Bulletin de l'Institut Scientifique, Rabat, Section Sciences de la Terre, 2025, n° 47, 25–38 ***e-ISSN : 2458-7184***

Performance comparison of Climate Research Units and Global Precipitation Climatology Centre datasets over the Awash River Basin, Ethiopia

Comparaison des performances des unités de recherche sur le climat et des ensembles de données du Centre mondial de climatologie des précipitations dans le bassin de la rivière Awash, en Éthiopie

Solomon Mulugeta[1*] and Clifford Fedler[2]

1. *Dire Dawa, Dire Dawa University, civil engineering department, Sabiyan, Dire Dawa, Ethiopia, (* solomon.mulugeta@ddu.edu.et; msteffera@gmail.com)*
2. *Texas Tech University, Department of Civil, Environmental, and Construction Engineering, Lubbock, TX 79409–1023, USA*

Abstract. The conventional sources of climate data are observation data from weather stations. However, there are a number of derived datasets with varying reliability that can be considered as alternative data sources especially for areas observational data are available or sparsely distributed. Among the several datasets, the two most commonly used derived datasets are Climate Research Units (CRU) and Global Precipitation Climatology Centre (GPCC). The purpose of this study was to compare the performance of two gridded global datasets (GPCC-Total Full V2018 and (CRU) Ts version 4.01) in simulating observed rainfall data from 10 selected weather stations in the Awash River Basin, Ethiopia.The performance of the GPCC and CRU datasets were evaluated by comparing the monthly rainfall data extracted from the two datasets with observation data (for period 1960- 2016) from the weather stations using statistical and graphical methods. Pearson correlation coefficient (CC), root mean square error (RMSE), scatter plots, box plots, Taylor diagram, and Kolmogorov-Smirnov test were used. Prior to using the gauge data for evaluation, the data were tested for temporal homogeneity using the standard normal homogeneity test (SNHT), the Buishand range test, and the Pettitt test. It was found the GPCC dataset showed generally higher correlations with gauge data (CC=0.85 for monthly time series) and lower errors (with averaged RMSE=45 mm/month) than that of the CRU dataset (CC=0.82 for monthly time series, with averaged RMSE over location =51 mm/mm). However, it was found that majority of the monthly rainfall data from both the CRU and GPCC datasets (71.2% for the CRU and 66% for the GPCC) failed to follow probability distribution with observation data.Still, it is clear that the GPCC dataset showed similar probability distribution with observed data for more number of time series than that of the CRU dataset. All graphical analysis also showed that GPCC dataset aligns more closely with gauge data than of the CRU. Overall, the GPCC dataset has showed better performance than the CRU dataset to simulate rainfall for the Awash River Basin. Relative to the CRU dataset, the GPCC dataset reduces the RMSE by an average of ~11% for monthly rainfall and by ~13.4% for annual rainfall in the Awash River basin. The GPCC dataset can be used as an alternative source of rainfall data for hydrological analysis and modelling required in the planning and design of water infrastructure, management of water resources, and climate and hydrological studies in the basin, especially for ungauged and data-scarce areas of the river basin. Further studies are crucial to identify datasets that can perform betteracross locations and seasons in reducing errors and bias and in replicating the probability distribution of observation data.

Key words: Rainfall (CRU, GPCC, CC), Statistical and graphical methods (SNHT, RMSE, CC), Taylor diagram, Kolmogorov-Smirnov test, performance, Awash River Basin (Ethiopia).

Résumé. Les sources conventionnelles de données climatiques sont les données d'observation provenant des stations météorologiques. Cependant, il existe un certain nombre de jeux de données dérivés de fiabilité variable qui peuvent être considérés comme des sources de données alternatives, en particulier pour les zones où les données d'observation sont disponibles ou peu distribuées. Parmi les nombreux ensembles de données, les deux ensembles de données dérivés les plus couramment utilisés sont les Climate Research Units (CRU) et le Global Precipitation Climatology Centre (GPCC). Le but de cette étude était de comparer la performance de deux ensembles de données mondiales en grille (GPCC- Total Full V2018 et (CRU) Ts version 4.01) dans la simulation des données de précipitations observées provenant de 10 stations météorologiques sélectionnées dans le bassin de la rivière Awash, en Éthiopie. La performance des ensembles de données GPCC et CRU a été évaluée en comparant les données mensuelles de précipitations extraites des deux ensembles de données avec les données d'observation (pour la période 1960-2016) des stations météorologiques en utilisant des méthodes statistiques et graphiques. Le coefficient de corrélation de Pearson (CC), l'erreur quadratique moyenne (RMSE), les nuages de points, les diagrammes en boîte, le diagramme de Taylor et le test de Kolmogorov-Smirnov ont été utilisés. Avant d'utiliser les données de jauge pour l'évaluation, les données ont été testées pour l'homogénéité temporelle en utilisant le test d'homogénéité normale standard (SNHT), le test de portée de Buishand et le test de Pettitt. Il a été constaté que l'ensemble de données GPCC montrait généralement des corrélations plus élevées avec les données de jauge (CC=0,85 pour les séries temporelles mensuelles) et des erreurs plus faibles (avec une RMSE moyenne = 45 mm/mois) que l'ensemble de données CRU (CC=0,82 pour les séries temporelles mensuelles, avec une RMSE moyenne par emplacement = 51 mm/mois). Cependant, il a été constaté que la majorité des données mensuelles de précipitations des ensembles de données CRU et GPCC (71,2 % pour le CRU et 66 % pour le GPCC) ne suivaient pas la distribution de probabilité avec les données d'observation. Cependant, il est clair que le jeu de données GPCC a montré une distribution de probabilité similaire aux données observées pour un plus grand nombre de séries temporelles que le jeu de données CRU. Toutes les analyses graphique sont également montré que l'ensemble de données GPCC s'aligne plus étroitement avec les données de jauge que celui du CRU. Dans l'ensemble, le jeu de données GPCC a montré de meilleures performances que le jeu de données CRU pour simuler les précipitations dans le bassin de la rivière Awash. Par rapport au jeu de données CRU, le jeu de données GPCC réduit le RMSE en moyenne de ~11% pour les précipitations mensuelles et de ~13,4% pour les précipitations annuelles dans le bassin de la rivière Awash. Le jeu de données GPCC peut être utilisé comme une source alternative de données de précipitations pour l'analyse hydrologique et la modélisation nécessaires à la planification et à la conception des infrastructures hydrauliques, à la gestion des ressources en eau, ainsi qu'aux études climatiques et hydrologiques dans le bassin, en particulier pour les zones non jaugées et pauvres en données du bassin fluvial. Des études supplémentaires sont cruciales pour identifier des ensembles de données qui peuvent mieux performer à travers les emplacements et les saisons en réduisant les erreurs et les biais et en reproduisant la distribution de probabilité des données d'observation.

Mots clés : Précipitations (CRU, GPCC), méthodes statistiques et graphiques (SNHT, RMSE, CC), diagramme de Taylor, test de Kolmogorov-Smirnov, performance, Bassin de la rivière Awash (Éthiopie).

INTRODUCTION

The conventional source of climate data are measurements from weather stations. However, obtaining reliable climate data from weather stations is not always easy for a number of reasons. Absent or sparsely distributed weather stations, lack of long-term data records from weathers stations, large missing data and inhomogeneity in data from weather stations are some of the major reasons making it difficult to obtain of climate data from weather stations (Tsidu 2012, Dinku *et al*. 2014, Ahmed *et al*. 2019, Van Vooren *et al*. 2019).

Ethiopia is among the countries with insufficient number of weather stations. There were about 910 weather stations in Ethiopia until recent times. New stations have added in recent years, currently there are about 1700 manned and automatic weather stations in Ethiopia (Ethiopian Meteorological Institute 2025). However, most of the weather stations are located in the central and western highland areas of the country. There is much smaller number of weather stations in the eastern low land areasof the country, which consists of most of the area of the Awash River basin. In addition, the weather stations in these areas are relatively new, making it difficult to get long year's climate data. Ethiopian Meteorological Institute has planned to increase the number of rainfall recording stations to 2589 until 2030 in order to meet the standard of world meteorological agency(National Meteorological Agency 2020).Together with this ongoing effort to meet the need, evaluating alternative data sources is crucial.

Alternative data sources have been explored for climate variables. Derived datasets are potential alternative data sources (Sun *et al*. 2018). Derived datasets include gauge-based, satellite-related, reanalysis, and hybrids of these datasets. Among these derived datasets, gauge based derived datasets are widely used in climate studies (Merino *et al*. 2021). In Ethiopia, a number of studies used these datasets (Wagesho *et al*. 2013, Dinku *et al.* 2014, Asfaw *et al*. 2018, Mulugeta *et al*. 2019). Gauge- based datasets are directly derived from actual meteorological observations. Because of this they show superior performance relative to others especially for areas with having good distribution and coverage of high quality weather stations. For this reason, they are also considered as the primary source to validate other products such assatellite and reanalysis data. Gauge based derived datasets have relative advantages over other derived dataset in that they have inherited well defined errors and uncertainties from the observation data and interpolation methods.Several gauge-based datasets have been developed so far (Sun *et al.* 2018, Ahmed *et al.* 2019). However, the Global Precipitation Climatology Centre (GPCC) datasets, and the Climatic Research Unit (CRU) datasets are the most widely used (Hu *et al*. 2018, Sun *et al*. 2018). Global Precipitation Climatology Centre (GPCC) uses data from large number of weather station, each of the climatology product are based on data approximately from 85,000 weather stations. The CRU datasets are developed by the University of East Anglia and comprise of data for a number of climate variables including rainfall. The CRU datasets are constructed based on monthly gauge data from National Meteorological Agencies (NMAs), the World meteorological organizations (WMO), the Centro Internacional de Agricultura Tropical, the Food and Agriculture Organization (FAO), and others. Data from around 4000 weather stations across the world are used for construction of the datasets. While the GPCC is developed at DeutscherWetterdienst in collaboration with the World Meteorological Organization (WMO) based on daily and monthly data from National Meteorological Agencies (NMAs), the World meteorological organizations (WMO), the CRU, the Food and Agriculture Organization (FAO), and the National Centers for Environmental Information as well as from international regional projects. Data acquired from more than 85.000 stations across the world are used to construct the datasets (Hu *et al.* 2018. Sun *et al*. 2018). There are different versions of datasets available for both the CRU and the GPCC datasets. Both datasets can provide high resolution (0.5x 0.5°) gridded global data for a period of more than 100 years. As the results, they are regularly used by climate scientists. They have been widely used in long-term climatic trend analysis and as "baseline" datasets for validations of other model outputs and satellite products (Hu *et al*. 2018, Sun *et al*. 2018, Ahmed *et al*. 2019). However, there are inherent biases and uncertainties in those datasetsassociated with density and coverage of the observation station, quality of observation data, amount of missing data, and the interpolation method used in building the datasets.Therefore, it is required to evaluate the reliability of datasets over a particular location.

A number of studies have also evaluated the Performances of gridded datasets for locations inEthiopia and for the East African region**.** Dinku et *al.* (2008) evaluated five gauge-based gridded products (three different products of GPCC, CRU, and CPC) against gauge data over the Ethiopian Highlands. They reporteda very good agreement between the datasets and gauge data. They also reported reasonably low systematic and random errors for the climate products.Woldemariam *et al.* (2017) evaluated four widely used reanalysis datasets (ERA-Interim (The European Centre for Medium-Range Weather Forecasts (**E**CMWF) **R**e-**A**nalysis, MERRA (Modern Era Retrospective-Analysis), NCEP-DOE R2 and CFSR) against on gauge data across Ethiopia. They found that the ERA-Interim and CFSR perform better in showing the different characteristics of daily rainfall over Ethiopia. Gebrechorkos *et al.*(2018) evaluated five climate datasets from multiple data sources [Africa Rainfall Climatology version 2.0 (ARC2), Climate Hazards Group InfraRed Precipitation (CHIRP), CHIRP with Station data (CHIRPS), Observational-Reanalysis Hybrid (ORH), and regional climate models (RCMs)] against gauge data in east African region. They found that CHIRPS, CHIRP, and ARC2 showed best performance, while ORH, I-RCM, and RCMs showed the worst performance. Asfaw *et al.*(2018) evaluated two datasets (GPCC and CRU) against station data over the north central Ethiopia using correlation coefficient and Kolmogorov-Smirnov test .

They found that data from GPCC dataset showed higher correlation with gauge data than that from the CRU dataset. They found that the probability distribution GPCC data follow the same distribution with observed data, but that of the CRU do not follow the same distribution with observed data.

The studies cited above have different issues limiting the information generated from them to be less useful for the Awash River basin. Dinku et *al.* (2008) evaluated the five gauge-based gridded products based ondata from large number of weather stations across Ethiopia. This study was the first of its kind and provided crucial information on performance of the datasets. However, firstly, the study was carried out at national level, so, no particular empathize can be given to the Awash River basin. Secondly, the study did not compare the probability distribution of gridded dataset with the observation data.Thirdly, one of the most important statistical parameter, the root mean square error was not employed in the study. Lastly, the study is a bit older now, indicating that the recent version of the datasets might be needed to be evaluated. Woldemariam *et al.* (2017) and Gebrechorkos *et al.* (2018) did not include the CRU and GPCC in their study. In addition, both of these studies are carried out for national and regional geographical areas, lacking details for the basin of interest in our study, the Awash River basin.

On the other hand, Asfaw *et al.* (2018) evaluated the GPCC and CRU over the Woleka sub-basin, in eastern part of the Abay River basin (which is not part of the Awash river basin). Asfaw *et al.* (2018) compared GPCC and CRU using correlation and correlation coefficient and Kolmogorov-Smirnov test alone. Some important performance metrics such as statistical errors and bias are not included. Therefore, it is important to evaluated performance of the recent version of GPCC and the CRU datasets at better level of detail and focus for the Awash River basin by employing a number of important statistical and graphical methods. To this end, this research was initiated to compare the performance of two globally commonly used gauge based global gridded datasets, the Global Precipitation Climatology Centre (GPCC) and the Climate Research Unit (CRU), against observation data over the Awash River Basin.

Description of the study area

Awash River Basin (ARB) is located in Ethiopia between 7°53’ N to 12° N and 37°57’ E to 43°25’ E. It covers an area of 116.374 square-km extending from the central highlands as high as 4195 m above mean sea level (M.S.L) to the lower arid areas as low as 210 m above M.S.L in the Danakil Depression. The basin is the most important basin in Ethiopia for the river in this basin isone of the most utilized river in Ethiopia. Thisriver is the source of water for most of the major cities, irrigated farms, and agro-industries of the country. The climate in the basin is characterized as humid subtropical at the upstream region, semi-arid in the middle region, and arid at the downstream region of the River Basin. The land uses in the Basin include around 50% agricultural land, 39% grassland and shrubs, and the remaining 11% under various uses *(Awulachew et al.* 2007, Kerim *et al.* 2016). Traditionally, the Basin is divided as the Upper, the Middle, and the Lower Awash. However, hydrologically, the Basin is divided into seven sub-basins (Fig.1). Figure 1 shows the major river basins in Ethiopia including the Awash River Basin and the sub-basins in of the ARB. The western 64.000 km^2 of the Basin contributes nearly the entire surface runoff of the Basin (Adeba *et al.* 2015). High stream distribution and density in the western side of the Basin (Fig. 1) also shows the relatively high contribution of surface runoff by the western side of the basin.

Data and methods

Monthly gauge rainfall data were obtained from the National Metrological Agency (NMA) of Ethiopia for 23 major gauge stations across the basin (Fig. 2).The major gauge stations considered in this study are those of class- A stations and having at least 30 years weather data. The percentage missing data were estimated for each station. Stations with data length not less than 30 years and with missing data less than 10% were originally targeted for selection. However, as the number of station meeting the criteria dropped significantly, stations with missing data around 15% were alsoincluded. Dinku *et al.*(2008) have used similar criteria among other.

They removed time-series less than 15 years and with excessive missing data. The missing observation data for the remaining gauge stations were replaced by linear interpolation. This method of interpolation is used to estimate values of unknown point from values of known data points based on linear relationship. It may not be accurate for rainfall due to his high natural variability. However, it can provide reasonable estimate for rainfall with limited missing data and a narrow time interval. In addition, the method is often the most practical especially underno nearby weather station where more precise spatial interpolation method such as Kriging interpolation would be difficult to carry out.Then, the temporal homogeneity of the data from selected stations were tested using the standard normal homogeneity test (SNHT), the Buishand range test, Pettitt test, and the Von Neumann ratio test. The tests are based a null hypothesis - the time series are homogenous. These methods are widely used to test temporal homogeneity of gauge data (Kang & Yusof 2012, Javari 2016, Ahmed *et al.* 2019).

The homogeneity tests were carried out because inhomogeneity in the data from weather station needs to be checked before using the data for further analysis (Zwiers & Zhang 2009). The results of the four tests were used together to classify the level of homogeneity in the time series as ‘useful’(Class A), ‘doubtful’(Class B) and ‘suspect’(Class C) depending on the number of homogeneity tests rejecting the null hypothesis. When none or one of the four tests rejects the null hypothesis, the homogeneity status is classified as Class A, such data can directly be used for further analysis. When any two of the four tests reject the null hypothesis, the homogeneity status of the time series is classified as Class B, suchtime series data have inhomogeneous signal and should be critically inspected before using for further analysis. When three or all of the four tests reject the null hypothesis, the homogeneity status of the time series is classified as Class C , such data cannot be used in further analysis before necessary corrections (Wijngaard *et al.* 2003, AL-Lami *et al*. 2014, Chang & Ghani 2017, Elzeiny *et al*. 2019a).

The NetCDF data files of Global Precipitation Climatology Centre Full Data Reanalysis (GPCC_FD) version Total Full V2018 (0.5x0.5) and the Climate Research Units (CRU) Ts version 4.01, were obtained from online- archives (https://data.ceda.ac.uk for the CRU dataset and https://www.esrl.noaa.gov for GPCC dataset).These versions of both datasets provide global monthly precipitation data at a 0.5° by 0.5° spatial resolution for the period 1901 to 2016. The CRU dataset also provides data for other climate variables such as air temperature and potential evapotranspiration (Schneider *et al*. 2015, Jones 2017).

There are two approaches to compare grid data with station observations: Grid to grid or point to point comparison. For grid to grid comparison, point data from weather stations need to be converted to grid data for the required grid size. For point to point comparison, the grid data need to be interpolated to a location of the point data (Ahmed *et al*. 2019). One method cannot inherently be consider as more accurate than the other. The suitable method depends the specific application and the purpose as both have strength and weakness. A point-to-point comparisons provide a direct, location-specific comparison of values, because of this it is ideal for validating the accuracy of gridded data against particular ground truth locations. A grid-to-grid comparison is suitable for large-scale analysis. However, the density and quality of the station data used to build the grid have a significant impact on its accuracy. In our case, there are limited weather stations over the study area with reliable and long term data within 0.5 degree grid. The performance of the gridded dataset are, thus, to be evaluated,against data over these locations. For these reasons, point to point comparison was applied in this study.

The monthly rainfall data were extracted from the respective NetCDF files of the two gridded datasets for four grid points surrounding each of the weather station. In order to get point data from the grid data, the grid data over the four grid point surrounding each weather station were interpolated to locations corresponding to the weather stations using Inverse distance weighted average (IDWA) interpolation method (power of 2).This interpolation method and parameter (power of 2) have been used for rainfall interpolation for a study in the Awash River basin (Getahun *et al*. 2021).To assess the strength of relationship between observed and grid data, the Pearson correlation coefficient were computed for the entire monthly data series. The correlation coefficient is widely used in similar studies to assess the level relationship between two datasets (Dinku *et al*. 2008, 2014, Asfaw *et al*. 2018, Gebrechorkos *et al.* 2018). Several statistical error and bias estimatorssuch as Mean Bias Errors (MBE), Mean Absolute Error (MAE), Modified Index of Agreement, androot mean square error, and relative efficiency have been used in similar studies (Dinku *et al.* 2008, 2014, Asfaw *et al*. 2018, Gebrechorkos *et al*. 2018, Ahmed *et al*. 2019). These estimators are similar in that they all are related to the magnitude of deviations between values of observed and grid data in different ways. The root mean square error (RMSE). RMSE offers a single, comprehensible measure of prediction error in the same units as the target variable, it is a popular and useful estimator for assessing the accuracy of gridded data. The probability distribution of the observed and gridded data were compared using different goodness-of-fit tests; the two most widely used test include Anderson–Darling (AD) testand Kolmogorov Smirnov (KS) test (Asfaw *et al*. 2018, Ahmed *et al*. 2019, Navidi Nassaj *et al.* 2022). The AD test estimates the deviation between distributions, but it places greater emphasis to the tail of the distributions. Because of this, it is especially useful for analysis of extreme values such as heavy rainfall. Kolmogorov Smirnov (KS) test is non parametric test based on estimation of the maximum difference between the cumulative probability distributions functions (CPDs) of two datasets. This test is sensitive to maximum difference near the center of the probability distributions.In this study, Kolmogorov Smirnov (KS) test was applied to assess the similarity in probability distribution of the observed and gridded data.In addition to statistical methods, graphical methods are also widely used in similar studied. Thus, the graphical methods such as scatter plots, box plots, and Taylor diagram were also used in this study for graphical comparison.

The R-Statistical Software was intensively used in this study for different purposes. It was used to extract the data from NetCDF by applying Rscript codes prepared by Uddameri(2017) , for filling missing data using linear interpolation using the ‘ImputeTS’ package (Moritz 2019), for IDWA interpolation using the ‘gstat’ package (Pebesma & Graeler 2020) , for homogeneity test using the ‘Trend package’ (Pohlert 2018), and for correlation, KS test, RMSE using the ‘Stat4’ package (R Core Team 2018). While the Geographic Information System (GIS) software was used for creating shapefiles, grid making, and mapping.

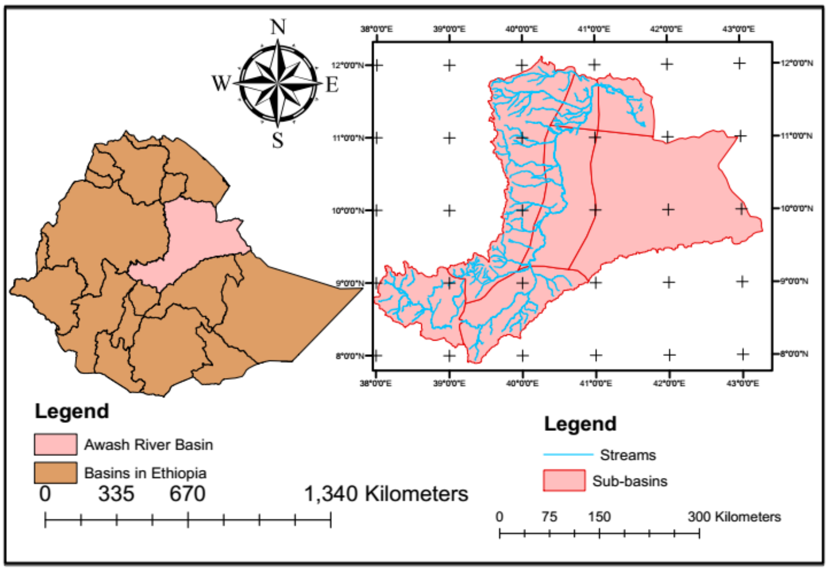

Figure 1. Major River Basins in Ethiopia and the Awash River Basin and its sub-basins.
Figures 1 is based on GIS files from the Ministry of Water Resources of Ethiopia

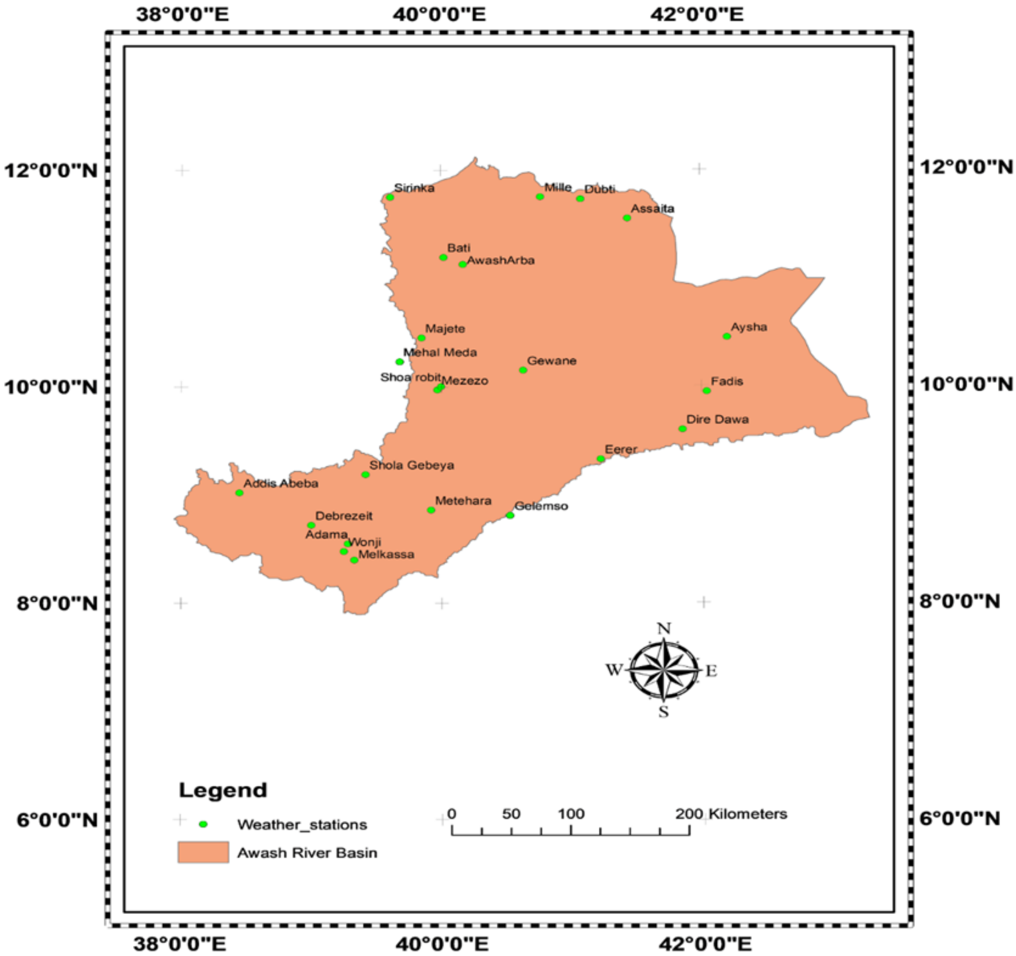

Figure 2. Selected 23 weather stations in the Awash River Basin.
Figures 2. is based on GIS files from the Ministry of Water Resources of Ethiopia and the coordinates of the stations are taken from the website of the National Metrological Agency of Ethiopia: (http://www.ethiomet.gov.et).

RESULTS AND DISCUSSIONS

Estimated percentage missing data

From 23 weather stations under consideration (Fig. 2), ten stations (Tab.1)with the least percentage missing data and relatively longer data records were selected. Stations with data length of 30 years and missing data less than 10% were originally targeted. However, as the number of station meeting the criteria dropped significantly, stations with missing data around 15% were included. The selected weather stations and the respective estimated percentage of missing data are shown in Table 1.

Homogeneity tests

The standard normal homogeneity test (Tab. 2) shows that the monthly rainfall time series are homogenous ($\alpha = 0.05$) for all months across 6 stations (Adama, Addis Ababa, Dire Dawa, Metahara, MehalMeda, and wonji) and for most of the months across the remaining 4 stations. According to this test, around 91% of monthly time series across the selected stations are homogenous ($\alpha = 0.05$).

The Buishand range test (Tab. 3) shows that the monthly rainfall time series are homogenous ($\alpha = 0.05$) for all months across 6 stations (Adama, Addis Ababa, Melkasa, Metahara, Mezezo, and Wonji) and for most of the months across the other 4 stations. According to this test, around 95% of monthly time series across the selected stations are homogenous ($\alpha = 0.05$).

The Pettitt test (Tab. 4) shows that the monthly rainfall time series are homogenous ($\alpha = 0.05$) for all months across 6 stations (Adama, Addis Ababa, Debrezeit, Mezezo, MehalMeda, and Wonji) and for most of the months across the other 4 stations. According to the Pettitt test, around 95 % of monthly time series are across the selected stations homogenous ($\alpha = 0.05$).

The Von Neumann ratio test (Tab. 5) shows that the monthly rainfall time series are homogenous ($\alpha = 0.05$) for all months across 4 stations (Addis Ababa, Debrezeit, Metahara, MehalMeda) and for most of the months across the remaining 6 stations. According to this test, around 90 % of monthly time series across the selected stations are homogenous ($\alpha = 0.05$).

The results show thatabout 94% the tests (112 out of 120) are rejected byone or none of the four tests. These monthly time series data can be classified as Class A, indicating that the data issuitable for further analysis. The tests for 5.8% of the tests (7 out of 120) are rejected by two of the four tests. The data can be classified as Class B, indicating that thehomogeneity of this data is questionable. The tests for a single monthly time series for February at Majete station (less than 1 %) is rejected by three of the four tests, the data can be classified as Class C, indicating that the data is spurious and cannot used for some analysis such as trend analysis. This means 94 % of the time series data are homogenous while around 6% of them have some signal of inhomogeneity. It is well known that inhomogeneity in data is potentially attributed to non-climatic factors such as relocation of weather stations, changes in measuring instruments, significant change in site surrounding condition, and alterations in observation procedures. The surrounding areas of the stations in Dire Dawa and Melkasa are known to have changed much due to increased urbanization around the stations which were known to be in the outskirt of the city area in old times. The station at Majate might have potentially also experienced similar change of surrounding area or update of instrumentation due to the fact that the station is among the oldest stations in Ethiopia.

Homogeneity testsfor monthly rainfall time series for locations in the Awash river Basinshave recently been investigated by a couple of studies. Similar to this study, Daba *et al*. and Getahun *et al*. (2021) employed the four tests (standar normal homogeneity test (SNHT), the Buishand range test, Pettitt test, and the Von Neumann ratio test), and classified the time series as class A, B, and C. Daba *et al*. (2020) assessed homogeneity of monthly as well asannual rainfall time series for stations in the upper Awash river basin. They have not indicated the percentage of time series in each classes, however, from the number of time series they provided the parentage of time series in each class can be estimated.

Accordingly, about 65.2% of the monthly time series are of class A, about 32.4% of the monthly time series are ofclass B, and about 2.3% of the monthly time series are of class C. Similarly, Getahun *et al*. (2021) assessed the homogeneity of basin- average monthly rainfall time series for Awash river basin to identify change detection periods. They found monthly time series of all months except for February and April are of class A. The time series for the month of February is found as of class C and that of April is found as class of B. Accordingly, more than 83% of the time series in this study are of class A. Some other studies have also investigated the homogeneity for annual rainfall time series for locations in Awash river basin (Adane *et al*. 2020, Edris *et al.* 2021). Edris *et al*. (2021) used all the four methodsused in this study while Adane *et al.* (2020) used double mass curve for the homogeneity testing. Edris *et al*. (2021) found that the annual rainfall times series are homogenous for 90% of the stations. It is clear that homogeneity is easier to achieve for annual time series than monthly time series data.Direct comparison of the results may be difficult due to difference data length, location, methods used, and spatial extent of analysis. However, generally the finding for monthly time series in this study showed higher level of homogeneity.The results imply that the monthly rainfall data for the selected weather stations in the Awash River Basin have not been significantly affected by artificial factors such as dislocation of weather stations, change in measurement instruments, procedure, and calibration and maintenance of the measuring devices. Therefore, the data from selected weather stations can be used as reference to assess the performance derived gridded datasets.

Correlation analysis

The entire monthly rainfall data from the gridded datasets were compared with observation data at the selected weather stations. The results of Pearson correlation (Tab. 6) show that rainfall data extracted from both the CRU and the GPCC datasets are generally significantly and highly correlated with the corresponding observation data. The GPCC showed higher correlation with observed data than the CRU for all stations except at Adama. Still, on average, the GPCC showed higher correlation coefficient than that of the CRU. The scatter plots of observed data versus datasets in Figure 3 shows that data from both the CRU and the GPCC datasets are not only highly and positively correlated to the observed data but also more closely and linearly follow the trend of the observed data. However, data from the GPCC dataset shows stronger positive correlation ($R^2 = 0.94$ for GPCC) with observed data than the CRU (with $R^2 = 0.88$ for CRU) (Fig. 4). In addition, the GPCC data follows the trend of the observed datamore closely with smaller number of outliers than that of the CRU. The box plot in Figure 5 shows the median value of the correlation coefficient for the GPCC is higher and with range smaller variability than that of the CRU.Overall, the GPCC dataset shows higher level of association with observed data than that of the CRU dataset.

Dinku *et al.* (2008) found an average correlation coefficients of 0.95 between observed rainfall data and GPCC-full and that of 0.90 between observed rainfall data and CRU for locations in the central Ethiopian highlands.

Table 1. Selected weather stations and estimated percentage missing data.

SN	Gauge Station	Latitude N	Longitude E	Data Period	Missing Data %
1	Addis Ababa	9.02	38.50	2015–1960	2.6
2	DebreZeit	8.72	39.00	2013–1960	13.1
3	Dire Dawa	9.60	41.85	2015–1960	2.1
4	MehalMeda	10.23	39.68	2015–1974	13.1
5	Melkasa	8.40	39.33	2013–1977	0.67
6	Metahara	8.86	39.92	2015–1984	7.2
7	Mezezo	9.97	39.97	2015–1986	16.4
8	Adama	8.55	39.28	2015–1980	14.3
9	Majete	10.45	39.85	1988-2015	8.3
10	Wonji	8.48	39.25	1960-2015	10.1

(The Coordinates of the stations are taken from the website of the National Metrological Agency of Ethiopia: http://www.ethiomet.gov.et).

Table 2. Test statistics (T) for SNHT over the gauge stations (* denotes T is statistically significant at 5% significance level).

Stations	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adama	4.2	3.9	2.1	2.1	4.1	7.0	6.1	6.8	2.9	1.3	3.4	4.4
Addis Ababa	3.2	3.4	2.5	2.2	1.0	2.3	3.8	4.1	5.5	1.2	2.3	2.2
Dire Dawa	3.6	6.8	3.2	5.1	1.8	4.1	6.2	7.2	3.1	3.8	7.1	4.0
DebreZeit	1.8	2.9	3.6	4.5	3.5	5.6	5.1	10.7*	11.8*	2.1	3.3	5.2
Melkasa	12.7*	3.6	2.1	10.0*	2.4	10.4*	10.4*	3.0	11.9*	15.6*	3.4	2.3
Metahara	2.0	7.6	5.0	7.1	5.4	5.4	3.2	3.9	7.7	3.0	7.4	1.0
Majete	5.7	14.7*	3.7	2.8	3.9	4.3	7.01	3.0	3.7	3.3	2.3	2.4
MehalMeda	1.7	6.0	2.7	3.9	1.2	1.1	4.8	2.2	6.5	5.2	0.8	1.1
Mezezo	4.2	10.9*	4.0	3.7	8.0*	4.5	6.7	1.7	3.9	2.4	4.0	2.6
Wonji	2.0	3.2	5.1	4.5	5.4	7.3	6.6	3.4	6.1	1.4	2.3	3.2

Table 3. Test statistics (R/sqrt (n)) for Buishand range test over the gauge stations (* denotes R/sqrt(n) is statistically significant at 5% significance level, P-value is less than 0.05).

Stations	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adama	0.9	1.3	0.9	1.0	1.4	1.3	1.3	1.2	1.1	1.1	1.0	0.9
Addis Ababa	1.1	1.3	0.7	0.9	0.8	0.9	0.9	1.2	1.0	1.0	1.1	1.0
Dire Dawa	1.4	1.5	1.1	1.6*	1.0	1.2	1.8*	1.3	1.1	1.1	1.2	1.2
DebreZeit	0.9	1.1	0.8	1.3	1.2	1.0	1.3	1.5	1.7*	1.0	0.9	1.6*
Melkasa	1.3	0.9	1.1	1.0	1.1	1.2	0.9	0.8	1.5	1.2	1.2	1.1
Metahara	0.9	1.4	1.3	1.2	1.2	0.9	1.0	1.1	1.4	1.0	1.2	0.8
Majete	1.2	1.6*	1.1	0.8	0.8	1.3	1.4	1.0	0.8	1.1	0.9	0.9
MehalMeda	0.9	1.5	1.0	0.9	0.8	0.6	1.7*	0.9	0.8	0.9	0.8	1.0
Mezezo	1.1	1.5	1.0	0.8	1.3	1.3	1.3	1.0	1.1	1.2	1.1	1.1
Wonji	1.0	1.3	1.4	1.0	1.0	0.9	1.0	0.9	1.4	1.0	1.2	1.3

Table. 4 Test statistics (U*) for Pettitt test over the gauge stations (* denotes U* is statistically significant at 5% significance level, P-value is less than 0.05).

Stations	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adama	232	291	129	139	227	254	279	157	204	124	267	171
Addis Ababa	173	226	160	140	120	162	176	144	246	94	180	198
Dire Dawa	206	292	263	257	176	218	*337	169	166	246	197	250
DebreZeit	179	99	119	151	186	196	202	266	258	124	139	111
Melkasa	80	117	90	117	80	80	106	101	*176	86	151	59
Metahara	57	*136	100	*138	84	50	46	70	98	66	144	52
Majete	91	*114	70	44	58	54	108	62	70	46	98	54
MehalMeda	99	171	84	134	59	127	206	72	76	88	125	73
Mezezo	65	102	84	58	70	80	82	48	52	68	103	54
Wonji	139	178	255	204	144	184	141	208	196	160	249	125

Table 5. Test statistics (RVN) for Von Neumann ratio test over the gauge stations (* denotes RVN is statistically significant at 5% significance level, P-value is less than 0.05).

Stations	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adama	1.7	1.8	2.1	2.0	2.3	2.1	1.7	1.8	2.3	1.9	*1.5	2.1
Addis Ababa	1.8	2.1	1.9	2.3	2.4	2.1	1.9	1.9	2.3	1.6	1.9	2.0
Dire Dawa	2.2	1.8	1.6	*1.5	1.7	1.9	1.9	1.9	1.9	2.0	2.2	2.0
DebreZeit	1.6	2.6	1.8	1.8	1.9	1.8	1.6	1.7	2.3	1.8	1.7	2.0
Melkasa	1.8	2.2	2.0	1.6	2.0	2.2	2.1	2.2	1.8	*1.4	1.9	1.8
Metahara	2.0	1.7	1.9	1.7	2.1	2.8	2.5	2.0	2.0	1.5	1.8	2.0
Majete	*1.4	1.5	1.9	2.6	2.5	*1.3	*1.3	*1.4	2.0	1.7	1.9	2.1
MehalMeda	1.8	1.6	2.0	2.0	2.5	2.3	1.5	2.1	2.3	1.6	1.9	2.3
Mezezo	1.9	*1.3	2.1	2.4	1.9	*1.2	1.2	1.5	1.9	1.9	1.8	2.4
Wonji	1.7	2.0	1.7	*1.5	1.7	2.0	2.1	2.2	1.8	1.9	2.0	1.6

The average correlation found by this study is higher than that of our study. In this study, data from large number of stations across central Ethiopian highlands are included. It is not clearly indicated in the study as to which locationsin the Awash River basin are also included. In addition, the data periods are quite different. These conditions might have caused difference in the findings. Yet, the finding by Dinku *et al.* (2008) is in agreement with that of our study in that there strong correlation between observed rainfall data and corresponding data from the CRU and the GPCC datasets and the correlation coefficient for the GPCC is higher than that of the CRU. For locations in northcentral Ethiopia, not in Awash river basin, Asfaw *et al.* (2018) found that data from GPCC are significantly and strongly correlated with station data ($r = 0.72$, $p < .001$) whereas they found that data from CRU dataset are not significantly correlated with observed data ($r = 0.27$, $p > 0.1$). The correlation coefficients for both GPCC and CRU of this study are lower than that ofcorresponding values in our study. Asfaw *et al.* (2018) studied locations in different basin, this among other factors clearly affect the possibility of getting comparable results. Yet, the findingis partly agreement with that of our study in that there strong correlation between observed rainfall data and corresponding GPCC.

Root mean square error

The root mean square error (RMSE) estimated are shown in Table 7 and Table 8. For the CRU dataset, the mean monthly RMSE over the stations varies from 21.3mm in November to 90 mm in July. For the GPCC dataset, the mean monthly RMSE over the stations varies from 20.6mm in December to 77.2 mm in August. The mean and median monthly RMSE for the CRU dataset are higher than that of the GPCC dataset for all months except for November (Tab. 7 and 8, Bar plot in Fig. 5, and Box plot in Fig. 6).

The annual RMSE over the stations for CRU dataset varies from 515.8mm at Dire Dawa to 1247.3 mm at Mezezo. The annual RMSE over the stations for GPCC datasets varies from 248.8mm at Addis Ababa to 1192.9 mm at Mezezo. The annual RMSE for the CRU dataset are higher than that of the GPCC dataset for all station except at Addis Ababa (Table 7 and 8). The annual mean RMSE for CRU dataset 728mm is far higher than annual mean RMSE for GPCCC dataset 541.57mm. Clearly, the monthly and annual RMSE values are generally higher for the CRU dataset than that of GPCC dataset. The GPCC dataset is more close to the observation data than that of the CRU dataset.

Looking at Figure 5 and 6, it can be noted that the root mean square errors for both CRU and GPCC datasets are highest for summer months (June to September, the main rainy season) followed by the spring months (February to May, the small rain season). The errors are the least for the winter months (October to January, the dry season). Studied show that high rainfall variability and high rate of rainfall in a season reduces the ability of rainfall product to predict the rainfall (Haile *et al*. 2010, Mekonnen *et al*. 2021, Asfaw *et al.* 2023, Li & Shao 2025). The RMSE can be normally higher for rainy season either due to high rainfall variability and/or more frequent heavy rainfall events in rainy season, which are typically more difficult to predict accurately.

In Ethiopia and in Awash River basin the variability of rainfall is generally higher in the dry season than in rainy season. Thus, the most likely reason for higher RMSE in rainy season than the dry season would be related to more frequent heavy rainfall events in rainy season. The monthly and annual Root mean square error (RMSE) is exceptionally highest at Mezezo which is also the gauge location with highest missing data. The lowest annual and monthly RMSE are found at Dire Dawa and Addis Ababa and where the missing data are also the lowest (Tab. 1). This implies the errors are highly associated with data quality at the weather stations. This indicates the performance of derived gauge data (GPCC and CRU) is markedly being influenced by the interpolation techniques employed to replace missing data during development of the gridded datasets.

None of the two related local studies, by Dinku *et al.* (2008) and by Asfaw *et al.* (2018), which compared GPCC and CRU, used mean root mean square for comparison. However, Dinku *et al.* (2008) used mean error (ME), and mean absolute error (MAE). They found that GPCC-full showed lower error than the CRU. Similar to the local studies, a study in Pakistan by Ahmed *et al.* (2019) did not use mean root mean square for comparison. Based on the Mean Bias Errors (MBE) and Mean Absolute Error (MAE), they found mixed results for different regions and months, but overall the GPCC dataset showed lowest error and bias than the CRU dataset. The results of the previous studiesare generally in agreement with the results of this study.

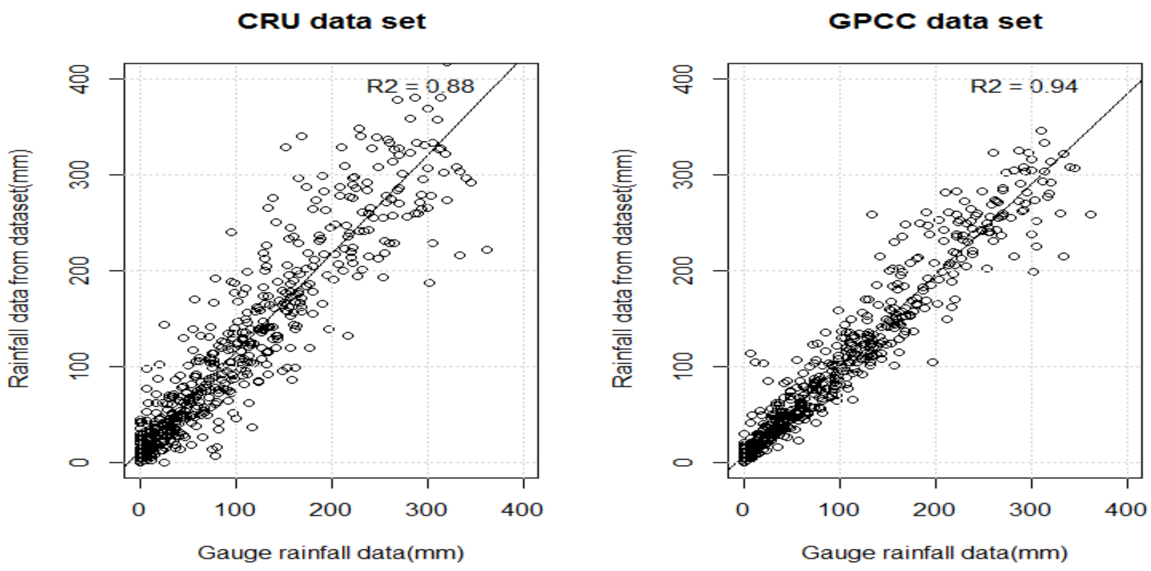

Figure 3. The scatter plots of monthly rainfall, observed versus gridded datasets at Addis Ababa for the period 1960-2015.

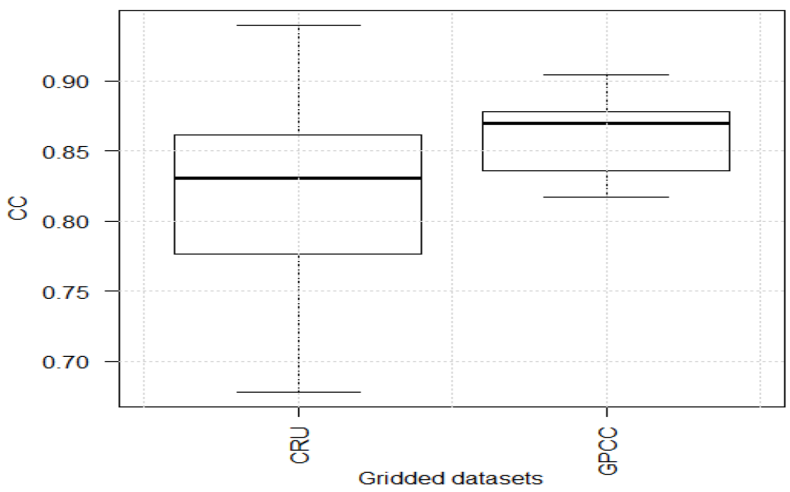

Figure 4. Comparison of the box plots for Pearson correlation coefficient (CC) for the CRU and GPCC datasets.

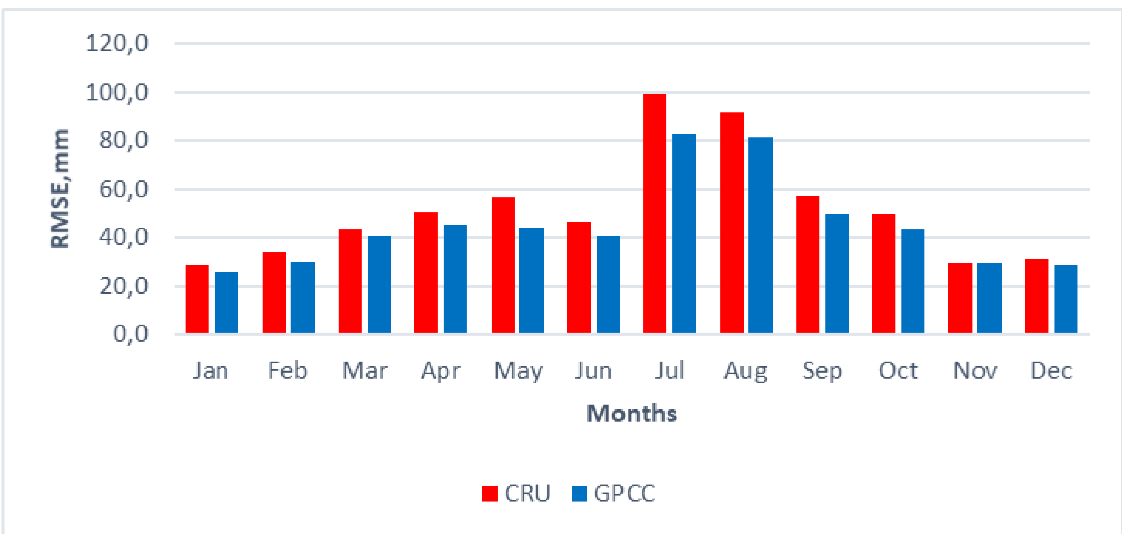

Figure 5. Bar plot for spatial mean of the root mean square error (RMSE) across elected stations for the CRU and GPCC datasets.

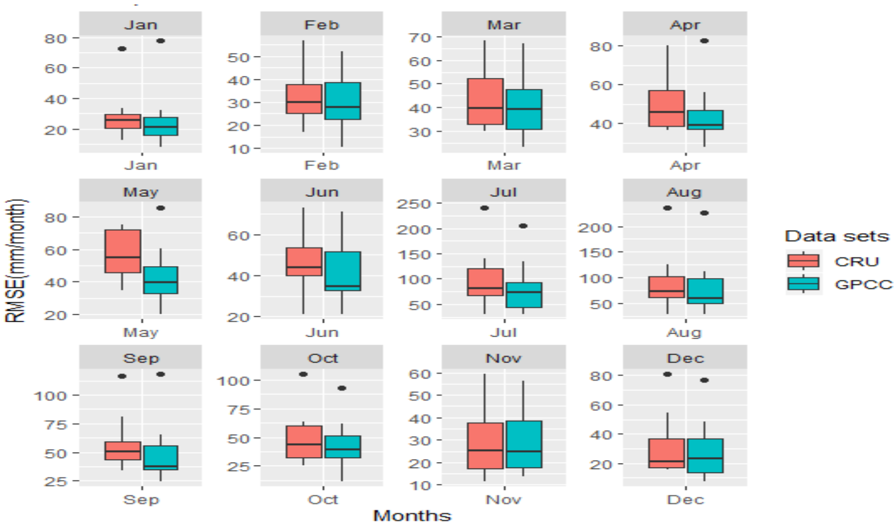

Figure 6. Box plots for monthly root mean square error (RMSE) across selected stations for the CRU and GPCC datasets.

Table 6. Pearson correlation coefficient (CC) between gauge and grid data.

Stations	CRU dataset	p-value	GPCC dataset	p-value
Adama/Nazert	0.68	0.00E+00	0.63	0.00E+00
Addis Ababa	0.94	1.70E-170	0.97	2.11E-187
DebreZeit	0.86	2.99E-204	0.84	2.01E-194
Dire Dawa	0.86	7.89E-179	0.87	1.11E-97
Majete	0.86	4.06E-141	0.87	3.03E-116
MehalMeda	0.77	1.12E-115	0.90	2.52E-77
Melkasa	0.84	7.45E-86	0.88	9.30E-79
Metahara	0.78	7.79E-48	0.87	2.65E-58
Mezezo	0.80	6.59E-101	0.82	4.27E-96
Wonji	0.82	2.99E-204	0.87	1.34E-161
Mean	0.82		0.85	

Note: The p-Value for both datasets and all stations are by far less than 0.05 and 0.01, indicating there is 99% confidencethat calculated correlation values are not because of random errors.

Table 7. Root mean square error (mm/month or year) for CRU datasets against gauge data.

Stations	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Adama	33.3	51.8	55.0	58.9	65.0	60.4	129.2	104.8	58.2	60.7	30.5	20.2	728
Addis Ababa	12.6	16.9	30.1	38.0	34.8	39.8	69.0	60.3	45.9	25.6	11.7	16.5	401.2
Dire Dawa	17.5	29.1	35.4	46.8	45.1	21.0	30.1	28.8	34.9	35.2	17.9	17.9	359.7
DebreZeit	19.2	37.8	52.4	38.6	47.8	42.9	68.0	67.2	48.9	43.2	27.7	22.1	515.8
Melkasa	23.1	22.3	31.0	36.6	49.8	39.2	67.9	53.6	33.7	31.2	23.1	27.2	438.7
Metahara	23.2	24.0	32.0	51.2	75.1	41.2	58.4	79.1	52.7	43.4	13.7	15.4	509.4
Majete	29.2	29.6	40.2	44.6	60.2	73.1	95.2	126.4	60.1	63.7	50.6	54.3	727.2
MehalMeda	29.9	29.9	51.4	65.9	74.2	45.0	140.6	91.5	81.0	58.3	39.9	40.3	747.9
Mezezo	72.5	56.7	68.2	80.2	74.4	56.5	240.8	236.3	116.8	105.0	59.2	80.7	1247.3
Wonji	28.7	38.0	39.3	39.8	40.2	44.4	92.8	65.3	42.3	27.8	16.8	16.1	491.5
Mean	28.9	33.6	43.5	50.0	56.7	46.4	99.2	91.3	57.5	49.4	29.1	31.1	728

Table 8. Root mean square error (mm/month or year) for GPCC datasets against gauge data.

Months	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Adama	32.1	52.1	60.9	56.1	60.5	55.5	134.9	100.5	53.7	61.5	32.3	22.4	722.5
Addis Ababa	7.8	13.2	23.1	27.7	20.2	25.6	40.7	32.3	24.7	11.7	13.9	7.9	248.8
Dire Dawa	17.5	29.1	35.4	46.8	45.1	21.0	30.1	28.8	34.9	35.2	17.9	17.9	359.7
DebreZeit	17.5	38.8	49.7	39.4	41.5	40.7	68.3	65.0	37.1	40.7	24.1	24.1	486.9
Melkasa	12.1	26.2	29.2	36.9	36.0	33.6	77.9	50.9	33.0	31.7	25.4	25.3	418.2
Metahara	15.3	10.3	24.2	36.8	29.5	32.5	39.4	51.5	35.8	38.5	17.9	11.4	343.1
Majete	28.2	21.5	42.0	46.8	51.0	71.1	53.2	111.5	64.9	53.5	48.3	48.2	640.2
MehalMeda	25.1	25.5	38.2	38.6	37.6	32.6	98.2	92.3	56.9	44.6	40.8	41.0	571.4
Mezezo	77.7	46.4	66.9	82.7	85.5	58.4	204.8	226.6	118.1	92.7	56.2	76.9	1192.9
Wonji	24.4	38.2	40.0	38.6	31.7	35.9	78.6	54.6	37.0	23.1	17.4	12.5	432
Mean	25.8	30.1	41.0	45.0	43.8	40.7	82.6	81.4	49.6	43.3	29.4	28.8	541.57

The Kolmogorov-Smirnov test

The similarity in probability distributions of the observed data and data from the two gridded datasets were tested by the Kolmogorov-Smirnov (KS) test. The results of the KS test (Tab. 9 and 10) releveled that among 120 monthly time series from each of the dataset (CRU and GPCC), only 28.8% of the time series for the CRU and 33.3% for the GPCC dataset showed similar probability distribution with gauge data ($\alpha = 0.05$). The remaining monthly time series (71.2% for the CRU and 66% for the GPCC) for the two datasets failed ($\alpha = 0.05$) to show similar probability distribution with respective gauge data.

Looking at the previous studies in Ethiopia, Dinku *et al.* (2008) did not compare the probability distribution of gridded dataset with the observation data. While Asfaw *et al.* (2018) compared GPCC and CRU probability distribution of data from the datasets and the observation station using Kolmogorov-Smirnov test. Asfaw *et al.* (2018) found that the data from GPCC dataset followed probability distribution of an average gauge data but that of CRU dataset failed to follow probability distribution of an average of gauge data over the locations in the northern central Ethiopia. The studies have not shown the results for each month. Similar study in Pakistan by Ahmed *et al.*(2019) found different results for the CRU and GPCC across climatic regions–semi-arid, arid, and hyper arid areas. They found that GPCC dataset replicated similar probability distributions with gauge data for 9, 7, and 9 months for semi-arid, arid, and hyper arid areas, respectively. Whereas the CRU dataset replicated similar probability distributions with gauge data for 6, 10, and 0 months for semi-arid, arid, and hyper arid areas, respectively.In semi-arid and hyper region, the GPCC dataset showed similarity with observed data for more number of months than that of the CRU dataset. While in arid region, the CRU dataset showed similarity with observed data for more number of months than that of the GPCC dataset. They found similar distributions between the datasets and gauge data in most of the months (for 25 out of 36 months for GPCC and 16 out 36 months for the CRU). This is in contradiction to the result of this study. Overall, the GPCC showed similarity with observed data for higher number of months than that of the CRU. This is in agreement with the result of this study.

The difference in the results of this study and previous studies are potentially attributed to the difference in density of weather stations and the quality of data from observation stations in the study areas. The difference in the versions of datasets and the method of analysis used in the studies would also have influence on the results.

The failure in the CRU and GPCC datasets in this study to replicate similar probability distribution with observed data for large majority of monthly time series could be attributed to inherent uncertainties in the gridded or observation data. On one hand, there are inherent uncertainties in the gauge based gridded datasets mainly associated with the density of gauge stations, quality of gauge data, and the interpolation techniques used during construction of the datasets. As a result, the capability of gridded data to replicate spatial and temporal climate variability might be limited (Nashwan *et al.* 2019, Tozer *et al.* 2012). However, the uncertainties associated with random and systemic errors in gridded datasets are often fairly low (Dinku *et al*. 2008). On the other hand, the quality of observation data used would affect the KS test.

In summary, the Pearson correlation coefficients, root mean square errors, and KS test ($\alpha = 0.1$) in this study showed that the GPCC dataset has better performance than the CRU dataset. Taylor diagram (Fig. 7) also illustrates that GPCC dataset performs better than CRU dataset. With respect to the similarity of the probability distribution of the dataset with observed that GPCC dataset showed similar probability distribution with observed data for more number of time series than that of the CRU dataset. In agreement with that, eastern African regional study by Dinku *et al.* (2008) reported that the GPCC showed best statistics best overall statistics over the CRU dataset. A number of other studies in different regions across the worldalso reported that the GPCC dataset performs better than the CRU dataset (Nashwan *et al*. 2019, Ahmed *et al*. 2019, Hu *et al.* 2018, Faiz *et al*. 2018). The relative higher performance of the GPCC dataset might have be owning to the fact that the GPCC dataset is derived based on ground observation data from much larger number of the weather stations (85,000) across the world compared to that of the CRU (4000). Furthermore, the highest RMSE for location with large missing data and the lowest RMSE at locations with lowest missing data indicate how significantly the data quality from observation station affect the performance of the derived climate products.

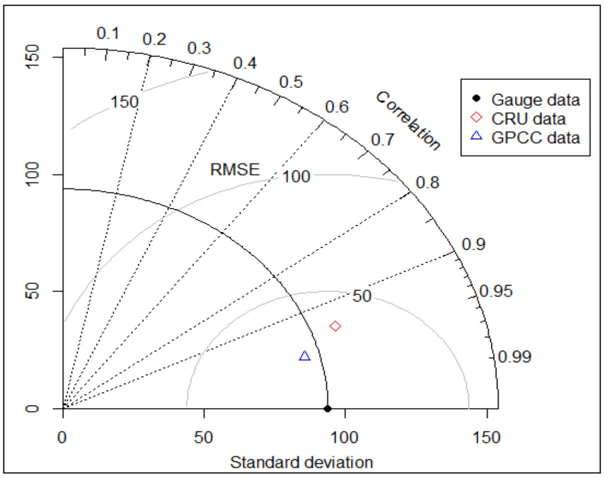

Figure 7. Taylor Diagram to compare the two datasets against observed data based on Centered Root Mean Square Error (RMSE), Correlation Coefficient, and Standard Deviation at Addis Ababa for the entire monthly data (1960-2015).

Table 9. Kolmogorov-Smirnov (KS) test statistics (D) for monthly time series from gauge and CRU dataset.

Months	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adama	*0.25	*0.25	0.22	**0.27	**0.26	**0.27	*0.23	*0.23	**0.27	**0.27	**0.27	**0.38
Addis Ababa	**0.29	0.19	0.14	0.08	0.16	**0.31	**0.31	**0.28	**0.29	**0.26	**0.44	**0.48
Dire Dawa	**0.39	0.20	0.13	0.09	*0.24	**0.34	**0.26	*0.25	*0.25	**0.29	**0.32	**0.41
DebreZeit	**0.40	**0.36	*0.25	*0.24	**0.30	*0.26	0.22	0.20	**0.27	**0.32	**0.38	**0.54
Melkasa	*0.27	0.24	0.21	**0.35	0.25	0.20	0.20	0.26	0.22	0.23	**0.32	0.23
Metahara	**0.42	**0.35	0.20	**0.48	**0.49	**0.58	**0.37	**0.57	**0.68	**0.43	**0.31	**0.37
Majete	*0.29	0.25	0.22	**0.33	0.22	0.24	**0.40	**0.50	**0.33	0.21	0.23	*0.31
MehalMeda	**0.32	0.21	0.21	**0.45	**0.45	**0.38	**0.28	0.16	**0.68	**0.42	**0.45	**0.39
Mezezo	**0.47	**0.33	**0.38	**0.50	0.16	**0.36	**0.80	**0.84	**0.64	**0.40	**0.32	**0.46
Wonji	**0.32	**0.40	0.16	**0.32	0.14	**0.32	**0.26	0.25	0.25	**0.26	**0.26	**0.37

Key: * and ** denote D is statistically significant at 10 % and 5% significance level, respectively.

Table 10. Kolmogorov-Smirnov (KS) test statistics (D) for monthly time series from gauge and GPCC dataset.

Months	Months											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adama	**0.41	**0.34	*0.23	**0.32	**0.28	**0.34	**0.31	**0.32	*0.23	**0.39	**0.43	**0.55
Addis Ababa	**0.36	**0.27	0.12	0.12	0.13	0.18	0.21	0.13	0.11	0.19	**0.51	**0.55
Dire Dawa	**0.39	0.20	0.13	0.09	0.24	**0.34	**0.26	*0.25	*0.25	**0.29	**0.32	**0.41
DebreZeit	**0.51	**0.41	*0.25	*0.24	0.18	0.15	0.17	0.18	0.20	0.27	**0.47	**0.60
Melkasa	**0.41	**0.33	0.23	**0.34	**0.33	**0.31	0.25	0.19	**0.31	**0.30	**0.47	**0.38
Metahara	**0.49	**0.42	0.25	**0.52	**0.40	**0.45	*0.30	**0.46	**0.54	**0.51	**0.39	**0.51
Majete	0.26	**0.32	0.18	*0.31	0.24	0.14	0.20	**0.36	**0.40	0.24	0.27	**0.38
MehalMeda	**0.30	0.20	0.12	0.23	**0.29	0.25	0.19	**0.30	**0.41	**0.30	**0.46	**0.42
Mezezo	**0.52	0.27	**0.34	**0.42	0.18	**0.47	**0.68	**0.70	**0.59	**0.36	0.22	**0.41
Wonji	**0.44	**0.47	0.25	**0.33	0.18	**0.33	0.16	**0.30	**0.32	**0.28	**0.39	**0.53

Key: * and ** denote D is statistically significant at 10 % and 5% significance level, respectively.

CONCLUSIONS

The monthly rainfall time series data from selected weather stations are dominantly (94%) homogenous across the selected stations at a significance level α=0.05. Rainfall data extracted from both the CRU and the GPCC datasets were significantly, and highly, correlated with the corresponding observation data from all weather stations. The GPCC dataset showed generally higher correlations with gauge data (CC=0.85 for monthly time series) and lower errors (with averaged RMSE=45 mm/month) than that of the CRU dataset (CC=0.82 for monthly time series, with averaged RMSE over location =51 mm/mm). Yet, majority of the monthly rainfall data from both the CRU and GPCC datasets (71.2% for the CRU and 66% for the GPCC) failed to follow probability distribution with observation data.Still, it is clear that the GPCC dataset showed similar probability distribution with observed data for more number of time series than that of the CRU dataset. All graphical analysis also showed that GPCC dataset aligns more closely with gauge data than of the CRU. Overall, the GPCC dataset has showed better performance than the CRU dataset to simulate rainfall for the Awash River Basin. Thus, the GPCC dataset can be used as better alternative source of rainfall data for hydrological analysis and modelling required in the planning and design of water infrastructure, management of water resources, and climate and hydrological studies in the basin, especially for ungauged and data-scarce areas of the river basin.Further studies are crucial to identify datasets that can perform better across locations and seasons in reducing errors and bias and in replicating the probability distribution of observation data.

ACKNOWLEDGEMENTS

The Ethiopian metrological agency, and the data sources for GPCC and CRU are acknowledged for providing data freely. The authors would like to express their gratitude to the editor of «Bulletin de l'Institut Scientifique» ACHAB Mohammed and the journal reviewer's, prof. Brahim BENZOUGAGH (Sc. Ins. Rabat) and Dr. Wahib HAMMOUDY (Dir. Méteo. Nat. Casablanca) for their significant contributions to the improvement of the manuscript.

REFERENCES

Adane G. B., Hirpa B. A., Song C. *et al.* 2020. Rainfall Characterization and Trend Analysis of Wet Spell Length across Varied Landscapes of the Upper Awash River Basin, Ethiopia. *Sustainability* ,12 ,21, 9221p. https://doi.org/10.3390/su12219221.

Adeba D., Kansal M. L. & Sen S. 2015. Assessment of Water Scarcity and Its Impacts on Sustainable Development in Awash Basin, Ethiopia. *Sustainable Water Resources Management,*1 , 1, 71–87. https://doi.org/10.1007/s40899-015-0006-7.

Ahmed K., Shahid S., Wang X. *et al.* 2019. Evaluation of griddedprecipitationdatasets over aridregions of Pakistan. *Water*, *11*, 2, 210p. https://doi.org/10.3390/w11020210.

AL-Lami A. M., AL-Timim Y. K., & AL-SalihiA. M. 2014. The homogeneityanalysis of rainfall time series for selectedmeteorological stations in Iraq. *Diyala Journal for Pure Science*, *10*, 2, 60-77.

Asfaw A., Simane B., Hassen A. *et al.* 2018. Variability and time series trend analysis of rainfall and temperature in northcentralEthiopia: A case study in Wolekasub-basin. *Weather and climateextremes*, *19*, 29-41. https://doi.org/10.1016/j.wace.2017.12.002.

Asfaw W., Rientjes T. & Haile A. T. 2023. Blending high-resolution satellite rainfallestimates over urbancatchmentusingBayesian Model Averagingapproach. *Journal of Hydrology:RegionalStudies*, *45*, 101287p.

Awulachew S. B., Yilma A. D., Loulseged M. *et al.* 2007. *Water resources and irrigation development in Ethiopia (Vol. 123)*. International Water Management Institute.

Chang C. K., Ghan A., PuayH. T. et al. 2017. Homogeneitytesting and trends analysis in long termrainfall data for Sungai Pahang River Basin over 40 years records. In *Proceedings of the 37th IAHR World Congress, Kuala Lumpur, Malaysia*, 13-18.

Daba M. H., Ayele G. T. & You S. 2020. Long-TermHomogeneity and Trends of Hydroclimatic Variables in UpperAwash River Basin, Ethiopia. *Advances in Meteorology*, *2020*(1), 8861959. https://doi.org/10.1155/2020/8861959.

Dinku T., Connor S. J., Ceccato, P. N. *et al.* 2008. Comparison of global griddedprecipitationproducts over a mountainousregion of Africa. *International Journal of Climatology,* 28, 12, 1627–38. https://doi.org/10.1002/joc.1669.

Dinku T., Hailemariam K., Maidment R. *et al.* 2013. Combined use of satellite estimates and rain gauge observations to generate high-qualityhistoricalrainfall time series over Ethiopia. *International Journal of Climatology*, *34*, 7, 2489-2504. https://doi.org/10.1002/joc.3855.

Salih E., Daniel T. & Amanuel Z. 2021. Assessment of Spatio-Temporal Relationship between Gauge and Satellite Rainfall Estimates over Awash River Basin, Ethiopia. *Journal of Remote Sensing & GIS*, 10, 7, 1-8.

Elzeiny R., Khadr M., Zahran, S. *et al.* 2019. Homogeneityanalysis of rainfallseries in the upperbluenile river basin, Ethiopia. *Journal of Engineering Research*, *3,* 46-53.

Ethiopian Meteorological Institute. 2025. "About." EMI. https://www.ethiomet.gov.et/organisation/about/.

Faiz M. A., Liu D., Fu Q. *et al.* 2018. How accurate are the performances of griddedprecipitation data products over NortheastChina?. *AtmosphericResearch*, *211*, 12-20. https://doi.org/10.1016/j.atmosres.2018.05.006.

Gebrechorkos S. H., Hülsmann S. &Bernhofer C. 2018. Evaluation of multiple climate data sources for managingenvironmentalresources in East Africa. *Hydrology and Earth System Sciences*, *22*, 8, 4547-4564. https://doi.org/10.5194/hess-22-4547-2018.

Getahun Y. S., Li M. H. &Pun I. F. 2021. Trend and change-point detection analyses of rainfall and temperature over the Awash River basin of Ethiopia. *Heliyon*, *7*(9) : e08024. https://doi.org/10.1016/j.heliyon.2021.e08024.

Haile A. T., Rientjes T., GieskeA. *et al.* 2010. Multispectral remotesensing for rainfalldetection and estimation at the source of the Blue Nile River. *International Journal of AppliedEarth Observation and Geoinformation*, *12*, S76-S82. https://doi.org/10.1016/j.jag.2009.09.001.

Hu Z., Zhou Q., Chen X. *et al.* 2018. Evaluation of three global griddedprecipitation datasets in central Asia based on rain gauge observations. *International Journal of Climatology*, *38*, 9, 3475-3493. https://doi.org/10.1002/joc.5510.

Javari M. 2016. Trend and homogeneityanalysis of precipitation in Iran. *Climate*, *4*, 3, 44. https://doi.org/10.3390/cli4030044.

Jones P. D. 2017. CRU TS4.01: Climatic Research Unit (CRU) Time-Series (TS) Version 4.01 of High-Resolution Gridded Data of Month-by-Month Variation in Climate (Jan. 1901- Dec. 2016). In *Centre for Environmental Data Analysis (CEDA)* *and University Of East Anglia Climatic Research Unit (CRU)*. https://doi.org/10.5285/58a8802721c94c66ae45c3baa4d814d0.

Kang H. M. &Yusof F. 2012. Homogeneity tests on dailyrainfallseries. *International Journal of Contemporary Mathematical Sciences*, *7*, 1, 9-22.

Kerim T., Abebe A., &Hussen B. 2016. Study of water allocation for existing and future demandsunderchangingclimatecondition: case of UpperAwashSub River Basin. *Journal of Environment and Earth Science*, *6*, 10, 18-31.

LiM. &Sha Q. 2025. Enhancinggriddedclimateproductswiththird party weather data in a rainfallstudyfrom Western Australia. *Scientific Reports*, *15*, 1, 36522. https://doi.org/10.1038/s41598-025-11145-0.

Mekonnen K., Melesse A. M. &Woldesenbet T. A. 2021. Spatial evaluation of satellite-retrievedextremerainfall rates in the UpperAwash River Basin, Ethiopia. *AtmosphericResearch*, *249*, 105297. https://doi.org/10.1016/j.atmosres.2020.105297.

Merino A., García-Ortega E., Navarro A. et al. 2021. Evaluation of griddedrain-gauge-basedprecipitationdatasets: Impact of station density, spatial resolution, altitude gradient and climate. *International Journal of Climatology*, *41*, 5, 3027-3043. https://doi.org/10.1002/joc.7003.

Moritz S. 2019. CRAN - Package imputeTS. *The R Journal* ,1 , 9, 207–18. https://doi.org/10.32614/RJ-2017-009.

Mulugeta S., Fedler C. &Ayana M. 2019. Analysis of long-term trends of annual and seasonalrainfall in the Awash River Basin, Ethiopia. *Water*, *11*, 7, 1498. https://doi.org/10.3390/w11071498.

Nashwan M. S., Shahid S. & Wang X. 2019. Uncertainty in estimated trends usinggriddedrainfalldata: a case study of Bangladesh. *Water*, *11*, 2, 349. https://doi.org/10.3390/w11020349.

National Meteorological Agency. 2020. RevisedMeteorological Station Network Master Plan for 2021-2030.

Navidi Nassaj B., Zohrabi N., Shahbaz A. N. *et al.* 2022. Evaluating the performance of eight global griddedprecipitationdatasetsacross Iran. *Dynamics of Atmospheres and Oceans*, *98*, 101297. https://doi.org/10.1016/j.dynatmoce.2022.101297.

Pebesma E. &Graeler B. 2020. *Gstat: Spatial and Spatio-Temporal Geostatistical Modelling, Prediction and Simulation*. Version 2.0-6. https://CRAN.R-project.org/package=gstat.

Pohlert T. 2018. *Trend: Non-Parametric Trend Tests and Change-Point Detection*. The R Project for Statistical Computing. https://CRAN.R-project.org/package=trend.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Schneider U., Becker A., Finger P. *et al.* 2015. "GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges Built on GTS-Based and Historic Data: Gridded Monthly Totals." Global Precipitation Climatology Centre (GPCC) at Deutscher Wetterdienst. https://doi.org/10.5676/dwd_gpcc/fd_m_v7_050.

Sun Q., Miao C., Duan Q. *et al.* 2018. A review of global precipitation datasets: Data sources, estimation, and intercomparisons. *Reviews of geophysics*, *56*, 1, 79-107. https://doi.org/10.1002/2017rg000574.

Tozer C. R., Kiem A. S. & Verdon-Kidd D. C. 2012. On the uncertaintiesassociatedwithusinggriddedrainfall data as a proxy for observed. *Hydrology and Earth System Sciences*, *16*, 5, 1481-1499. https://doi.org/10.5194/hess-16-1481-2012.

Tsidu G. M. 2012. High-resolutionmonthlyrainfalldatabase for Ethiopia:Homogenization, reconstruction, and gridding. *Journal of Climate*, *25*, 24, 8422-8443. https://doi.org/10.1175/JCLI-D-12-00027.1.

Uddameri V. 2017. Using R to Read NetCDF Files. WateR - Tutorials: NetCDF Available Online: Osf.Io/8xche. http://dx.doi.org/10.17605/OSF.IO/8XCHE.

Van Vooren S., Van Schaeybroeck B., Nyssen J. *et al.* 2019. Evaluation of CORDEX rainfall in northwestEthiopia:sensitivity to the model representation of the orography. *International Journal of Climatology*, *39*, 5, 2569-2586. https://doi.org/10.1002/joc.5971.

Wagesho N., Goe N. K., & Jain M. K. 2013. Temporal and spatial variability of annual and seasonalrainfall over Ethiopia. *Hydrological Sciences Journal*, *58*, 2, 354-373. https://doi.org/10.1080/02626667.2012.754543.

Wijngaard J. B., Klein Tank A. M. G. &KönnenG. P. 2003. Homogeneity of 20th century Europeandailytemperature and precipitationseries. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *23*, 6, 679-692. https://doi.org/10.1002/joc.906.

Tesfaye T. W., Dhanya C. T. &Gosain A. K. 2017. Evaluation of ERA-Interim, MERRA, NCEP-DOE R2 and CFSR Reanalysisprecipitation Data using Gauge Observation over Ethiopia for a period of 33 years. *AIMS Environmental Science*, *4*, 4, 596–620. https://doi.org/10.3934/environsci.2017.4.596.

Zwiers F.W & Zhang X. 2009. *Guidelines on Analysis of Extremes in a Changing Climate in Support of Informed Decisions for Adaptation*. World Meteorological Organization (WMO). https://library.wmo.int/index.php?lvl=notice_display&id=138#.XTguo_IzbIU.

Manuscrit reçu le 04/03/2025
Version révisée acceptée le 08/12/2025
Version finale reçue le 25/12/2025
Mise en ligne le 31/12/2025