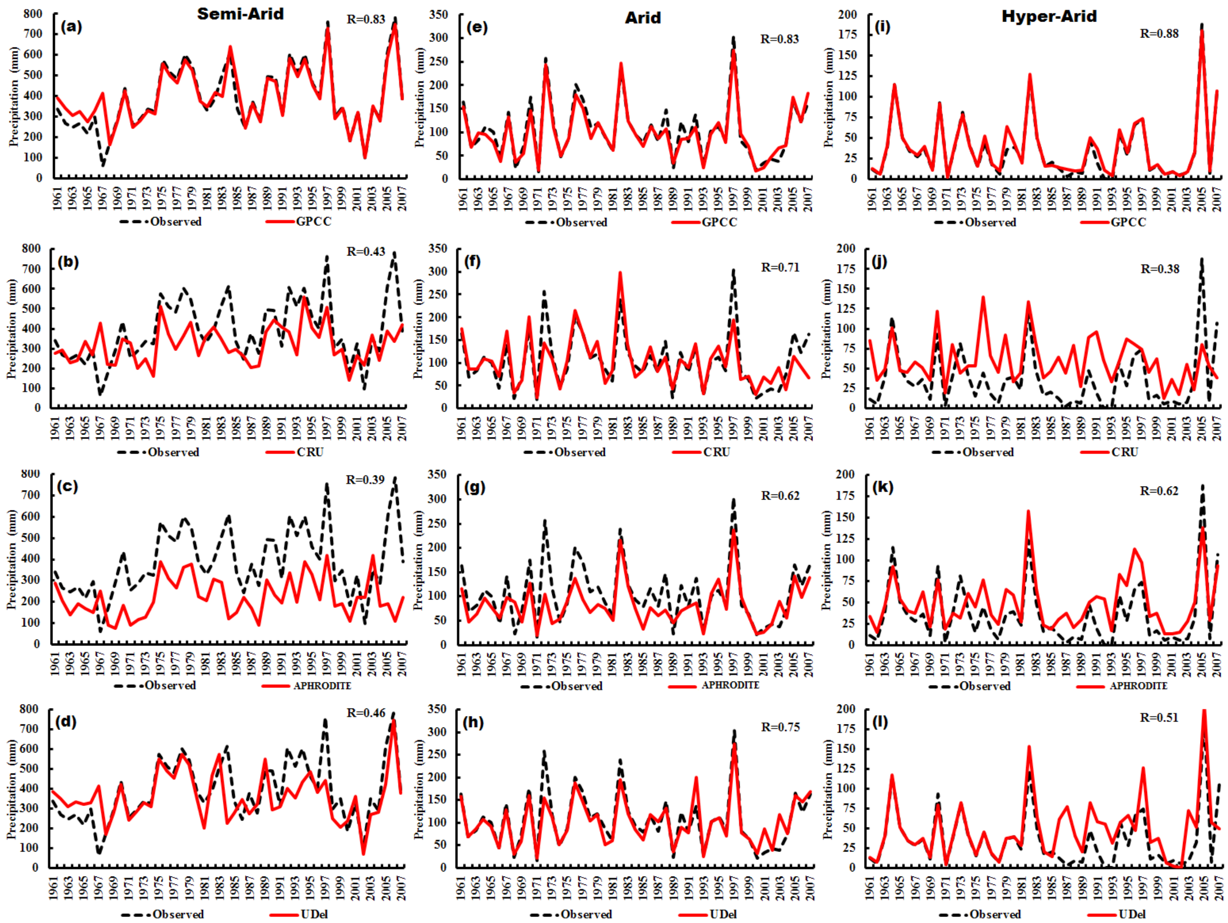


Fig. 5. Comparison of annual observed and gauge-based precipitation time series at semi-arid (a,b,c,d), arid (e,f,g,h), and hyper-arid (i,j,k,l) stations.

(Figs. 5e–h) show that GPCC (R = 0.83), CRU (R = 0.71) and UDel (R = 0.75) can capture the annual precipitation better. At hyper arid station (Figs. 5i–l), GPCC (R = 0.88) and APH (R = 0.62) are found better. It can be noted that UDel (R = 0.51) captured the precipitation successfully till the year 1985, after that it over-estimated rainfall.

### *4.3. Temporal variability of seasonal precipitation*

The rainfall time series of two major precipitation seasons namely, monsoon and winter were also compared to assess the performance of data products. Comparison of monsoon and winter rainfall at semi-arid station is shown in Fig. 6. GPCC (R = 0.86, 0.85) was able to reproduce the rainfall time series in both seasons much better compared to other products. UDel was able to replicate both the time series until 1985. On the other hand, both CRU and APHRODITE were found to under-estimate rainfall in both the seasons.

The comparison of observed and gauge-based precipitation time series during monsoon and winter at arid station is shown in Fig 7. The arid station receives most of its precipitation during winter season. It can be observed that almost all datasets are able to reproduce winter precipitation better than monsoon season. However, the monsoon precipitation is also well captured by GPCC (R = 0.70), CRU (R = 0.69) and UDel (R = 0.69).

The comparison of observed and gauge-based precipitation time series during monsoon and winter at hyper-arid station is shown in Fig 8. The amount of rainfall at this station during monsoon is very low, while it receives substantial amount of rainfall during winter. All data products were able to replicate the winter precipitation at this station. The GPCC (R = 0.82) was found to replicate better as compared with other products. The precipitation at hyper-arid station during monsoon is nearly zero in most of the year; however some extreme rainfall values were recorded in some years. The GPCC (R = 0.52) was found to capture well the extreme rainfalls compared to other data products.

### *4.4. Validation of results*

#### *4.4.1. Root mean square error (RMSE)*

Table 2 shows the RMSE in observed and gauge-based precipitation data. It can be seen that GPCC data gave the lowest error in all months at semi-arid station. On the other hand, different products are found to have the lowest

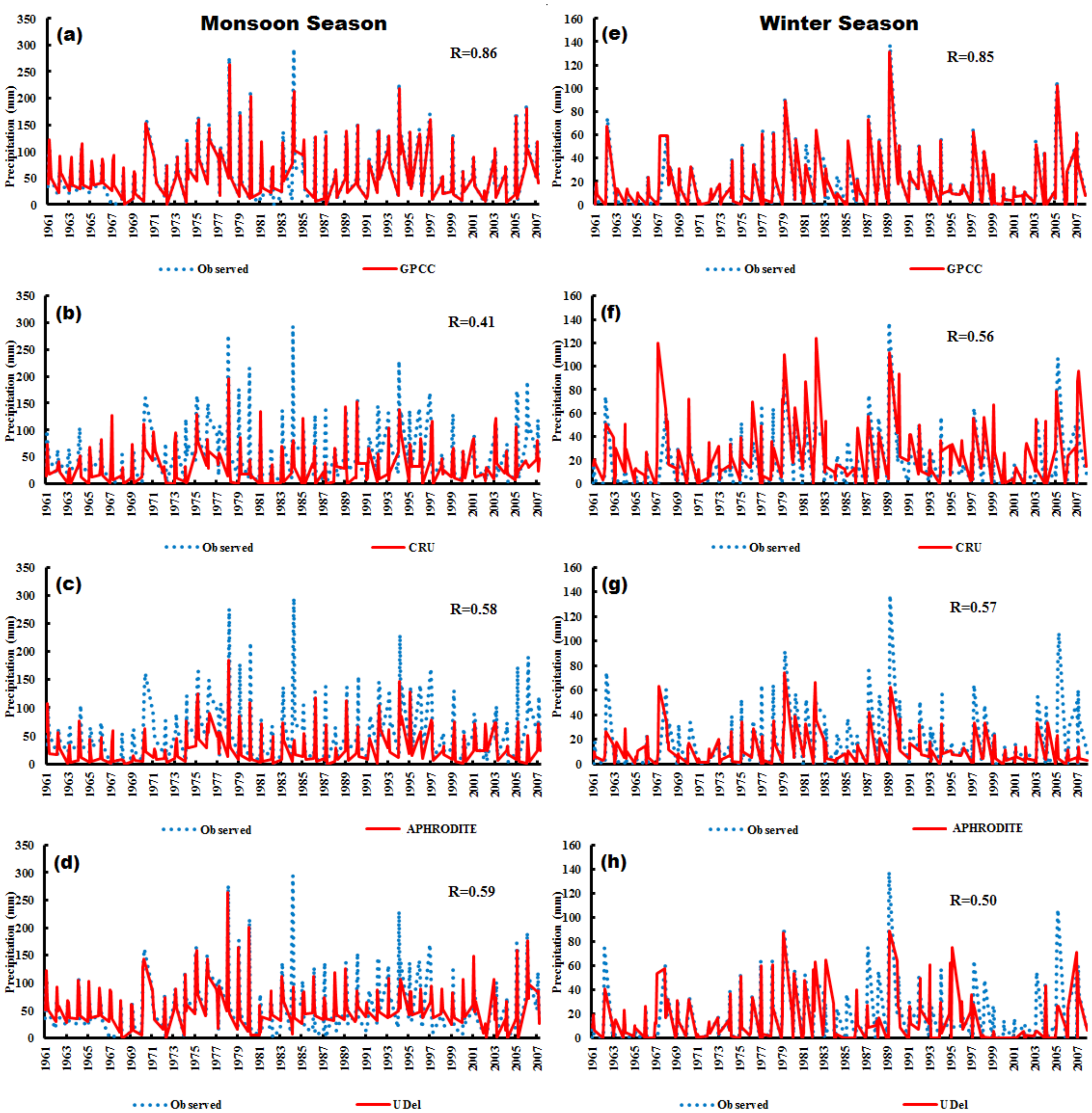


Fig. 6. Comparison of observed and gauge-based precipitation time series during monsoon and winter at semi-arid station.

RMSE in different months at arid station. For example, the GPCC gave the lowest RMSE in the months of February, April, June, September and December, while CRU showed the lowest RMSE in the months of January, March and May, and UDel in the months of July, August, October and November. The RMSE values at hyper-arid station showed that GPCC has lowest errors in all months. Overall, averaged RMSE results shows that GPCC has the lowest errors at all the stations. The CRU showed the highest RMSE in semi-arid and hyper-arid stations, while the APHRODITE showed the highest RMSE in the arid station.

RMSE is highly sensitive to outliers in the rainfall data. Therefore, the obtained RMSE values were compared with normalized room mean square error (NRMSE) to check the consistency. The results were found consistent with other statistics.

*4.4.2. Bias*

Biases in different gauge-based precipitation datasets are presented in Table 3. It can be seen that biases in different gauge-based precipitation vary between different climatic zones. GPCC data showed better performance at semi-arid station, where it showed the lowest biases in all months except May and June. The maximum bias in GPCC is observed in the month of July (–4.43 mm). The biases in other gauge-based precipitation products are relatively high at this station.

At arid station, GPCC showed the lowest bias in the months of January, February, March, June, September, and October; CRU showed the lowest bias in April, and May, while UDel showed the lowest bias in May, July, August and December. The biases in APHRODITE are found very high in all the stations. Similarly, GPCC product showed the lowest

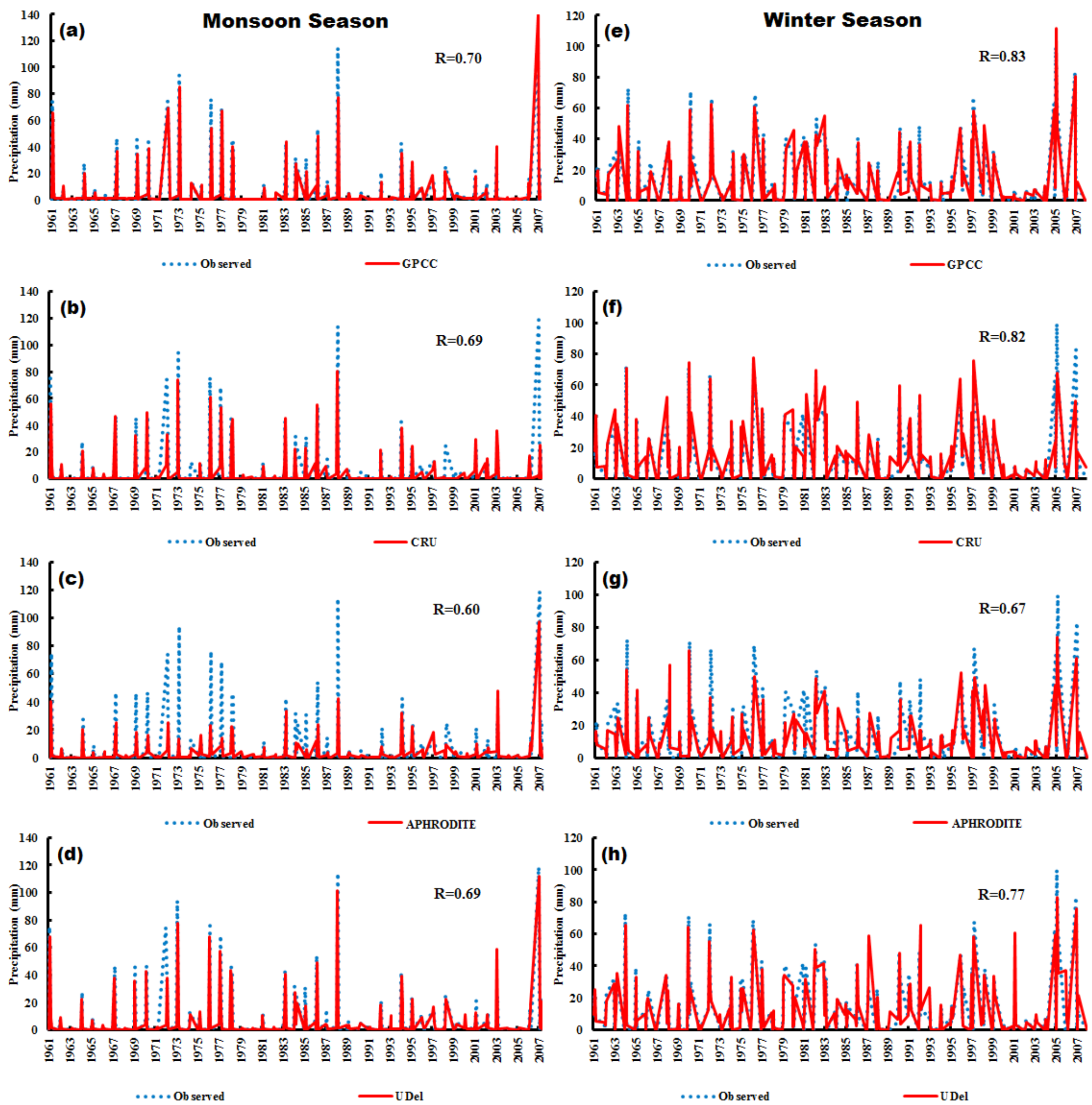


Fig. 7. Comparison of observed and gauge-based precipitation time series during monsoon and winter at arid station.

biases in hyper-arid station. Since the rainfall in the region is very low, the biases are also found very low. GPCC showed the maximum biases of –0.75 mm in the month of February. On the other hand, other gauge-based precipitation products showed higher biases at hyper-arid station in different months.

The sign of biases are found to vary for different gauge-based precipitation data at different months and locations. Overall, the performance of GPCC is found better at all stations in term of biases showing lowest negative bias (–0.62 mm) at semi-arid, (0.38 mm) at arid and (–0.19 mm) at hyper-arid station.

### 4.5. Correlation analysis

The results of non-parametric correlation analysis are given in Table 4. Results shows that highest Kendall's rank correlations is found between observed and GPCC data in almost all months of semi-arid. At arid station, GPCC gave the highest correlation in the months of February, April, and December. On the other hand, CRU showed the highest correlations in the months of March, May, June, July, October, and November, and UDel showed the highest correlation in January, August and September. GPCC was also found to show highest correlations at hyper-arid station in almost all months except August, October and December. Overall, GPCC is found to perform better in semi-arid and hyper-arid stations, while CRU showed better performance in arid station in term of Kendall's correlation.

It can be noted that correlation could not be calculated for CRU dataset for few months at arid and hyper-arid stations due to zero precipitation estimated by CRU in those months. Besides, correlation could also not be calculated at

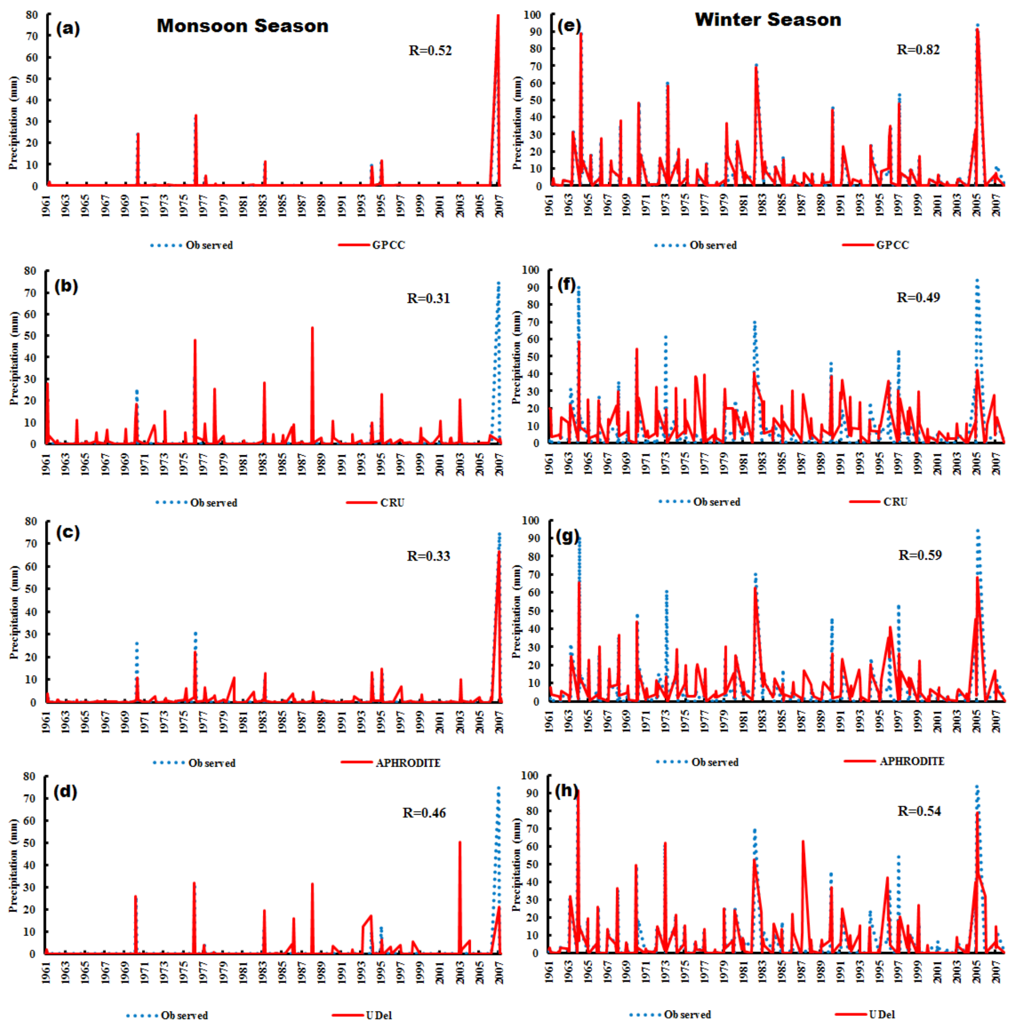


Fig. 8. Comparison of observed and gauge-based precipitation time series during monsoon and winter at hyper-arid station.

Table 2
Root mean square error in gridded precipitation datasets at stations located in semi-arid, arid and hyper-arid regions

| Month | Semi-arid | | | | Arid | | | | Hyper-arid | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GPCC | CRU | APH | UDel | GPCC | CRU | APH | UDel | GPCC | CRU | APH | UDel |
| Jan | **2.24** | 14.97 | 11.67 | 13.14 | 4.41 | **2.73** | 9.06 | 3.83 | **1.54** | 14.47 | 10.59 | 8.02 |
| Feb | **4.23** | 20.70 | 14.74 | 14.12 | **3.73** | 13.01 | 10.63 | 12.31 | **2.39** | 13.73 | 6.47 | 5.91 |
| Mar | **10.72** | 29.51 | 26.94 | 27.91 | 6.21 | **6.03** | 10.20 | 9.30 | **2.91** | 13.71 | 8.62 | 12.68 |
| Apr | **13.76** | 29.03 | 23.61 | 21.08 | **4.25** | 9.47 | 8.83 | 8.79 | **1.77** | 4.19 | 3.61 | 9.65 |
| May | **6.41** | 27.98 | 25.46 | 17.92 | 3.18 | **0.98** | 4.09 | 2.45 | **0.81** | 1.05 | 1.73 | 2.20 |
| Jun | **12.31** | 38.84 | 33.91 | 27.48 | **3.36** | 19.60 | 10.78 | 10.42 | **0.78** | 10.92 | 2.49 | 8.38 |
| Jul | **17.10** | 60.01 | 57.61 | 45.67 | 7.94 | 8.43 | 22.33 | **7.66** | **0.33** | 11.11 | 2.55 | 8.74 |
| Aug | **21.04** | 62.93 | 58.12 | 44.51 | 3.95 | 3.46 | 8.72 | **2.52** | **0.51** | 4.17 | 2.55 | 2.75 |
| Sep | **10.60** | 34.08 | 36.43 | 19.98 | **1.21** | 5.51 | 3.57 | 2.53 | **0.08** | 1.15 | 0.49 | 2.05 |
| Oct | **2.04** | 17.80 | 18.70 | 16.87 | 1.79 | 6.25 | 1.80 | **0.98** | **0.42** | 1.30 | 1.73 | 8.02 |
| Nov | **0.71** | 8.16 | 4.97 | 20.23 | 1.97 | 4.92 | 2.10 | **0.86** | **0.69** | 0.90 | 2.05 | 4.45 |
| Dec | **3.85** | 10.19 | 9.20 | 10.66 | **4.16** | 10.33 | 7.25 | 8.38 | **1.16** | 8.25 | 6.48 | 8.20 |
| Avg | **8.75** | 29.52 | 26.78 | 23.30 | **3.85** | 7.56 | 8.28 | 5.84 | **1.12** | 7.08 | 4.11 | 6.75 |

Bold numbers indicate the lowest error among the data products.

Table 3
Biases in gridded precipitation datasets at stations located in semi-arid, arid and hyper-arid regions

| Month | Semi-arid | | | | Arid | | | | Hyper-arid | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GPCC | CRU | APH | UDel | GPCC | CRU | APH | UDel | GPCC | CRU | APH | UDel |
| Jan | **0.40** | –8.35 | 2.17 | –2.06 | **0.37** | –1.48 | 1.02 | 2.01 | **0.31** | –5.68 | –0.91 | –1.08 |
| Feb | **0.14** | –8.65 | 6.90 | 3.10 | **–0.02** | –1.40 | 1.71 | –2.97 | –0.75 | –4.05 | –1.93 | **–0.41** |
| Mar | **–1.70** | –8.94 | 11.38 | 7.68 | **1.22** | –1.76 | 3.10 | 1.63 | **–0.44** | –3.81 | –1.69 | 0.58 |
| Apr | **1.90** | –2.51 | 10.70 | 9.83 | 1.22 | **0.17** | 1.95 | 0.44 | **–0.15** | –0.59 | –1.49 | –2.77 |
| May | 0.56 | **0.46** | 9.03 | 1.27 | –0.45 | **0.01** | –0.70 | **–0.01** | –0.45 | **0.37** | –0.84 | –0.79 |
| Jun | –2.56 | 18.67 | 22.16 | **1.41** | **–0.83** | 3.13 | 1.02 | 1.71 | **–0.15** | 0.70 | –0.68 | 0.42 |
| Jul | –4.43 | 32.22 | 39.45 | **3.53** | 2.69 | 3.49 | 9.36 | **1.31** | **–0.03** | –5.48 | –0.90 | –1.52 |
| Aug | **–0.18** | 29.20 | 37.77 | 14.92 | 1.43 | –1.16 | 2.70 | **–0.03** | **–0.04** | –1.72 | –0.26 | –0.71 |
| Sep | **–1.96** | 19.53 | 23.26 | –2.95 | **–0.24** | 1.65 | 0.25 | –0.54 | **–0.02** | –0.49 | –0.23 | –0.53 |
| Oct | 0.24 | 3.38 | 2.56 | **0.17** | **0.02** | –0.24 | –0.15 | –0.23 | **–0.19** | 0.31 | –0.54 | –2.15 |
| Nov | **0.04** | –1.57 | 1.23 | –6.01 | –0.45 | 0.54 | –0.09 | –0.26 | **–0.14** | 0.30 | –0.93 | –0.90 |
| Dec | **0.11** | –4.78 | 1.35 | –2.03 | –0.46 | –1.31 | 1.22 | **0.06** | **–0.22** | –3.78 | –3.35 | –2.45 |
| Avg | **–0.62** | 5.72 | 14.00 | 2.41 | 0.38 | **0.14** | 1.78 | 0.26 | **–0.19** | –1.99 | –1.15 | –1.03 |

Bold numbers indicate the lowest error among the data products.

Table 4
Mann-Kendall rank correlation coefficients between gridded precipitation datasets and observed rainfall at stations located in semi-arid, arid and hyper-arid region

| Month | Semi-arid | | | | Arid | | | | Hyper-arid | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | GPCC | CRU | APH | UDel | GPCC | CRU | APH | UDel | GPCC | CRU | APH | UDel |
| Jan | **0.87** | 0.57 | 0.53 | 0.46 | 0.80 | 0.86 | 0.62 | **0.87** | **0.80** | 0.44 | 0.52 | 0.46 |
| Feb | **0.88** | 0.50 | 0.69 | 0.56 | **0.85** | 0.82 | 0.67 | 0.75 | **0.89** | 0.65 | 0.77 | 0.69 |
| Mar | **0.80** | 0.38 | 0.43 | 0.41 | 0.82 | **0.89** | 0.62 | 0.76 | **0.84** | 0.45 | 0.64 | 0.64 |
| Apr | **0.85** | 0.50 | 0.51 | 0.60 | **0.86** | 0.79 | 0.62 | 0.72 | **0.80** | 0.33 | 0.42 | 0.42 |
| May | **0.86** | 0.40 | 0.59 | 0.40 | 0.64 | **0.83** | 0.45 | 0.61 | **0.48** | – | 0.39 | 0.46 |
| Jun | **0.82** | 0.25 | 0.47 | 0.56 | 0.65 | **0.66** | 0.46 | 0.72 | **0.50** | 0.04 | 0.23 | 0.12 |
| Jul | **0.84** | 0.35 | 0.46 | 0.44 | 0.85 | **0.87** | 0.65 | 0.79 | **0.79** | 0.43 | 0.46 | 0.79 |
| Aug | **0.82** | 0.19 | 0.47 | 0.46 | 0.79 | 0.79 | 0.65 | **0.80** | 0.51 | 0.48 | 0.41 | **0.71** |
| Sep | **0.85** | 0.40 | 0.57 | 0.57 | 0.39 | – | 0.42 | **0.60** | – | – | – | – |
| Oct | **0.85** | 0.64 | 0.47 | 0.52 | 0.61 | **0.74** | 0.55 | 0.52 | 0.33 | – | 0.52 | **0.68** |
| Nov | **0.83** | 0.57 | 0.52 | 0.36 | 0.56 | **0.77** | 0.56 | 0.67 | **0.71** | – | 0.51 | 0.61 |
| Dec | **0.81** | 0.64 | 0.52 | 0.53 | **0.85** | 0.81 | 0.73 | 0.79 | **0.78** | 0.42 | 0.49 | 0.39 |
| Avg | **0.84** | 0.45 | 0.52 | 0.49 | 0.72 | **0.80** | 0.58 | 0.72 | **0.68** | 0.41 | 0.49 | 0.54 |

Bold numbers indicate the highest correlation among the data products.

hyper arid region in the month of September due to no precipitation during 1961–2007.

### 4.6. Trend detection results

Precipitation datasets were converted from monthly to annual and seasonal time series for detecting trends in the annual and seasonal rainfall. Mann-Kendall test was applied at each station and season separately. The $Z$ values for annual, monsoon and winter seasons obtained for different datasets at each station are presented in Table 5. It can be noticed that most of the data products have similar $z$ values (same sign + ve or –ve) as observed at all stations. For example, at semi-arid station GPCC, CRU and APH have positive trend in annual and monsoon season while in winter season only GPCC and CRU have positive trend. On the other hand, UDel have negative trend in all seasons that is opposite to observed trends.

The results obtained for arid station showed declining trend for observed rainfall in annual and other seasons. The declining trends were only replicated by GPCC and CRU. APH showed positive trends for annual and monsoon and negative in winter season. Similar to semi-arid station, UDel also showed opposite trend of observed rainfall. At hyper-arid station, decreasing trend was noticed in observed annual and winter rainfall. The same was replicated by all

Table 5
Mann-Kendal $Z$ statistics for annual observed and gridded precipitation datasets at semi-arid, arid and hyper-arid stations

| Station | Dataset | Annual | Monsoon | Winter |
|---|---|---|---|---|
| Semi-arid | Observed | **2.27** | 1.42 | 1.14 |
| | GPCC | 1.32 | 0.53 | 0.64 |
| | CRU | 1.27 | 1.25 | 0.59 |
| | APH | 1.12 | 1.26 | –0.68 |
| | UDel | –0.02 | –0.18 | –0.56 |
| Arid | Observed | –0.57 | –0.55 | –0.75 |
| | GPCC | –0.26 | –0.35 | –0.45 |
| | CRU | –1.50 | –0.72 | –0.97 |
| | APH | 0.28 | 0.33 | –0.68 |
| | UDel | 0.36 | 0.29 | 0.04 |
| Hyper-arid | Observed | –1.45 | 0.14 | –1.37 |
| | GPCC | –1.55 | 0.22 | –1.67 |
| | CRU | –0.78 | –0.26 | –0.96 |
| | APH | –0.34 | 0.75 | –1.04 |
| | UDel | 1.14 | **2.58** | –0.34 |

Bold letter indicates the significance at $p< 0.05$.

gauge-based datasets in winter rainfall while UDel showed contrast to annual observed rainfall trend. The monsoon season at this station showed increasing trend that is replicated by all gauge-based data except CRU where decreasing trend is observed. Overall, the results show that GPCC precipitation $z$ values are much closer to observed $z$ values, while $z$ values of UDel precipitation were found completely different from observed $z$ values. The performance of CRU is also found satisfactory as it shows similar to observed trend.

## 5. Discussion

Several studies has been conducted in past to assess the reliability of various gauge-based datasets around the world [4,6,48–52]. Most the studies showed the superiority of GPCC compared to other gauge-based precipitation data product in replicating observed rainfall. Negrón-Juárez et al. in 2009 evaluated the performance of six gridded precipitation products for tropical South America and concluded that GPCC presented the major agreement over South America [50]. El Kenawy and McCabe in 2015 assessed the performance of rainfall data from six widely used gauge-based products over Saudi Arabia, and suggested GPCC along with CRU, APHRODITE and PRINCETON perform well against most metrics [52]. Prakash et al. in 2015 compared the performance of four gauge-based land-only rainfall products with the Indian meteorological department (IMD) gridded rainfall dataset and reported that APHRODITE and GPCC datasets perform better relative to the others in terms of a variety of skill metrics [4]. Kishore et al. in 2016 investigates the characteristics of Indian rainfall using observations and reanalysis datasets and reported that GPCC data shows similar features as that of IMD with high degree of comparison [6]. Present study also collaborates with the finding of those studies. In the present study, GPCC data was found better compared to other three data products used in terms of all statistical measures used.

Various factors define the performance of gauge-based gridded precipitation data in a particular area which include number of observed stations used, distribution of available meteorological stations, data quality of the observed stations, method of interpolation, and topography of the area [53,54]. One of the major causes of the better accuracy of GPCC data is relatively more observed stations used for constructing the GPCC dataset compared to other data products. Furthermore, the quality of observed time series data in GPCC is controlled through successive automatic and visual checks. Additional statistical and visual evaluations are also performed in GPCC to confirm the reality of any outlier or extreme data [34]. Besides, GPCC uses Smart interpolation, which has the ability to consider the systematic relationship between elevation and station observations and consequently, the ability to enhance estimation accuracies [34]. Schneider et al. in 2014 reported that GPCC products better represent the rainfall amount and pattern over rough topography across the World [35]. The rigorous quality control system, use of more observations, and robust interpolation method probably has made the GPCC more reliable compared to other data products.

## 6. Conclusion

The performance of four gauge-based gridded precipitation analysis data products, namely, GPCC, CRU, APHRODITE, and UDel, are evaluated using various statistical approaches to assess their efficacy in replicating monthly rainfall in Balochistan province of Pakistan. The results revealed clear superiority of GPCC data over other gauge-based precipitation data in the region by capturing different threshold values and following annual and seasonal precipitation. The RMSE and biases values indicate the consistency of GPCC data in giving lowest errors in most of the months of a year at all the stations. Correlation analysis indicates a good association between observed and GPCC precipitation in all the months. In addition to this, GPCC data is found to reveal very similar to observed annual precipitation trends in all the stations. The major factors that contribute to better performance of GPCC data in the study area may be due to (i) the use of comparatively more data points from the region to generate the gridded precipitation; and (ii) the use of smart interpolation approach which has the ability to consider the large topographic variations in interpolating rainfall.

Due to the limitation in availability of observed data, it was not possible to compare the spatial pattern of observed and gridded precipitation in the study area. Gridded precipitation data can be further validated by comparing the spatial structure of precipitation in similar climatic region where more observed data are available. Furthermore, studies can be conducted in future to validate other gauge-based or satellite based gridded precipitation datasets. It is expected that the finding of this study will help to identify the appropriate gauge-based precipitation data product for hydrological or other applications in Balochistan.

**Acknowledgment**

We are grateful to Universiti Teknologi Malaysia (UTM) for providing financial support of this research through research grant (Vote No. 06H36). We are also grateful to the developers of GPCC, CRU, APHRODITE, and UDel for providing gridded precipitation datasets.

**References**

[1] S. Feng, Q. Hu, W. Qian, Quality control of daily meteorological data in China, 1951–2000: a new dataset, Int. J. Climatol., 24 (2004) 853–870.
[2] S. Prakash, R. Gairola, A. Mitra, Comparison of large-scale global land precipitation from multisatellite and reanalysis products with gauge-based GPCC data sets, Theor. Appl. Climatol., (2014) 1–15.
[3] J.-D. Creutin, M. Borga, Radar hydrology modifies the monitoring of flash-flood hazard, Hydrol. Process, 17 (2003) 1453–1456.
[4] S. Prakash, A.K. Mitra, I.M. Momin, E.N. Rajagopal, S. Basu, M. Collins, A.G. Turner, K. Achuta Rao, K. Ashok, Seasonal intercomparison of observational rainfall datasets over India during the southwest monsoon season, Int. J. Climatol., 35 (2015) 2326–2338.
[5] Y. Yang, G. Wang, L. Wang, J. Yu, Z. Xu, Evaluation of gridded precipitation data for driving SWAT model in area upstream of three gorges reservoir, PloS one., 9 (2014) e112725.
[6] P. Kishore, S. Jyothi, G. Basha, S.V.B. Rao, M. Rajeevan, I. Velicogna, T. Sutterley, Precipitation climatology over India: validation with observations and reanalysis datasets and spatial trends. Clim. Dynam., (2015) 1–16.
[7] L. Hermida, L. López, A. Merino, C. Berthet, E. García-Ortega, J.L. Sánchez, J. Dessens, Hailfall in southwest France: Relationship with precipitation, trends and wavelet analysis, Atmos. Res., 156 (2015) 174–188.
[8] S. Beharry, R. Clarke, K. Kurmarsingh, Precipitation trends using in-situ and gridded datasets. Theor. Appl. Climatol., 115(3–4) (2014) 599–607.
[9] M. Souvignet, P. Laux, J. Freer, H. Cloke, D.Q. Thinh, T. Thuc, L. Ribbe, Recent climatic trends and linkages to river discharge in Central Vietnam, Hydrol. Process., 28(4) (2014) 1587–1601.
[10] D. Mullan, J. Chen, X. Zhang, Validation of non-stationary precipitation series for site-specific impact assessment: comparison of two statistical downscaling techniques, Clim. Dynam., (2015) 1–20.
[11] A. Shirvani, W.A. Landman, Seasonal precipitation forecast skill over Iran, Int. J. Climatol., 2015.
[12] J. Chen, X.J. Zhang, F.P. Brissette, Assessing scale effects for statistically downscaling precipitation with GPCC model, Int. J. Climatol., 34(3) (2014) 708–727.
[13] T. Raziei, D. Martins, I. Bordi, J. Santos, M. Portela, L. Pereira, A. Sutera, SPI Modes of drought spatial and temporal variability in Portugal: comparing observations, PT02 and GPCC gridded datasets, Water Resour. Manag., 29(2) (2015) 487–504.
[14] K. Ahmed, S. Shahid, S. Haroon, W. Xiao-jun, Characterization of seasonal droughts in Balochistan Province, Pakistan, StochEnv. Res. Risk A, (2015) 1–16.
[15] T. Raziei, I. Bordi, L. Pereira, An application of GPCC and NCEP/NCAR datasets for drought variability analysis in Iran, Water Resour. Manag., 25(4) (2011) 1075–1086.
[16] E.E. Moreira, SPI drought class prediction using log-linear models applied to wet and dry seasons, Physics and Chemistry of the Earth, Parts A/B/C, 2015.
[17] H. Varikoden, J.V. Revadekar, Y. Choudhary, B. Preethi, Droughts of Indian summer monsoon associated with El Niño and Non-El Niño years, Int. J. Climatol., 35(8) (2015) 1916–1925.
[18] M. Saber, T. Hamaguchi, T. Kojiri, K. Tanaka, T. Sumi, A physically based distributed hydrological model of wadi system to simulate flash floods in arid regions, Arab J. Geosci., 8(1) (2015) 143–160.
[19] N. CandoganYossef, L.P.H. van Beek, J.C.J. Kwadijk, M.F.P. Bierkens, Assessment of the potential forecasting skill of a global hydrological model in reproducing the occurrence of monthly flow extremes, Hydrol. Earth Syst. Sci., 16(11) (2012) 4233–4246.
[20] M.R. Jury, Malawi's Shire river fluctuations and climate, J. Hydrol., 15(5) (2014) 2039–2049.
[21] K. Ahmed, S. Shahid, S. Haroon, W. Xiao-jun, Multilayer perceptron neural network for downscaling rainfall in arid region: A case study of Baluchistan, Pakistan, J. Earth System Sci., 124(6) (2015) 1325–1341.
[22] A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider, M. Ziese, A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth Syst. Sci. Data., 5 (2013) 71–99.
[23] I. Harris, P. Jones, T. Osborn, D. Lister, Updated high-resolution grids of monthly climatic observations–the CRU TS3, 10 Dataset, Int. J. Climatol., 34 (2014) 623–642.
[24] A. Yatagai, O. Arakawa, K. Kamiguchi, H. Kawamoto, M.I. Nodzu, A. Hamada, A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges, Sola., 5 (2009) 137–140.
[25] K. Matsuura, C. Willmott, Terrestrial precipitation: 1900–2010 gridded monthly time series (1900–2010)(v 3.01 added 6/14/12). University of Delaware. 2012.
[26] K. Ahmed, S. Shahid, S. Haroon, Spatial interpolation of climatic variables in a predominantly arid region with complex topography, Environ. Syst. Decisions, 34(4) (2014) 555–563.
[27] M.M. Sheikh, Drought management and prevention in Pakistan. in COMSATS 1st meeting on water resources in the south: present scenario and future prospects, Islamabad. 2001.
[28] M. Ashraf, J.K. Routray, M. Saeed, Determinants of farmers' choice of coping and adaptation measures to the drought hazard in northwest Balochistan, Pakistan, Nat. Hazards, 73(3) (2014) 1451–1473.
[29] M.H. Hussein, Bonded labour in agriculture: a rapid assessment in Sindh and Balochistan, Pakistan. 2004, International Labour Organization.
[30] M. Shafiq, M. Kakar, Effects of drought on livestock sector in Balochistan Province of Pakistan, Int. J. Agri. Biology (Pakistan), 2007.
[31] J. Kajornrit, K.W. Wong, C.C. Fung, Estimation of missing precipitation records using modular artificial neural networks. In International Conference on Neural Information Processing. Springer, pp. 52–59.2.(2012).
[32] T.C. Peterson, D.R. Easterling, T.R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullett, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimrey, J. Salinger, E.J. Forland, I. Hanssen-Bauer, H. Alexandersson, P. Jones, D. Parker, Homogeneity adjustments of in situ atmospheric climate data: a review, Int. J. Climatol., 18(13) (1998) 1493–1517.
[33] G.M. Tsidu, High-resolution monthly rainfall database for ethiopia: homogenization, reconstruction, and gridding. J. Climate, 25(24) (2012) 8422–8443.
[34] C. Funk, G. Husak, J. Michaelsen, T. Love, D. Pedreros, Third generation rainfall climatologies: satellite rainfall and topography provide a basis for smart interpolation. in Proceedings of the JRC—FAO Workshop, Nairobi, Kenya. 2007.
[35] U. Schneider, A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, B. Rudolf, GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, Theor. Appl. Climatol., 115 (2014) 15–40.
[36] C.J. Willmott, S.M. Robeson, Climatologically aided interpolation (CAI) of terrestrial air temperature, Int. J. Climatol., 15(2) (1995) 221–230.
[37] G. Ali, G. Rasul, T. Mahmood, Q. Zaman, S. Cheema, Validation of APHRODITE precipitation data for humid and sub

humid regions of Pakistan, Pakistan J. Meteorology (Pakistan), 2012.
[38] R.S. Teegavarapu, Floods in a Changing Climate: Extreme Precipitation. 2012: Cambridge University Press.
[39] M. Lovino, N.O. García, W. Baethgen, Spatiotemporal analysis of extreme precipitation events in the Northeast region of Argentina (NEA), J. Hydrology: Reg. Studi., 2 (2014) 140–158.
[40] C.J. Willmott, Some comments on the evaluation of model performance, B AMMeteorol Soc., 63(11) (1982) 1309–1313.
[41] Y. Shen, A. Xiong, Y. Wang, P. Xie, Performance of high-resolution satellite precipitation products over China, J. Geophys. Res-Atmos. (1984–2012)., (2010) 115 (D2).
[42] K. Johnston, ArcGIS 9: using ArcGIS geostatistical analyst. 2004: Esri Press.
[43] F.-W. Chen, C.-W. Liu, Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan, Paddy Water Environ., 10(3) (2012) 209–222.
[44] R.J. Stone, Improved statistical procedure for the evaluation of solar radiation estimation models, Solar Energy, 51(4) (1993) 289–291.
[45] H.B. Mann, Nonparametric Tests Against Trend, Econometrica, 13(3) (1945) 245–259.
[46] World Meteorological Organization (WMO). Analyzing long time series of hydrological data with respect to climate variability. WCAP-3, WMO/TD- No: 224. WMO: Geneva; 1–12. (1988).
[47] P. Sonali, D. Nagesh Kumar, Review of trend detection methods and their application to detect temperature changes in India, J. Hydrol., 476 (2013) 212–227.
[48] R. Chokngamwong, L.S. Chiu, Thailand daily rainfall and comparison with TRMM products, J. Hydrol., 9 (2008) 256–66.
[49] T. Dinku, S.J. Connor, P. Ceccato, C.F. Ropelewski, Comparison of global gridded precipitation products over a mountainous region of Africa, Int. J. Climatol., 28 (2008)1627–1638.
[50] R.I. NegrónJuárez, W. Li, R. Fu, K. Fernandes, A. de Oliveira Cardoso, Comparison of precipitation datasets over the tropical South American and African continents, J. Hydrol., 10 (2009) 289–99.
[51] D. Duethmann, J. Zimmer, A. Gafurov, A. Güntner, D. Kriegel, B. Merz, S. Vorogushyn, Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modeling, Hydrol. Earth Syst. Sc., 17 (2013) 2415–2434.
[52] A.M. El Kenawy, M.F. McCabe, A multi-decadal assessment of the performance of gauge-and model-based rainfall products over Saudi Arabia: climatology, anomalies and trends, Int. J. Climatol., (2015).
[53] Y. Fu, J. Xia, W. Yuan, B. Xu, X. Wu, Y. Chen, H. Zhang, Assessment of multiple precipitation products over major river basins of China, Theor. Appl. Climatol., (2014) 1–12.
[54] Q. Sun, C. Miao, Q. Duan, D. Kong, A. Ye, Z. Di, W. Gong, Would the 'real' observed dataset stand up? A critical examination of eight observed gridded climate datasets for China, Environ. Res. Lett., 9 (2014) 15001.