

*Research article*

## **Evaluation of ERA-Interim, MERRA, NCEP-DOE R2 and CFSR Reanalysis precipitation Data using Gauge Observation over Ethiopia for a period of 33 years**

**Tewodros Woldemariam Tesfaye, C.T. Dhanya\* and A.K. Gosain**

Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi-110016, India.

\* **Correspondence:** Email: [dhanya@civil.iitd.ac.in](mailto:dhanya@civil.iitd.ac.in); Fax: +91-11-26581117; Tel: +91-11-2659 7328.

**Abstract:** The vital demand of reliable climatic and hydrologic data of fine spatial and temporal resolution triggered the employment of reanalysis datasets as a surrogate in most of the hydrological modelling exercises. This study examines the performance of four widely used reanalysis datasets: ERA-Interim, NCEP-DOE R2, MERRA and CFSR, in reproducing the spatio-temporal characteristics of observed daily precipitation of different stations spread across Ethiopia, East Africa. The appropriateness of relying on reanalysis datasets for hydrologic modelling, climate change impact assessment and regional modelling studies is assessed using various statistical and non-parametric techniques. ERA-Interim is found to exhibit higher correlation and least root mean square error values with observed daily rainfall, which is followed by CFSR and MERRA in most of the stations. The variability of daily precipitation is better captured by ERA, CFSR and MERRA, while NCEP-DOE R2 overestimated the spread of the precipitation data. While ERA overestimates the probability of moderate rainfall, it is seemingly better in capturing the probability of low rainfall. CFSR captures the overall distribution reasonable well. NCEP-DOE R2 appears to be outperforming others in capturing the probabilities of higher magnitude rainfall. Climatological seasonal cycle and the characteristics of wet and dry spells are compared further, where ERA seemingly replicates the pattern more effectively. However, observed rainfall exhibits higher frequency of short wet spells when compared to that of any reanalysis datasets. MERRA relatively underperforms in simulating the wet spell characteristics of observed daily rainfall. CFSR overestimates the mean wet spell length and mean dry spell length. Spatial trend analysis indicates that the northern and central western Ethiopia show increasing trends, whereas the Central and Eastern Ethiopia as well as the Southern Ethiopia stations show either no trend or decreasing trend. Overall, ERA-Interim and CFSR are better in depicting various characteristics of daily rainfall in Ethiopian region.

---

**Keywords:** precipitation; spatial variability; trends; reanalysis; wet and dry spells; synoptic stations

---

## 1. Introduction

Long term low frequency variations in rainfall intensity and spatial distribution are reported to be adversely affecting many parts of the world in the recent period [1]. Signatures of similar variations are also evident over Ethiopia, a region characterized by substantial climate variability on inter seasonal, inter annual and long term scales. Often, agriculture, water supply and hydropower energy sectors of this country have been adversely affected by the abrupt changes in the frequency and intensity of dry/wet spells and also by unexpected extreme climate events such as droughts and floods. The socio-economic systems exhibit low adaptive capacity and high sensitivity to these climate variabilities, which in turn demands an intensive integrated hydrological and climatological modelling of this region [2].

Simulation and assessment of potential climate change impacts through hydrological models over any region demands reliable hydrological and meteorological variables for current (observed) and future climate conditions [3]. However, water resources management and hydrologic modelling studies are limited in Ethiopia, mainly due to the unavailability of long term observed data of hydrologic variables [4]. Limited availability of data also restricts undertaking of any regional climate modelling exercises and further assessment of impact of climate change on water resources over the region. The absence of sufficient observed dataset for hydrologic modelling and climate change impact studies forced researchers to simulate the systems with proxy datasets (global datasets or reanalysis datasets). General Circulation Models (GCMs) are found to be good in reproducing the seasonal precipitation; but fail to satisfy water balance requirements [5]. Satellite-based rainfall estimates (SRFE) are used also as a replacement of observed datasets for hydrological simulations. The use of Version 7 TRMM Multi-satellite precipitation analysis (TMPA) has improved simulations in mountainous areas [6]. Stisen and Sandholt showed the importance of satellite based rainfall estimates in simulating continuous distributed hydrological models together with rain gauge observations [7-9]. Considerable improvement in large-scale hydrological modelling is also reported using GRACE (Gravity Recovery and Climate experiment) datasets [10].

Most of the related studies have been found to employ the global datasets for precipitation and temperature [11-13], in the absence of observed datasets. Global Weather Experiments [14] aimed at generating global datasets through a comprehensive space based observing system was an attempt to overcome this major obstacle [15]. Though this was a major step, it is noted that the global observational records prior to 1979 have considerable deficiencies. Last decade witnessed major attempts to overcome this obstacle by creating comprehensive climate records for extended past periods, by preferably reconstructing the three-dimensional state of the atmosphere of the past adopting the similar methodology of Numerical Weather Prediction (NWP) by Bengtsson and Shukla [16]. The available observations are re-analysed with advanced data-assimilation techniques to provide a consistent initial state for the next short term forecast, thus generating a continuous stream of three-dimensional fields of meteorological variables [17].

In addition to global datasets, various reanalysis datasets have found wide usage in modelling exercises, due to the range of variables and the time span made available by various reanalysis. Reanalysis datasets are frequently been used to provide initial and lateral boundary conditions for

regional climate model simulations. Reanalyses have been popular for a wide range of studies due to their homogenous nature compared to the raw observations and are an important source in climate studies related issues for addressing long term climatic conditions [18]. Reanalysis data are classified into three major phases corresponding to the advancements in the observing system: the ‘early’ period from 1940 to 1957, when the first upper air observations were established; the ‘modern radiosonde network’ era from 1958 to 1978; and the era of ‘modern satellite’ from 1979 to present [19]. A few frequently used reanalysis datasets are: European Centre for Medium-Range Weather Forecasts (ECMWF’s) ERA-40 and ERA-interim, NASA’s Modern-Era Retrospective Analysis for Research and Applications (MERRA), Japanese Meteorological Agency’s JRA-55, National Center for Environmental Prediction/National Center for Atmospheric Science (NCEP/NCAR) Reanalysis I, NCEP-Department of Energy (DOE) and NCEP Climate Forecast System Reanalysis (CFSR) [20-28].

The reanalysis datasets are continuously under improvement due to the significant changes in the density of observations and modifications in the data assimilation systems [29-32], which in turn compel the centres to improve their data assimilation systems used for numerical weather prediction. As mentioned before, in recent years, different global reanalysis datasets with high spatial and temporal resolution have been used to compensate the lack of direct observations. However, the quality and reliability of these datasets need to be assured, since reanalysis data are bound to have regional biases. This is highly significant for a country like Ethiopia where influence of geography in its climate is of particular interest, with the central part of Ethiopia being dominated by the East Africa highlands which split the country climatically into different regions by creating a spatially and temporally inhomogeneous distribution of rainfall [33].

Accurate representation of the hydrological cycle in reanalysis data poses a special challenge since it involves many parameters that are constrained only indirectly by observations. Estimates of precipitation associated with the reanalysis data are produced by the forecast model, based on temperature and humidity information derived from the assimilated observations. Approximations used in the model’s representation of hydrologic processes strongly affect the quality and consistency of the hydrological cycle. Any imbalances in the analysed fields relative to the model equations can cause large initial changes in the forecast [29]. The strengths and weaknesses of different reanalysis datasets in simulating various components of hydrologic cycle across the world are examined by many studies [34-40]. Studies focussing on different aspects of precipitation such as climatology, inter-annual variation, long term trend has pointed out the superiority of MERRA and ERA-Interim datasets [38,39]. Global comparison of ERA-Interim and GPCC (Global Precipitation Climatology Centre) indicate more rainfall than GPCC in most of the Northern Hemisphere, and in parts of South America and ERA-Interim probably overestimates the decrease in rainfall in the central African region [21]. Dee et al. compared ERA with Global Climatology Precipitation Project (GPCP), a dry bias in GPCP over the northwest and western regions of Ethiopia is highlighted in Tsidu [21,41]. Nijssen et al. and Adler et al. also pointed out the insufficiency of GPCP, especially in mountainous regions by highlighting its underestimation of precipitation in regions with orographic features [42,43]. However, number of such comparison studies of reanalysis datasets over Ethiopia is limited. Only a few studies have attempted comparison of reanalysis datasets over African continent. Diro et al. found that NCEP/NCAR and ERA-40 captured the spatial pattern of the climatology of the main rainy season “Kiremt” (June to September) in Ethiopia [33]. However, it is also shown that NCEP/NCAR and ERA-40, overestimate the mean rainfall in the North West, West and Central regions and underestimate the same in the South and East regions. Overestimation by NCEP/NCAR is prominent in “Belg” (from

February to May) season; whereas ERA-40 overestimates the mean rainfall in Kiremt. ERA-40 is further shown to have captured the annual cycle over the country much better than NCEP/NCAR. ERA-Interim reanalysis data is also found to be satisfactorily capturing the climatic trends and variability over the entire African continent [44]. In addition, ERA-Interim's regional heterogeneity is also showed by highlighting its inability to reproduce the inter-annual variability over the West Sahel and Central South Africa, and on the other hand its ability in effectively capturing the variability over South Africa and East Sahel. Wu et al. compared MERRA, ERA-40, NCEP-DOE R2 and JRA-25 with precipitation datasets from the Global Precipitation Climatology Project (GPCP) to represent the circulation over the African Monsoon Region, and showed that MERRA produces the best estimate of seasonal precipitation over that region [45]. Diaconescu et al. modelled Western Sahel region using regional climate models (RCMs) from the AFRICA-CORDEX experiment, driven with gridded observed datasets and reanalysis products, and showed that ERA-Interim and MERRA datasets overestimate the number of wet days over Sahel region [46]. A shift in the frequency distribution toward lower daily precipitation magnitude is also highlighted. CORDEX (COordinated Regional Climate Downscaling Experiment) is contributed by the World Climate Research Program (WCRP) and it is a climate model data not a reanalysis data. The RCM hindcast covers 1989–2008 with boundary conditions from the ERA-Interim reanalysis data [47]. The ability of the CORDEX RCMs in simulating large-scale global climate forcing signals was assessed in East Africa by compositing the El Nino-Southern Oscillation (ENSO) and Indian Ocean dipole (IOD) events. It is found that most RCMs reasonably simulate the main features of the rainfall climatology over the three sub regions in East Africa and also reproduce the majority of the documented regional responses to ENSO and IOD forcing's. Reda et al. used ERA-Interim to drive seven regional climate models and showed that the simulations reproduce the temporal variability and the geographical distribution of precipitation reasonably well [48]. Dee et al. compared ERA with GPCP, a dry bias in GPCP over the northwest and western regions of Ethiopia is highlighted in Tsidu [21,41]. Nijssen et al. and Adler et al. also pointed out the insufficiency of GPCP, especially in mountainous regions by highlighting its underestimation of precipitation in regions with orographic features [42,43]. Cheung et al. showed no significant changes in annual rainfall at the national or watershed level in Ethiopia [49]. On the other hand, Wagesho et al. highlighted increasing trend in limited regions in the eastern Ethiopia, decreasing trend in most parts of eastern and northeastern peripheries and negligible trend in most of the regions [50].

Performance evaluation of reanalysis data in simulating the climatic variability is of particular interest especially for a region like Ethiopia, which is geographically heterogeneous with the East Africa highlands in the central part dividing the country climatically and resulting in in-homogeneous distribution of rainfall, with different seasonal cycles. Hence the present study is aimed at evaluating the comparative performance of different reanalysis datasets in simulating the observed daily precipitation of different stations spread across the country, using various statistical and non-parametric techniques and assess the spatio-temporal variability, so as to recommend the most appropriate reanalysis dataset to employ in future hydrologic modelling exercises, climate change impact assessments and regional modelling studies over the region.

## 2. Study region and Synoptic and Reanalysis Data

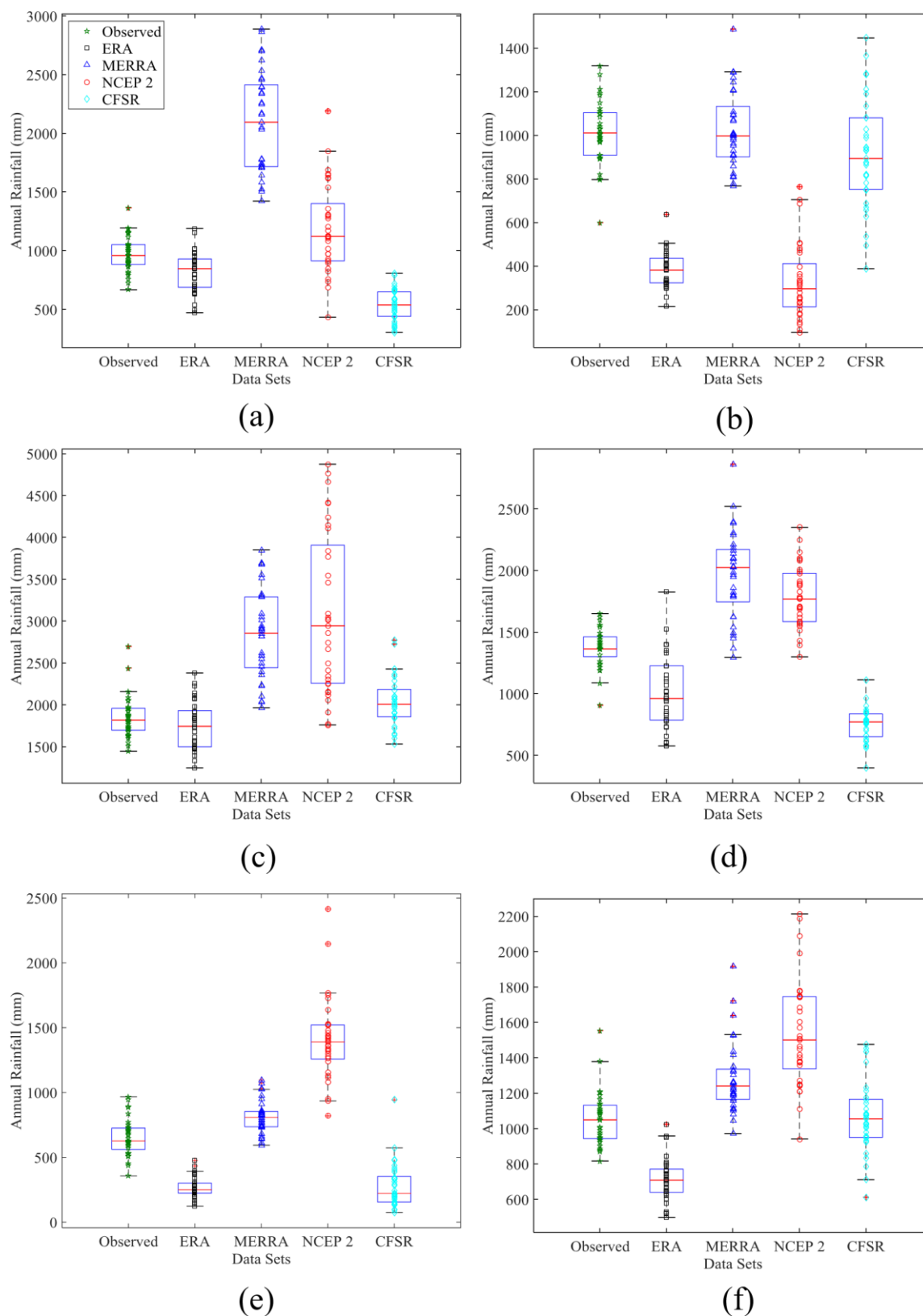
### 2.1. Study Region

Ethiopia is situated in the horn of the East Africa between latitudes 2.9 to 15.3 °N and longitudes 32.7 and 48.3 °E. The climate of the country is tropical in the South eastern lowland regions, and much cooler in the large central highland regions. Annual average precipitation is around 848 mm. The topography of Ethiopia is ragged, with elevation ranging from as low as 125 m below sea level at the Denakil depression in the North east to as high as 4620 m at Ras Dashen in the North west within a near vicinity [49] (Figure 1a). Elevation differences significantly influence the local microclimate, thereby causing large variability and trends in the regional rainfall [51]. The spatio-temporal distribution and magnitude of rainfall is highly variable across the country, with regions located at higher elevations receiving more rainfall when compared to the low arid regions [48,52]. While, the macro-scale pressure systems and monsoon flows which are related to the changes in the pressure systems cause the seasonal and annual rainfall variations in the country [53], the changes in the intensity, position, and direction of the weather systems causes the spatial variations of the rainfall [54]. National Meteorological Services Agency identified the common weather systems that cause rain over Ethiopia as Sub Tropical Jet, Inter Tropical Convergence Zone, Red Sea Convergence Zone, Tropical Easterly Jet, Somalia Jet, African Easterly Jet, and state of the El Niño Southern Oscillation (ENSO) [55,56]. The seasons in country are mainly classified into three: (i) Main rainy season (Kiremt/Summer) from June to September; (ii) Dry season (Bega) from November to January and (iii) Small rainy season (Belg) from February/March to May. The northern and central western part of the country has a single rainy season June to September (Kiremt). Central and eastern Ethiopia has two rainy periods March to May (Belg) and June to September (Kiremt). Southern Ethiopia has two rainy seasons, the long rain (from March to May) and the short rain (from September/October to November) [55].

### 2.2. Synoptic and Reanalysis Data

In this study, four reanalyses rainfall datasets are considered: ERA-Interim, NCEP-DOE R2, MERRA and CFSR. These four reanalysis datasets have frequently been used in hydrologic and climate studies, across the world. Reanalysis datasets are compared with the observed rainfall synoptic stations. Based on the availability of observed data, all datasets are considered for the period 1981–2013. Observed station data is obtained from Ethiopian National Metrologic Agency (ENMA). ERA-Interim reanalysis, NCEP-DOE R2 (NCEPR2) reanalysis, MERRA and CFSR reanalysis are obtained from the data servers of ECMWF (<http://www.ecmwf.int/>), NOAA ([www.esri.noaa.gov](http://www.esri.noaa.gov)) NASA (<http://disc.sci.gsfc.nasa.gov>) and NCAR (<http://rda.ucar.edu>) respectively. The precipitation field is given as a precipitation rate ( $\text{kg m}^{-2} \text{s}^{-1}$ ) in NCEP and CFSR, as total surface precipitation flux ( $\text{kg m}^{-2} \text{s}^{-2}$ ) in MERRA and as total precipitation (m) in ERA Interim.

The pre-processing of observed and reanalysis datasets is necessary because of the difference in its spatial resolution and format. The location of the synoptic stations and spatio-temporal details of all datasets are described below.



**Figure 1.** Box Plots of annual observed and reanalysis rainfall (mm) for the period 1981–2013 shows low (25 percentile) and upper quartiles (75 percentiles) for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa.

### 2.2.1. Synoptic Data

Observed precipitation data from 16 stations, Debre Zeit, Metehara, Awassa, Neghele, Mekele, Nekemte, Combolcha, Debre Markos, Gore, Arba Minch, Gondar, Bahir Dar, Gode, Dire Dawa, Jimma, and Addis Ababa stations are used for the analysis. The locations of these 16 stations are shown in Figure S1a. The stations are well-spread across the country and well-capture the spatial heterogeneity of the region. Various quality checks are conducted in the observation datasets before comparing it with reanalysis data. It is found that missing data is as low as 0.39% for Combolcha station and as high as 19.34% for Gode station. The limitations are mainly due to the stations are not accessible, a problem of data collectors and in addition the civil war in the country may also have contributed for the missing data in the north part of the country (personal communication). In addition to the above limitation, the existing meteorological network density of Ethiopia (about 832) doesn't satisfy the World Meteorological Organization (WMO's) recommendations (2399–5428) [57]. The percentage of missing data for various stations is shown in Figure S1b. The choice of interpolation method depends on the type of data used to fill the gap. Correlation of precipitation data of the nearby station can be used but as precipitation can be quite variable on a short distance, the correlation may not be very good. The conventional centroid method and time-dynamic Voronoi tessellation method were used for SWAT simulation. Time-dynamic Voronoi tessellation utilizes available data more efficiently and significantly reduces the amount of missing data in sub basin climate time series [58]. Missing data at a test station are estimated by weighted averages of observations at neighbouring stations. The weights are inversely proportional with some power of the distance between the test station and the neighbour stations. Spatial interpolation of missing data by weighted average of a spatial data such as Inverse Distance Weighting (IDW), Ordinary Kriging (OK), Universal Kriging (UK), Spline (S) and Topo to Raster (TR) are also used by researchers. Since precipitation data from only 16 stations are available, an intensive missing data replacement (e.g., any spatial interpolation technique) is not advisable. Therefore, missing observed data are infilled through climatological daily precipitation wherever needed. The bias that may arise due to this will be minimal, since the data is missing at different time periods and a huge gap period is not present in any precipitation time series.

Outliers separated from extreme events by homogeneity test. Outliers are defined to be the marginal values of a climate time series, which are very distant from the mean value. They can be due to measurement errors or extreme meteorological events. Outliers that are known to be wrong measurements should be excluded from the data set, but for those that may have a physical background, whether correction should be made or not is an important question [59,46]. The extreme values, if they are correct, carry valuable climatological information that should not be dismissed [60]. On the other hand, outliers can affect the estimation of sample statistics during the use of non-resistant techniques. These techniques are known to be sensitive to the presence of outliers [61]. In order to retain the information of extreme events while not influencing non-resistant statistics too much, outliers can be replaced by a threshold value specific for each time series [59]. Outlier detection is performed using boxplots and using a threshold suggested by González-Rouco et al. [60]. Outliers trespassing a maximum threshold should be replaced by  $P_{out} = Q_{0.75} + 3 \text{ IQR}$ , where  $Q_{0.75}$  is the third quantile and IQR is the interquartile range. No significant outliers are detected in the daily precipitation data of 16 stations. Consistency and homogeneity are assured by performing a double mass curve analysis. For this purpose, the commonly adopted three rainfall regimes based on the rainfall homogeneity determined by the topographical variation, seasonal cycles and regional & global weather systems are

considered [33,62]. Stations in western Ethiopia (from southwest through to northwest) such as Mekale, Debre Markose, Bahir Dar, Jimma, Gore, Nekemte, and Gonder with a mono-modal rainfall pattern fall in regime A and those stations in Central and Eastern Ethiopia such as Awassa, Methara, Debre Zeit, Dire Dawa, Combolcha, and Addis Ababa with bi-modal long rainy season fall in regime B. Whereas stations in South and South Eastern Ethiopia such as Arba Minch, Gode and Neghele with two distinct dry and wet seasons lie in regime C. For more details, readers are referred to Diro et al. and Dawit et al. [33,62]. As an illustration, double mass curve of Debre Markose precipitation is plotted with the cumulative average precipitation of six other stations lying on regime A and is shown in Figure S1c. A straight line with no deviations indicates no significant change in the Debre Markose precipitation regime. Similar exercise is performed for all 16 synoptic stations, considering its rainfall regime and no inconsistency is found in the synoptic precipitation data of 16 stations. These checks also prove that 16 synoptic stations used in this study well represent the three rainfall regimes of the country and are adequate for any comparison study.

Frequency of wet and dry days to check whether the reanalysis hit or miss the reference daily observations was calculated by several categorical indices based on the contingency table. Probability of Detection (POD), False Alarm Ratio (FAR) and Critical Success Index (CSI) were used for this performance test [63]. POD, FAR and CSI were used for this performance test [63]. All three indices range from zero to one. While POD and CSI take zero as worst and one as best, FAR takes zero as best and one as worst. A wet (dry) day is defined as a day with precipitation more (less) than 1 mm. Since most regions of Ethiopia receive significant percentage of its total rainfall during the months of May to October much similar to the seasonal cycle over the northern tropical regions lying between equator and 30 °N, hence, the period between May to October (MJJASO; 184 days) is only considered for this analysis [51,62].

### 2.2.2. Reanalyses Datasets

Since, the four reanalyses datasets differ in their spatial resolution, an interpolation technique is employed to obtain the precipitation values corresponding to the 16 station locations. Interpolation is done for reanalysis data and not for synoptic data in order to preserve the variations in observed data, if any. The interpolation is done by regridding through a nearest neighbour method. However, while assessing spatial variations and trends, a spline interpolation technique is adopted using Geographic Information Systems (GIS) platform at several irregularly spaced station locations [65].

## 3. Results and discussion

Spatio-temporal characteristics of four reanalysis datasets are examined for its efficacy in reproducing long and short term characteristics of observed precipitation values across Ethiopia. Skill of different reanalysis datasets in simulating the annual and monthly precipitation characteristics are initially analysed. Further, a detailed investigation is carried out to compare the daily precipitation characteristics.



**Table 1.** Showing the detail description of the reanalysis datasets (as mentioned briefly in [21,25,66,67,27])

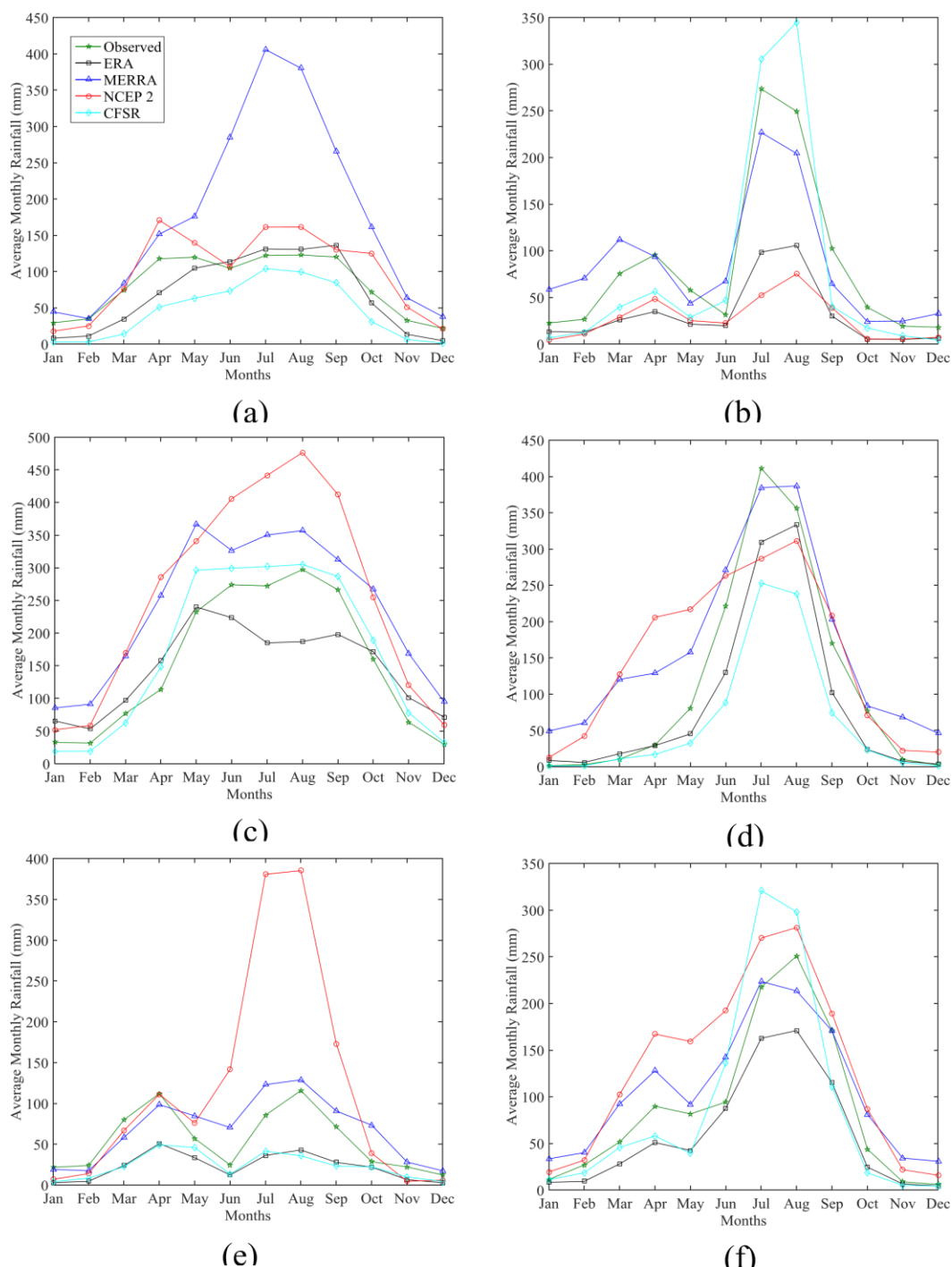
Reanalysis Data	Availability	Spatial Resolution	Data Assimilation Description	Sources
ERA-Interim	1979 to present	$0.25^{\circ} \times 0.25^{\circ}$	4-dimensional variational assimilation system. Horizontal resolution of the dataset is 80 km (T255 spectral) with 60 vertical levels from the surface up to 0.1 hPa	ECMWF
NCEP-DOE R2	1979 to Present	$0.5^{\circ} \times 0.5^{\circ}$	Horizontal resolution of the NCEP-DOE R2 model is T62 (210 km) with 28 vertical sigma levels. The data assimilation system used is 3-dimensional variational analysis technique	NASA
MERRA	1979 to Present	$2/3^{\circ} \times 1/2^{\circ}$ with 72 vertical levels	Uses the Goddard Earth Observing System Data Assimilation (GEOS-5). 3-dimensional variational analysis assimilation	NASA
CFSR	CFSR (1979 to 2010) and CFS version 2, CFSv2 (2011 till present)	$1/2^{\circ} \times 1/2^{\circ}$ global grid with 64 vertical levels	high-resolution coupled atmosphere-ocean-land surface-sea ice system designed to out-perform the older products	NASA
CORDEX AFR	1981 to 2005	$0.44^{\circ} \times 0.44^{\circ}$	lateral boundary forcing obtained from ERA-Interim reanalysis	MPI-M, Germany

### 3.1. Characteristics of Annual Rainfall

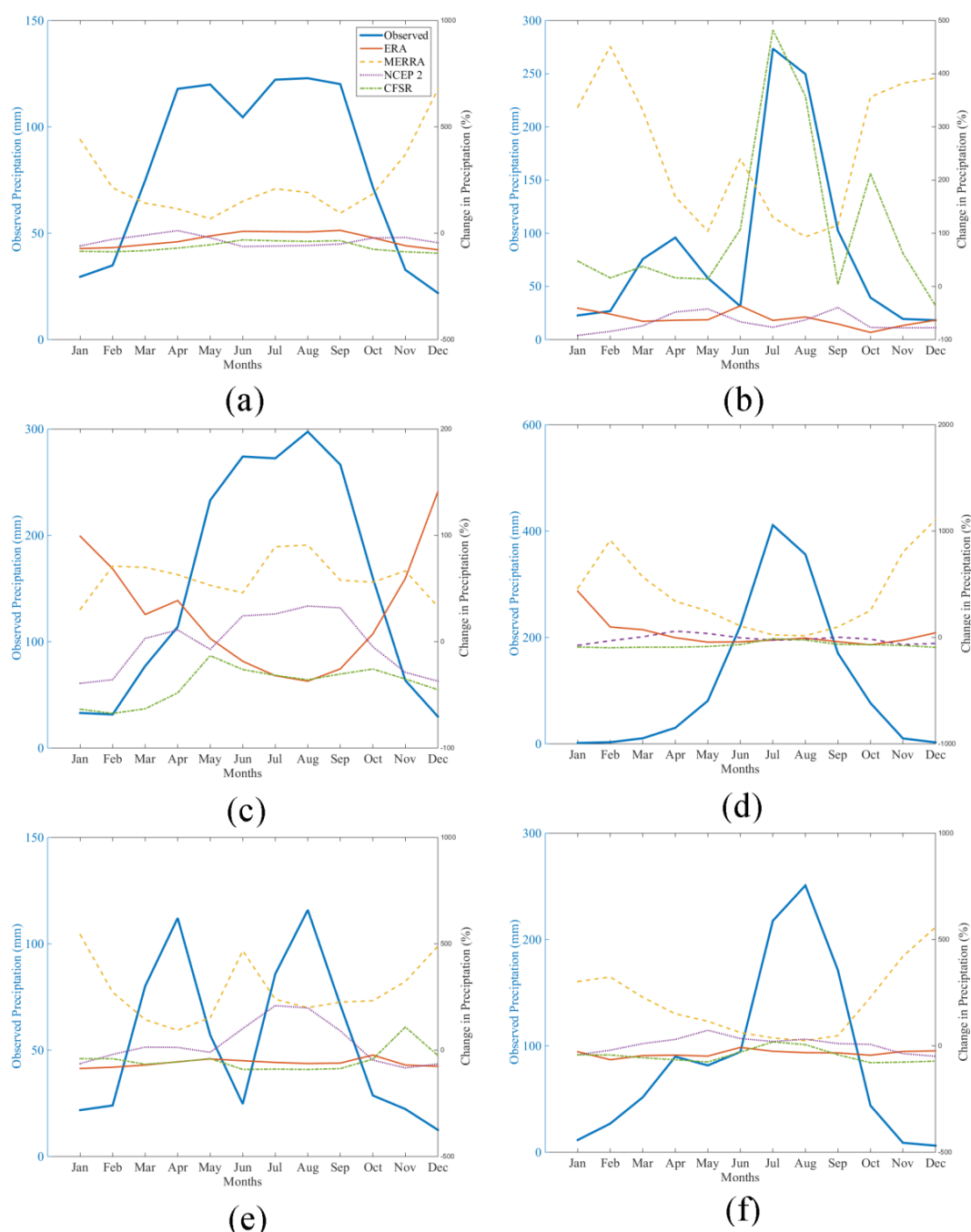
Often, water resources management studies demand annual rainfall and seasonal cycle rather than daily rainfall. Hence the accurate simulation of annual rainfall series need to be checked in reanalysis datasets. The box plots of annual rainfall of observed as well as reanalysis datasets are shown in Figure 1. The behaviour of reanalysis sets are found to be varying for different stations. While NCEP-DOE R2 is significantly different, ERA, MERRA and CFSR tends relatively closer to the spread of the observed annual rainfall, with ERA and CFSR being more close to the observed rainfall in most of the stations. The climatological annual mean is simulated better by ERA and CFSR when compared to NCEP-DOE R2 and MERRA. The inability of NCEP-DOE R2 in simulating the long-term averages is evident from the boxplots.

Further analysis is carried out by computing the climatological annual cycle, to check the capability of reanalysis datasets in simulating the seasonal variation of rainfall over the country. The mean annual cycle computed from observed and reanalysis datasets is shown in Figure 2. While NCEP-DOE R2 is evidently inferior, MERRA tends closer to the observed long-term monthly cycle. However, the seasonal variations in rainfall are simulated well by ERA, MERRA and CFSR, though the magnitudes differ. The seasonality is different for different stations, which further indicates the wide spatial variability of rainfall across the country. The eastern part of the country (Dire Dawa station)

receives rain twice in a year in the months of March to May and in months of July to October. The South-Central part of the country (Awassa station) receives rain in the months of March to June and also in the months of June to September. Regions following a bimodal rainfall pattern are well captured by ERA and MERRA datasets. The month January for ERA shows a fourfold increase of rainfall for Bahir Dar and MERRA a fivefold increase for Dire Dawa (Figure 3).



**Figure 2.** Variation of mean annual cycle of observed and reanalysis rainfall for the period 1981–2013 for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa.



**Figure 3.** Percentage change of the reanalysis relative to the observed Rainfall value for the period 1981–2013 for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa.

- **Spatial Variability and Trends of Annual Rainfall**

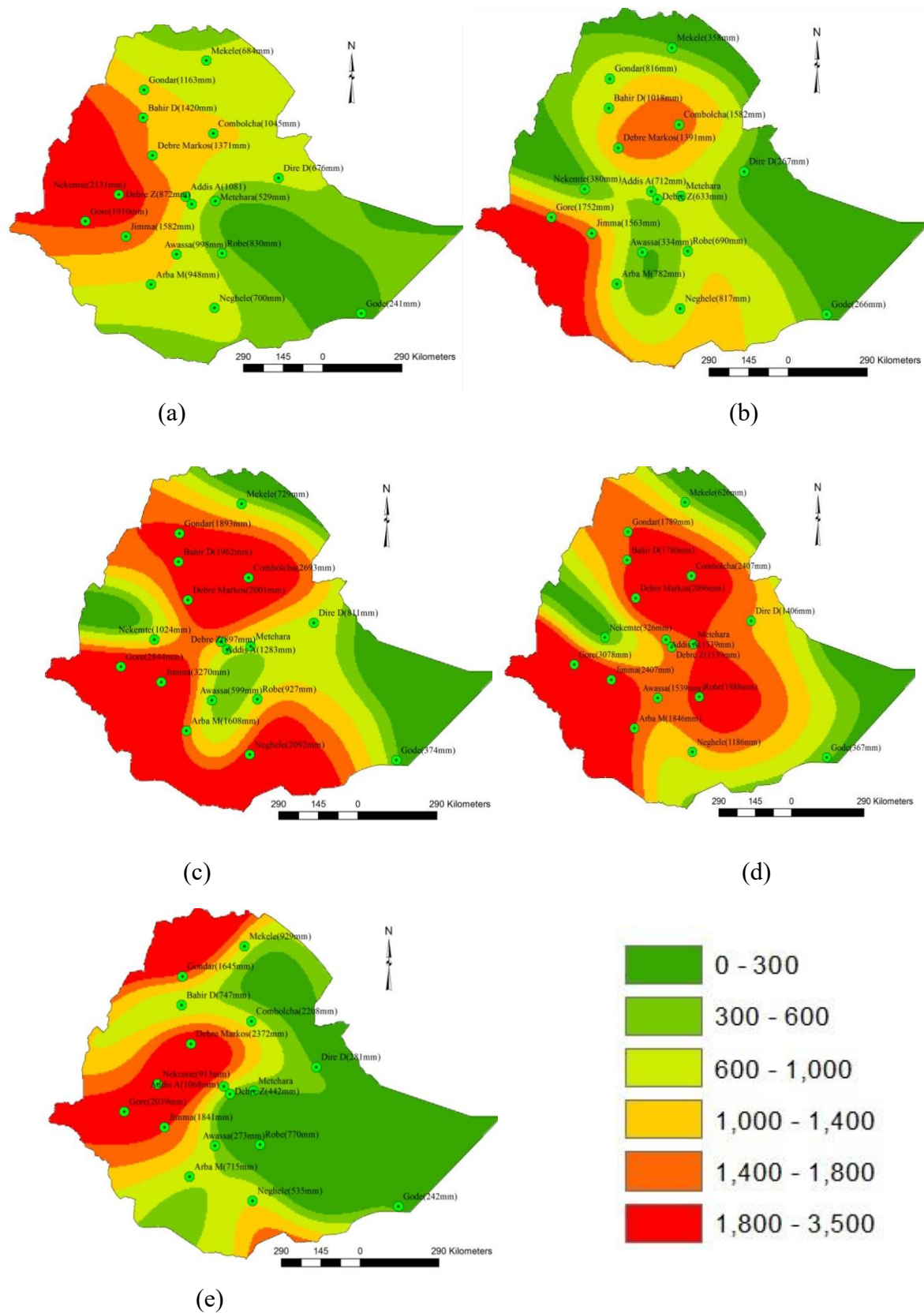
Areal average rainfall is computed from the station data using the spline interpolation in GIS platform, since spline interpolation is recommended for sparse observed data [68]. The spatial variation of long term average annual rainfall of observed and reanalysis rainfall series is shown in Figure 4. Ethiopia receives heavy rainfall in higher elevations which are densely populated and have cultivated

lands. On the contrary, lower elevation areas are more arid and receive less rainfall. The observed annual precipitation is high in the Northern and Central Western Ethiopia. The Central and Western Ethiopia as well as the Southern Ethiopia regions receive ample precipitation (Figure 4). The precipitation pattern is different for four reanalysis rainfall series. While ERA and CFSR tend to under predict the annual precipitation, MERRA and NCEP-DOE R2 tend to over predict the same. MERRA annual precipitation shows high values in the South-Western parts whereas scarce rainfall in the Eastern and upper Northern parts. NCEP-DOE R2 annual precipitation pattern shows low rainfall in the Northern and Southern part of the country. ERA annual precipitation distribution, though almost similar to observed annual rainfall distribution, shows low magnitudes in the South-Eastern part of the country. CFSR annual precipitation shows high values of rainfall in the Central Western part of Ethiopia that shows spatial similarity to the ERA and observed values. The underestimation by ERA appears to disagree with the earlier studies conducted on this region. This might be due to the difference in comparison approaches adopted. Earlier studies had taken a grid-wise comparison while the present study has adopted a station-wise comparison. It is observed that, while for Gore, ERA Interim over estimates the observed rainfall in 8 months (October to March) and underestimates in 4 months (June-September); for Bahirdar, ERA Interim underestimates the observed rainfall in most of the months. The distinct behavior of each station is hence evident from the present analysis, which may not be apparent in grid-wise (spatial averaging) analysis. Hence, a possible overestimation of wet bias in ERA cannot be avoided. The long term mean precipitation demonstrated in these studies may not reveal the station-wise monthly precipitation variations.

A detailed trend analysis is also performed to determine the efficacy of reanalysis series in reproducing the trends in observed data was done by Man-Kendall non-parametric test. Most of the stations located in the northern and central western Ethiopia shows  $Z_c$  value shows a very highly significant (99%) increasing trend. The Central and Eastern Ethiopia as well as the Southern Ethiopia stations show either no trend or a decreasing trend at a 90% significance. The increasing trend is more or less positively correlated with the elevation. While similar results are obtained with previous results, contradictory results are obtained for other regions, which could be due to the difference in time period and the stations considered. The performance of reanalysis datasets is not satisfactory in replicating the observed trends. While NCEP-DOE R2 shows an increasing trend in the North-Western and Eastern parts of the country and a very highly significant decreasing trend in the Western regions, MERRA shows an increasing trend in almost all parts of the country. ERA dataset shows a decreasing trend in the Central part of the country and an increasing trend in the Southern part. CFSR shows an increasing trend in the Northern parts of the country. NCEP-DOE R2 shows the highest decrease of 65 mm of rainfall per year in the Western parts and MERRA shows a decrease of 7 mm in the Northern parts. MERRA shows a maximum increasing trend of 45 mm per year for the South-Western parts of the country. In short, reanalysis datasets fail in reproducing the exact spatial patterns of trends of annual rainfall across the country (See Table S1).

### 3.2. Characteristics of Monthly Rainfall

The four reanalysis datasets are used to derive the rainfall values at 16 locations employing nearest neighbour interpolation technique. The comparison of ERA-Interim, MERRA and NCEP-DOE R2 datasets with the observational data for 16 stations are carried out by estimating various moments and performance measures. The correlation, standard deviation and root mean square (RMSE) of monthly rainfall (mm/day) of observed and reanalyses datasets are compared through Taylor diagrams.



**Figure 4.** Spatial variation of long term average annual precipitation in mm/year for (a) Observed, (b) ERA-Interim, (c) MERRA, (d) NCEP-DOE R2 and (e) CFSR datasets.

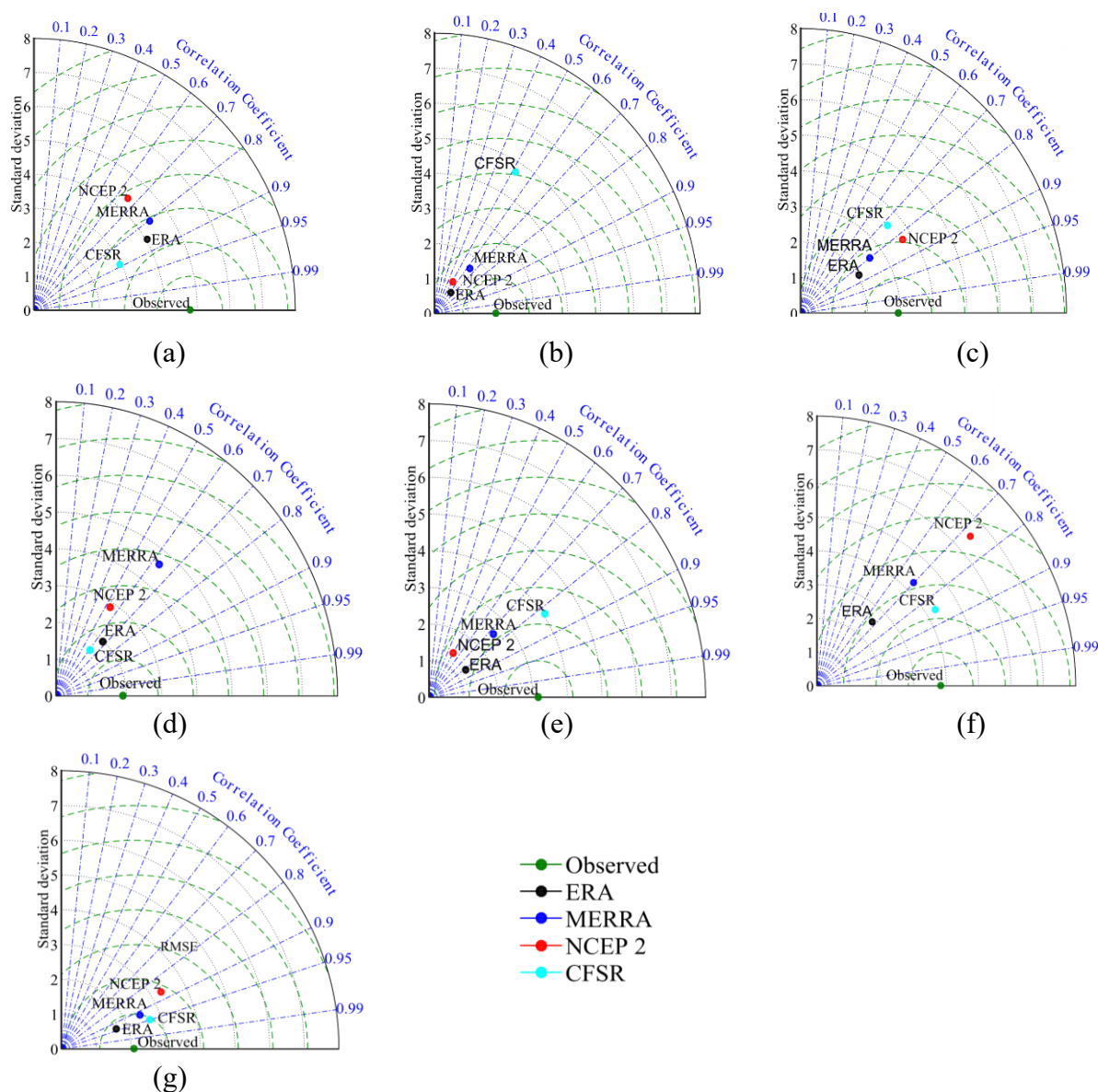
For the purpose of illustration, results of only six stations, (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa (f) Addis Ababa and (g) mean of all 16 stations are presented. These six stations were chosen as they represent each administrative region in Ethiopia, encompassing the cardinal and middle parts of the country so as to well-represent the spatial rainfall pattern of the country. ERA-Interim exhibit higher correlation values with observed daily rainfall, in most of the stations. High correlation coefficients of around 0.8–0.9 are found for ERA, which is followed by CFSR and MERRA in most of the stations. Performance in terms of least root mean square error is also better for ERA as can be seen from Figure 5. RMSE values are comparatively large for NCEP-DOE R2, when compared with that of ERA, CFSR and MERRA datasets. Further analysis is done by estimating various statistical moments, mean, skewness and kurtosis of monthly rainfall (mm/day) for 16 stations, from observed and four reanalysis sets are shown in Figure 5. As shown in Figure 5c, mean daily precipitation values of CFSR are much closer to the observed values for most of the stations. This is followed by ERA and MERRA reanalysis values. It is noted that NCEP-DOE R2 and MERRA are overestimating the mean daily precipitation values, with NCEP-DOE R2 being the worst. The standard deviation of daily precipitation is also better captured by ERA, MERRA and CFSR, while NCEP-DOE R2 overestimates the spread of precipitation data (Figure 6).

Though all datasets exhibit a positive skewness, with more values aggregated below the mean, it is evident from Figure 6b that MERRA underestimates the skewness value for most of the stations. ERA is relatively closer but marginally overestimates the skewness for some of the stations. Kurtosis values are the least for MERRA, and the highest for NCEP-DOE R2 values (Figure 6c). The positive kurtosis exhibited by observed precipitation indicates heavy tailed distribution with sharper peak when compared to the normal distribution. Higher kurtosis values of ERA, CFSR and NCEP-DOE R2 indicate that the variability in data is mainly due to values lying near to the mean than due to occurrence of extreme values. A comparison of the probability distributions and quantile-quantile plots should reveal more details on this aspect.

### 3.3. Probable Occurrence of Daily Rainfall Values

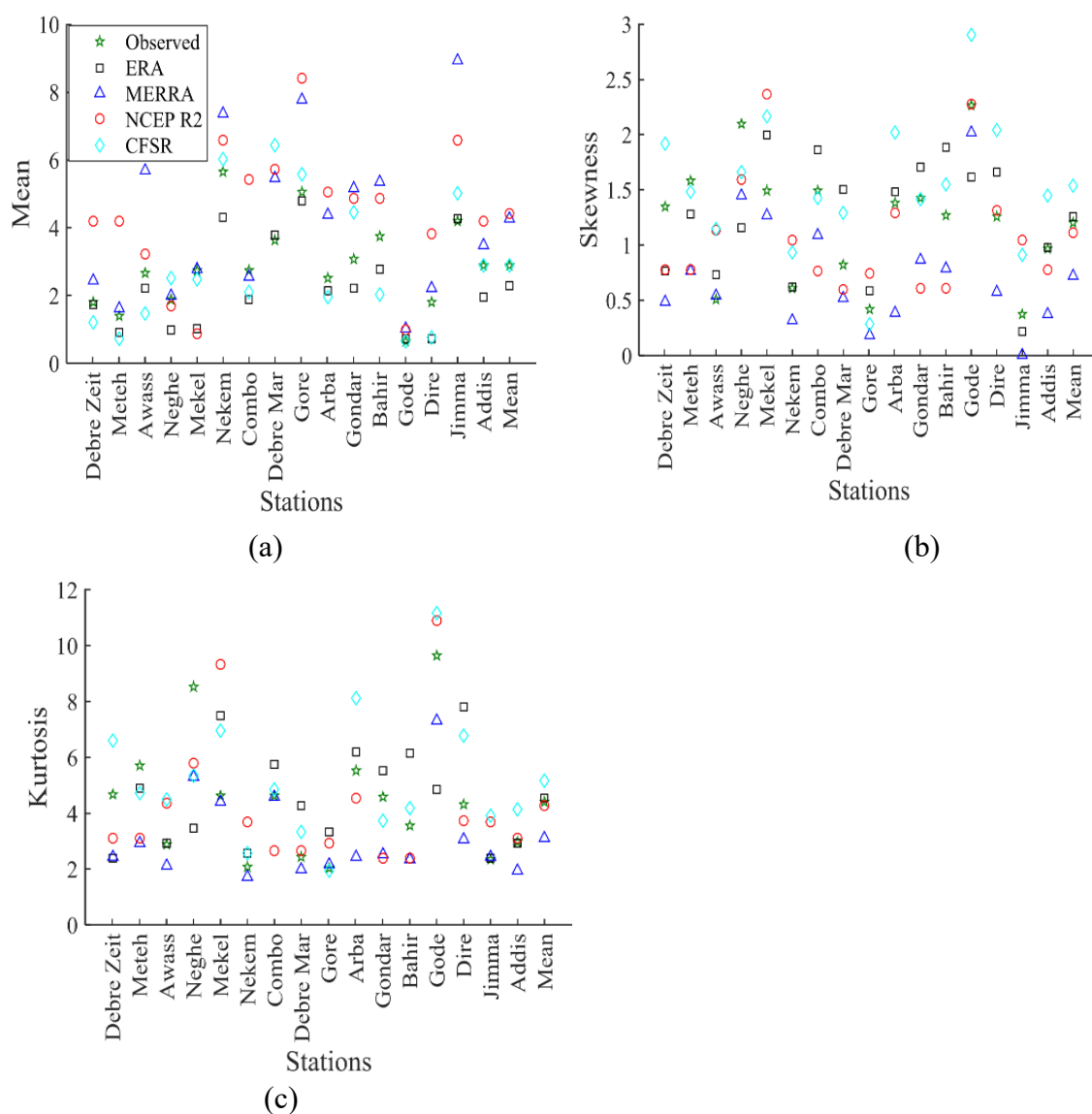
A comparison of cumulative distribution functions (CDFs) of the daily precipitation from observed dataset and reanalysis datasets are shown in Figure S2. For the purpose of illustration, results of only six stations, (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa are presented. While goodness of fit test fails for all the four reanalysis datasets, a visual comparison shows that MERRA precipitation underestimates the probabilities of both zero rainfall or dry days and heavy rainfall. However, it seemingly replicates the probability of moderate rainfall values. ERA appears to be overestimating the probability of moderate rainfall, while capturing the probability of low rainfall fairly well. While, NCEP-DOE R2 is better in capturing the high magnitude rainfall values, CFSR outperforms all three reanalyses in simulating the overall distribution of observed daily rainfall. A percentile-wise performance analysis through quantile-quantile plots is further carried out, for obtaining a clearer picture.

Quantile-quantile (Q-Q) plots of daily rainfall of observed and reanalysis datasets are shown in Figure 7 for the six stations. Although reanalysis datasets are unable to capture the lower percentile values accurately, the superiority of ERA and CFSR is clearly visible from the Q-Q plots. The performance of NCEP-DOE R2 is the worst when compared to other reanalysis datasets.



**Figure 5.** Comparison of monthly precipitation (mm/day) from various stations (observed) and from ERA-Interim, MERRA, NCEP-DOE R2 & CFSR dataset for the period of 1981–2013 using Taylor diagram for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa (f) Addis Ababa and (g) Mean rainfall of all 16 stations.



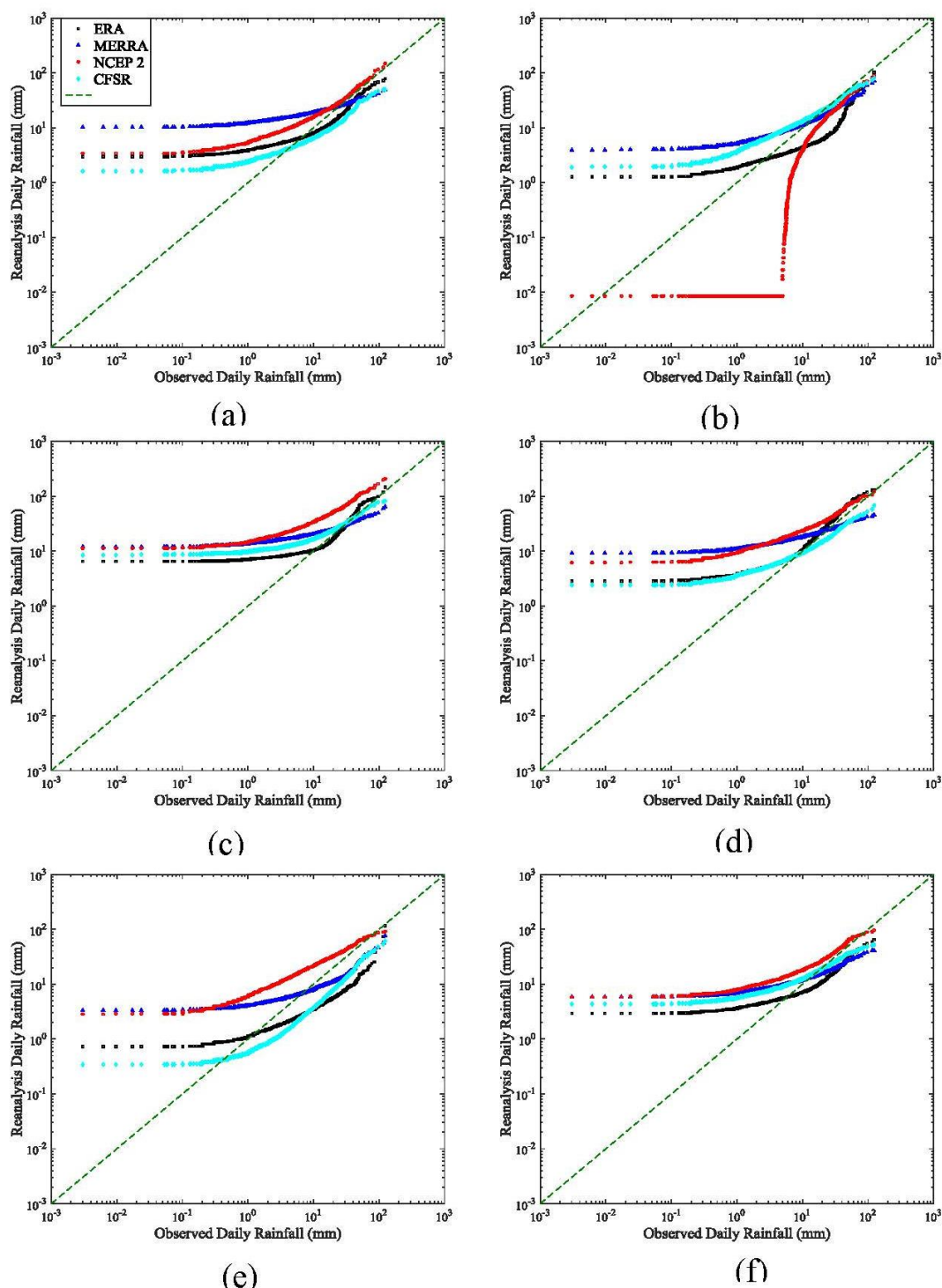


**Figure 6.** Statistics (moments) of observed, ERA-Interim, MERRA, NCEP-DOE R2 and CFSR monthly precipitation (mm/day) for the period of 1981–2013 for 16 stations (a) Mean (b) Skewness and (c) Kurtosis.

### 3.4. Hits and Misses of Daily Rainfall

The dry season and wet season classified as stated in section 2.1.1 and the period between May to October is considered for explaining statistical performance of the frequency of wet and dry days by using the probability of detection (POD), False Alarm Ratio (FAR) and Critical Success Index (CSI). Table S2 gives the summary of these indices for wet and dry days for four reanalysis datasets. Except NCEP-DOE R2, all reanalysis is comparable to yield the hits and misses of both wet and dry days. Overall, NCEP-DOE R2 is inferior in all indices, with worst POD (and CSI) and high FAR. MERRA is seemingly superior in detecting the POD of dry days for most of the stations. Nevertheless, ERA, MERRA and CFSR are in good agreement with each other in simulating different characteristics.





**Figure 7.** Quantile-Quantile plots of daily rainfall (mm) of observed and reanalysis datasets for the period 1981–2013 for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa.

### 3.5. Characteristics of Wet and Dry Spells

The accurate simulation of wet and dry spells is important for the application of any reanalysis dataset as a proxy data for studies related to agricultural and irrigation management such as design of

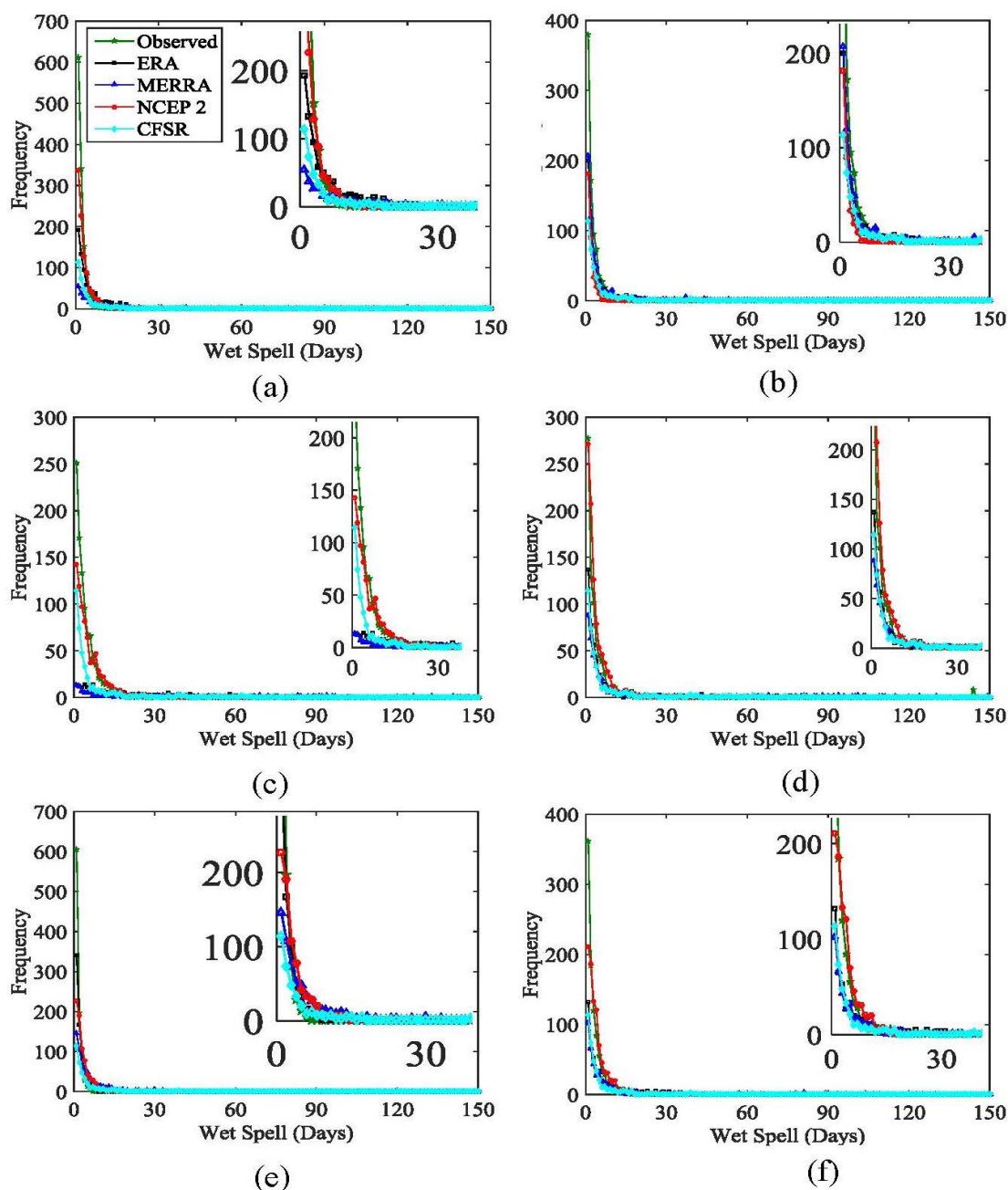
rain water harvesting system, risk assessment of crop destruction through water stress, design of supplementary irrigation system etc. Similar to the previous analysis, period between May to October is only considered for spell analysis. The wet spell length is estimated as the consecutive days (number of days) with precipitation more than 1 mm/day. The variation of wet spell frequency with length is shown in Figure 8. Observed rainfall exhibits higher frequency of short wet spells when compared to that of reanalysis datasets. MERRA is inferior in simulating the wet spell characteristics of observed daily rainfall.

Similarly, dry spell is defined as the consecutive days with precipitation value less than 1 mm/day [69,70]. The variation of dry spell frequency with length is shown in Figure 9. Reanalysis datasets under-estimate the frequency of short dry spell. NCEP-DOE R2 shows higher frequency of short dry spell when compared to that of another reanalysis. As shown in Figure 10, a comparison of mean dry spells indicates that MERRA and NCEP-DOE R2 under-estimate the observed mean dry spell length, whereas ERA and MERRA overestimate the observed mean wet spell length. While, ERA simulates observed mean dry spell length relatively well, NCEP-DOE R2 is effective in capturing the observed mean wet spell length. CFSR is overestimating both the mean wet spell and dry spell lengths. However, considering the ability in replicating diverse characteristics of observed rainfall series, ERA followed by CFSR are evidently more effective when compared with MERRA and NCEP-DOE R2 in simulating the rainfall characteristics.

The results detailed above should be examined with caution since 16 stations may under-represent the precipitation variability of Ethiopia. Though the preciseness of the analysis directly depends on the density of rain gauges, in this study a good spread of stations is ensured across the country to represent the remarkable precipitation zones. While more number of stations will definitely improve the reliability of the results, the under-representation of a few regions like southeast and Western regions may not significantly affect the results obtained since enough number of stations are present for each rainfall regime. Though dense precipitation data from as large as 233 gauge stations over Ethiopia has been used by a few studies, the quality of data was not assured [41]. The quality of data used in the present study over the 16 selected synoptic stations are assured, however.

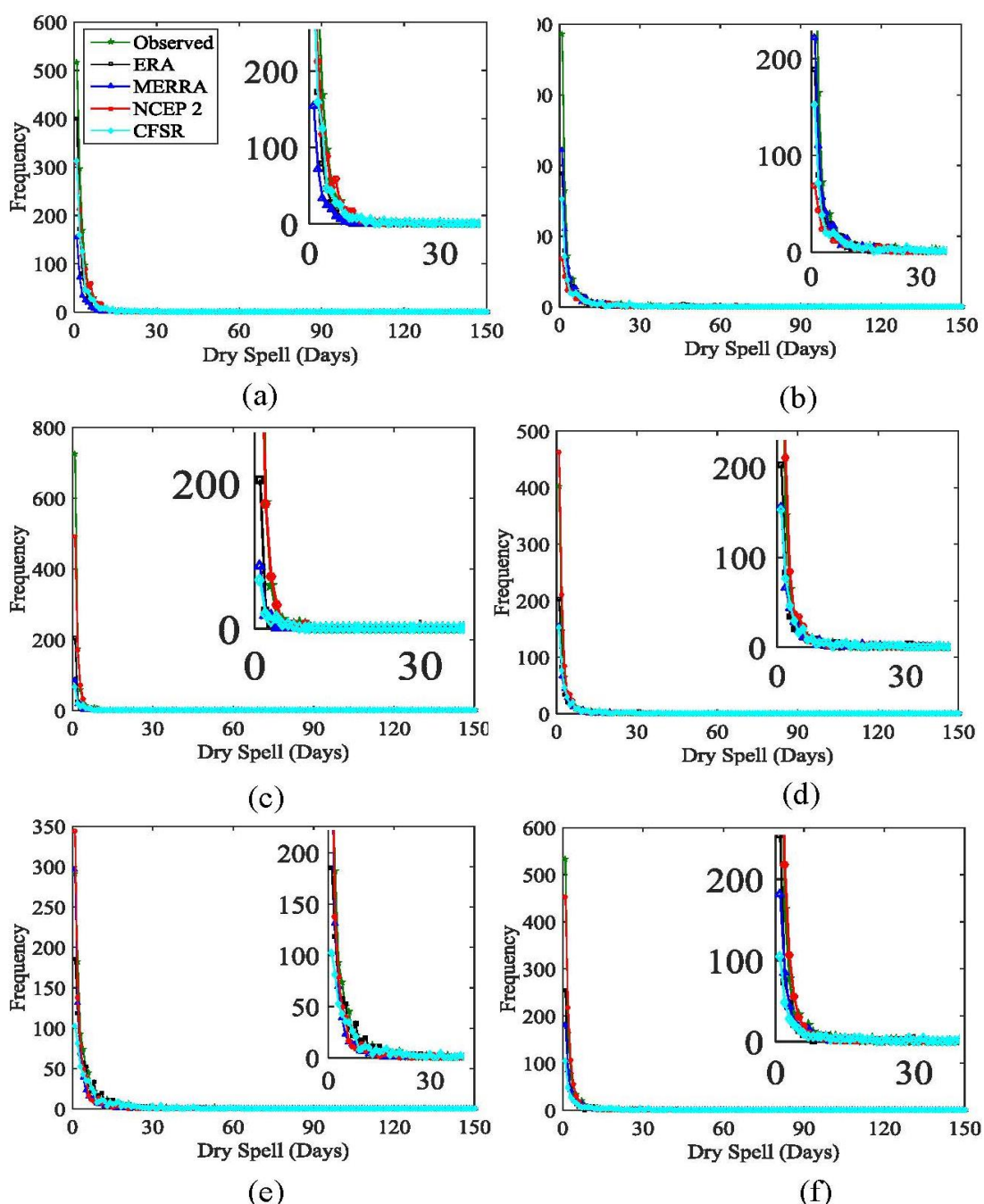
#### 4. Conclusions

Hydrologic modelling and climate change impact studies need reliable and long term information about various climatic variables. The lack of sufficient observed dataset to fulfil this requirement puts demand on alternate proxy datasets. The performance of four reanalysis precipitation data is evaluated and compared with observed station rainfall data. The analysis reveals that mean annual precipitation is well captured spatially over the entire country by all the four reanalysis datasets. The performance of ERA appears to be the best in terms of capturing the diverse characteristics of daily and annual rainfall, followed closely by CFSR. NCEP-DOE R2 is found overestimating the daily, annual and longterm averages. However, reanalysis datasets in general failed to capture the long-term trends in annual rainfall spatially.



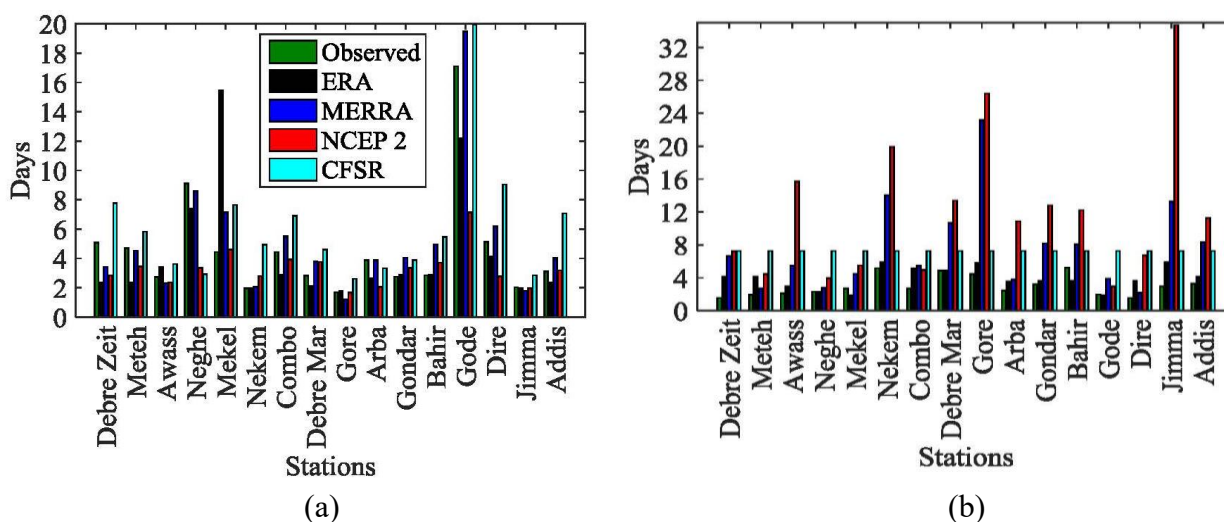
**Figure 8.** Frequency distribution of wet spell lengths of observed and reanalysis rainfall for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa (Zoomed version of moderate wet spell frequency are shown as insets).

However, it is noted here that the presence of gaps in the observed rainfall data (though less when compared to other studies) in a few stations may limit the performance analysis. The lack of observed data affects the accuracy of reanalysis datasets, since reanalysis datasets are derived from the combination of observed data and model simulations. In particular, reanalysis data is expected to perform poorly over Ethiopian region because of the highly heterogeneous rainfall pattern which is majorly attributed to its topography. Various pressure systems and circulation patterns predominantly influence the spatio-temporal characteristics of Ethiopian rainfall, hence limiting its predictability.



**Figure 9.** Frequency distribution of dry spell lengths of observed and reanalysis rainfall for stations (a) Awassa (b) Mekele (c) Gore (d) Bahir Dar (e) Dire Dawa and (f) Addis Ababa (Zoomed version of moderate dry spell frequency are shown as insets).

Long period reliable observed meteorological data over this region should be made available to improve the reanalysis model outputs like precipitation. In addition, reanalysis products through improved spatio-temporal resolution and model modifications would be able to remove the above highlighted deficiencies. Nevertheless, further investigation on climatic variables other than the precipitation needs to be performed for a holistic accuracy analysis.



**Figure 10.** Comparison of (a) mean wet spell length and (b) mean dry spell length in days for all stations.

As presented above, the detailed performance analysis highlights the superiority of ERA-Interim and CFSR among other reanalysis datasets in replicating the rainfall characteristics over Ethiopian region and hence may be employed in performing hydrologic and climate modelling studies that go with their level of accuracy and resolution in the region.

## Acknowledgements

We would like to acknowledge the following organizations for providing the datasets: National Meteorological Agency (NMA) of Ethiopia for synoptic station precipitation data; ECMWF for ERA-Interim data; NASA for MERRA; and NCEP/NCAR for Reanalysis II datasets and Climate Forecast System reanalysis (CFSR). We would also like to thank the editor, associate editor and anonymous reviewers for their valuable inputs which helped in significantly improving the manuscript.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Kite GW, Haberlandt U (1999) Atmospheric model data for macroscale hydrology. *J Hydrol* 217: 303-313.
2. Stocker TF, Qin D, Plattner GK, et al (2013) Climate change 2013: The physical science basis. Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5), New York: Cambridge Univ Press.
3. ABABA A (2007) Climate change national adaptation programme of action (Napa) of Ethiopia. National Meteorological Services Agency, Ministry of Water Resources, Federal Democratic Republic of Ethiopia, Addis Ababa.
4. Feyen J, Zambrano RFV (2015) Modeling hydrological consequences of climate and land use Change-Progress and Challenges. *Maskana* 2: 83-100.

5. Kim J, Waliser DE, Mattmann CA, et al (2014) Evaluation of the CORDEX-Africa multi-RCM hindcast: systematic model errors. *Climate dynamics* 42: 1189-1202.
6. Wagner PD, Fiener P, Wilken F, et al (2012) Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *J Hydrol* 464: 388-400.
7. Kite GW, Haberlandt U (1999) Atmospheric model data for macroscale hydrology. *J Hydrol* 217: 303-313.
8. Petty GW, Krajewski WF (1996) Satellite estimation of precipitation over land. *Hydrol Sci J Taylor & Francis* 41: 433-451.
9. Stisen S, Sandholt I (2010) Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. *Hydrol Proc* 24: 879-891.
10. Zhang X, Alexander L, Hegerl GC, et al (2011) Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdiscip Rev Clim Change* 2: 851-870.
11. Abdo KS, Fiseha BM, Rientjes THM, et al (2009) Assessment of climate change impacts on the hydrology of Gilgel Abay catchment in Lake Tana basin, Ethiopia. *Hydrol Process* 23: 3661-3669.
12. Beyene T, Lettenmaier DP, Kabat P (2010) Hydrologic impacts of climate change on the Nile River Basin: implications of the 2007 IPCC scenarios. *Clim Change* 100: 433-461.
13. Dibike YB, Coulibaly P (2005) Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *J Hydrol* 307: 145-163.
14. Hulme M, Doherty R, Ngara T, et al (2001) African climate change: 1900-2100. *Clim Res* 17: 145-168.
15. Fleming RJ, Kaneshige TM, McGovern WE (1979) The global weather experiment 1. The observational phase through the first special observing period. *Bull Am Meteorol Soc* 60: 649-661.
16. Fellous JL (2008) Towards a global climate observing system. *Interdiscip Sci Rev* 33: 83-94.
17. Bengtsson L, Shukla J (1988) Integration of space and in situ observations to study global climate change. *Bull Am Meteorol Soc* 69: 1130-1143.
18. Bengtsson L, Hagemann S, Hodges KI (2004) Can climate trends be calculated from reanalysis data? *J Geophys Res: Atmos* 109: D11.
19. Mooney PA, Mulligan FJ, Fealy R (2011) Comparison of ERA-40, ERA-Interim and NCEP/NCAR reanalysis data with observed surface air temperatures over Ireland. *Int J Clim* 31: 545-557.
20. Kistler R, Collins W, Saha S, et al (2001) The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull Am Meteorol Soc* 82: 247-267.
21. Uppala SM, Kållberg PW, Simmons AJ, et al (2005) The ERA-40 re-analysis. *Q J R Meteorol Soc* 131: 2961-3012.
22. Dee DP, Uppala SM, Simmons AJ, et al (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137: 553-597.
23. Rienecker MM, Suarez MJ, Gelaro R, et al (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. *J Clim* 24: 3624-3648.
24. Ebita A, Kobayashi S, Ota Y, et al (2011) The Japanese 55-year Reanalysis "JRA-55": an interim report. *Sola* 7: 149-152.
25. Kobayashi S, Yukinari OTA, Harada Y, et al (2015) The JRA-55 reanalysis: General specifications and basic characteristics. *J Meteorol Soc Japan Ser II* 93: 5-48.
26. Kalnay E, Kanamitsu M, Kistler R, et al (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77: 437-471.



27. Kanamitsu M, Ebisuzaki W, Woollen J, et al (2002) Ncep-doe amip-ii reanalysis (r-2). *Bull Am Meteorol Soc* 83: 1631-1643.
28. Saha S, Moorthi S, Pan H-L, et al (2010) The NCEP climate forecast system reanalysis. *Bull Am Meteorol Soc* 91: 1015-1057.
29. Saha S, Moorthi S, Wu X, et al (2014) The NCEP climate forecast system version 2. *J Clim* 27: 2185-2208.
30. Dee DP, Uppala S (2009) Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q J R Meteorol Soc* 135: 1830-1841.
31. Poli P, Healy SB, Dee DP (2010) Assimilation of Global Positioning System radio occultation data in the ECMWF ERA-Interim reanalysis. *Q J R Meteorol Soc* 136: 1972-1990.
32. Behringer DW, Xue Y (2004) Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. Proc. Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface.
33. Penny SG, Behringer DW, Carton JA, Kalnay E (2015) A hybrid global ocean data assimilation system at NCEP. *Mon Weather Rev* 143: 4660-4677.
34. Diro GT, Grimes DIF, Black E, et al (2009) Evaluation of reanalysis rainfall estimates over Ethiopia. *Int J Clim* 29: 67-78.
35. Zhao Y, Li J (2006) Discrepancy of mass transport between the Northern and Southern Hemispheres among the ERA-40, NCEP/NCAR, NCEP-DOE AMIP-2, and JRA-25 reanalysis. *Geophysical res letters* 33.
36. Thorne PW, Vose RS (2010) Reanalyses suitable for characterizing long-term trends. *Bull Am Meteorol Soc* 91: 353.
37. Bosilovich MG (2013) Regional climate and variability of NASA MERRA and recent reanalyses: US summertime precipitation and temperature. *J Appl Meteorol Clim* 52: 1939-1951.
38. Chen G, Iwasaki T, Qin H, Sha W (2014) Evaluation of the warm-season diurnal variability over East Asia in recent reanalyses JRA-55, ERA-Interim, NCEP CFSR, and NASA MERRA. *J Clim* 27: 5517-5537.
39. Lin R, Zhou T, Qian Y (2014) Evaluation of global monsoon precipitation changes based on five reanalysis datasets. *J Clim* 27: 1271-1289.
40. Huang D, Zhu J, Zhang Y, et al (2016) Assessment of summer monsoon precipitation derived from five reanalysis datasets over East Asia. *Q J R Meteorol Soc* 142: 108-119.
41. Hutchinson MF (1998) Interpolation of rainfall data with thin plate smoothing splines. Part I: Two dimensional smoothing of data with short range correlation. *J Geogr Inf Decis Anal* 2: 139-151.
42. Tsidu GM (2012) High-resolution monthly rainfall database for Ethiopia: homogenization, reconstruction, and gridding. *J Clim* 25: 8422-8443.
43. Nijssen B, O'Donnell GM, Lettenmaier DP, et al (2001) Predicting the discharge of global rivers. *J Clim* 14: 3307-3323.
44. Prakash S, Gairola RM, Mitra AK (2015) Comparison of large-scale global land precipitation from multisatellite and reanalysis products with gauge-based GPCC data sets. *Theor Appl Clim* 121: 303-317.
45. Sylla MB, Coppola E, Mariotti L, et al. (2010) Multiyear simulation of the African climate using a regional climate model (RegCM3) with the high resolution ERA-interim reanalysis. *Clim Dyn* 35: 231-247.

46. Wu M-LC, Reale O, Schubert SD, et al. (2012) African easterly jet: Barotropic instability, waves, and cyclogenesis. *J Clim* 25: 1489-1510.
47. Güntner A (2008) Improvement of global hydrological models using GRACE data. *Surv Geophys* 29: 375-397.
48. Diaconescu EP, Gachon P, Scinocca J, et al. (2015) Evaluation of daily precipitation statistics and monsoon onset/retreat over western Sahel in multiple data sets. *Clim Dyn* 45: 1325-1354.
49. Reda DT, Engida AN, Asfaw DH, et al. (2015) Analysis of precipitation based on ensembles of regional climate model simulations and observational databases over Ethiopia for the period 1989-2008. *Int J Clim* 35: 948-971.
50. Sen PK (1968) Estimates of the regression coefficient based on Kendall's tau. *J Am Stat Assoc* 63: 1379-1389.
51. Cheung WH, Senay GB, Singh A (2008) Trends and spatial distribution of annual and seasonal rainfall in Ethiopia. *Int J Clim* 28: 1723-1734.
52. Seleshi Y, Zanke U (2004) Recent changes in rainfall and rainy days in Ethiopia. *Int J Clim* 24: 973-983.
53. Funk CC, Rowland J, Eilerts G, et al. (2012) A climate trend analysis of Ethiopia. *US Geol Sur*.
54. Walker S, Getahun YG, Tesfaye K, et al. (2003) The use of agroclimatic zones as a basis for tailored seasonal rainfall forecasts for the cropping systems in the Central Rift-valley of Ethiopia. *Insights and Tools for Adaptation: Learning from Climate Variability*.
55. Tadesse T (2000) Drought and its predictability in Ethiopia, In: Wilhite, D.A, *Drought: A Global Assessment* 135-142.
56. National Meteorological Services Agency (NMSA) (1996) Climate and Agro climatic Resources of Ethiopia. Meteorological research report series 1.
57. Xue X, Hong Y, Limaye AS, et al. (2013) Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J Hydrol* 499: 91-99.
58. Belachew A (2002) Practical approach to upgrade the hydro-meteorological network in Ethiopia. *Ethiopian J. Water Sci Tech* 3.
59. Barnett V, Lewis T (1994) *Outliers in statistical data*, New York: Wiley.
60. Zaroug MAH, Giorgi F, Coppola E, et al. (2014) Simulating the connections of ENSO and the rainfall regime of East Africa and the upper Blue Nile region using a climate model of the Tropics. *Hydrol Earth Syst Sci Discuss* 11: 2233-2262.
61. Lanzante JR (1996) Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *Int J Clim* 16: 1197-1226.
62. González-Rouco JF, Jiménez JL, Quesada V, et al. (2001) Quality control and homogeneity of precipitation data in the southwest of Europe. *J Clim* 14: 964-978.
63. Wagesho N, Goel NK, Jain MK (2013) Temporal and spatial variability of annual and seasonal rainfall over Ethiopia. *Hydrol Sci J* 58: 354-373.
64. AghaKouchak A, Mehran A (2013) Extended contingency table: Performance metrics for satellite observations and climate model simulations. *Water Resour Res* 49: 7144-7149.
65. Kennedy AD, Dong X, Xi B, et al (2011) A comparison of MERRA and NARR reanalyses with the DOE ARM SGP data. *J Clim* 24: 4541-4557.



66. Dawit A, Future climate of Ethiopia from PRECIS Regional Climate Model Experimental Design 2010. Available from:  
[http://www.metoffice.gov.uk/media/pdf/o/9/PRECIS\\_Experimental\\_Design\\_Dawit.pdf](http://www.metoffice.gov.uk/media/pdf/o/9/PRECIS_Experimental_Design_Dawit.pdf).
67. Suarez MJ, Rienecker MM, Todling R, et al (2008) The GEOS-5 Data Assimilation System-Documentation of Versions 5.0.1, 5.1.0, and 5.2.0.
68. Hartkamp AD, De Beurs K, Stein A, et al. (1999) Interpolation techniques for climate variables. Geographic Information Systems Series 99-01.
69. Ratan R, Venugopal V (2013) Wet and dry spell characteristics of global tropical rainfall. *Water Resour Res* 49: 3830-3841.
70. Douguedroit A (1987) The variations of dry spells in Marseilles from 1865 to 1984. *J Clim* 7: 541-551.



AIMS Press

© 2017 C.T. Dhanya et al., licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)