OPEN ACCESS Freely available online

Journal of Remote Sensing & GIS

Research Article

Assessment of Spatio-temporal Relationship between Gauge and Satellite Rainfall Estimates over Awash River Basin, Ethiopia

Salih Edris[1,2*], Daniel Teka[3], Amanuel Zenebe[1,3]

[1] *Institute of Climate and Society, Endayesus Campus, Mekelle University, Ethiopia*

[2] *College of Natural and Computational Science, Samara University, Ethiopia*

[3] *Department of Land Resource Management and Environmental Protection, Endayesus*

ABSTRACT

Accurate information on rainfall data is essential for numerous operational and research fields. Conventionally, ground-based measurement is the main source of rainfall data. However, in developing countries, networks of ground-based measurements are very sparse or nonexistent. An alternative to this measurement could be satellite-based rainfall estimates (SREs). However, SREs need to be validated as their accuracy can be affected by topography and climate. This study seeks to investigate the spatiotemporal relationship between gauge and SREs over the Awash River basin. The Climate Hazards Group Infrared Precipitation (CHIRP), CHIRP combined with station observations (CHIRPS), and African Rainfall Climatology version 2 (ARC2) are evaluated at dekadal (10-day), monthly and annual time-scales for selected normal years against 37 ground-based measurements located at different elevations of the basin. A point-to-grid-based comparison is adopted, using continuous statistical validation tools. Temporal and spatial analysis indicates the basin exhibits tremendous spatial variability in the rainfall amount which varying from 190 mm in lowland to 1300 mm yr-1 in highland, with significant correlation. From the overall analysis at dekadal, monthly, and annual temporal scale, CHIRPS followed by CHIRP exhibited better performance in comparison to ARC2. ARC2 product is poorly performed SREs with underestimations of high rainfall rate. The agreement between the SREs and ground-based measurement improved with increase in time scale from dekadal (for instance CHIRPS has correlation > 0.77, Nash-Sutcliff efficient coefficient (Eff) > 0.59, root mean square error (RMS) < 22.1, and bias ≤ 1.1) to monthly (correlation > 0.89, Eff > 0.79, RMSE < 39.0 and bias ≤ 1.0), but the performance of these products decrease when aggregated into annual time scale (correlation > 0.40, Eff > -0.56, RMSE < 161.10). In general, the SREs shows good agreement with ground-based measurements over the highland parts of the basin at dekadal and monthly time scale, however, at an annual time scale, all products show better performance over the lowland parts of the basin. This study demonstrates the reliability of satellite rainfall estimates in regions where the spatial network of the gauge is highly sparse and inaccessible.

Keywords: Ground-based Measurement; Awash Basin; Satallite; Rainfall Estimation

INTRODUCTION

Accurate information on rainfall is essential for numerous operational and research fields of water management, hydrological applications, and agricultural forecasts (Sunilkumar et al., 2015). Conventionally, the ground-based measurement is the main source of rainfall data. However, in many parts of the world and especially in developing countries, ground-based measurements are very sparse or nonexistent (Behrangi et al., 2011) and hence supplementing rainfall estimates from satellite-

Correspondence to: Salih Edris, Institute of Climate and Society, Endayesus Campus, Mekelle University, Ethiopia, Tel: 251914143929; E-mail: salihedris99@gmail.com

Received: November 16, 2020; **Accepted:** September 01, 2021; **Published:** September 10, 2021

Citation: Edris S (2021) Assessment of Spatio-temporal Relationship Between Gauge and Satellite Rainfall Estimates over Awash River Basin, Ethiopia. J Remote Sens GIS 10: p151

based sensors is important. However, these satellite estimates need local validation.

The number of ground-based rainfall measurement in Ethiopia is still less than the minimum station density recommended by the world meteorological Organization (WMO, 1996), as a result, the data provided by National Meteorological Agency (NMA) of Ethiopia pause many inadequacies like many National Meteorological Stations in Africa (Dinku et al., 2014). For instance, ground-based rainfall measurement in the Awash River basin, especially in the lowland parts of the basin, is extremely sparse with no stations over some parts of the basin. In addition, the existence of ground-based measurements is located along main roads in a town and cities. Because of the unrepresentative distribution of ground-based measurement, the dependability on ground-based measurement to estimate the spatial distribution of rainfall over large areas of Ethiopia is considerably reduced (Ayehu et al., 2018).

Advances in remote sensing science have provided an opportunity to estimate rainfall from satellite observations and are becoming an important alternative source of rainfall data (Ayehu et al., 2018). Satellite rainfall estimates (SRE) are advantageous in terms of temporal and spatial coverage, as well as they can provide data in ungagged basins (Katsanos et al., 2015). At present, there are now relatively new satellite-based rainfall products with good spatial (0.05° latitude/longitude) and temporal (daily, pentad, dekadal, and monthly) resolution, as well as quasi-global coverage (50° S–50° N) for a longer duration. These are the Climate Hazards Group (CHG) Infrared Precipitation (CHIRP) and CHIRP combined with station data (CHIRPS) from the University of California at Santa Barbara and U.S. Geological Survey (Funk et al., 2014; 2015a). ARC2 is also another new product with relatively high spatial (0.1° and 0.0375°, respectively) and temporal (daily) resolutions. Despite its importance as an alternative source of rainfall data, satellite rainfall products need to be validated as their accuracy can be affected by geographical position, topography, and climate, as well as by the algorithms used to derive rainfall from satellite measurements (Meng et al., 2014; Xue et al., 2013).

Assessment and inert-comparison of different rainfall products over complex topography like Awash River basin are essential to determine which products are representative. A number of studies have been conducted over different parts of Ethiopia to compare satellite rainfall products with ground-based measurements at different spatial and temporal scale (e.g. Gebremichael et al., 2014; Worqlul et al., 2014; Beyissa et al., 2017; Gebremichael et al., 2017; Ayehu et al., 2018; Dinku et al., 2007; 2018; Gebrechorkos et al., 2018). However, many of these studies have mainly focused on the Upper Blue Nile basin and to some extent on whole and central Ethiopia. The results of these studies indicated that the skills of SREs vary with the characteristics of local climate, topography, and seasonal distributions of rainfall. Therefore, the reliability of satellite rainfall needs to be assessed against ground measurements to a specific area and temporal scales before it can be used in any subsequent application (Feidas, 2010). Other studies (e.g. Romilly and Gebremichael, 2011; Hirpa et al., 2010) documented validation of SREs over Awash River basin. None of these studies evaluated CHIRP, CHIRPS, and ARC2 SREs reliability in detail over the Awash River basin. Moreover, these studies used small numbers of ground-based measurements across the basin and lack validation based on a stratified sample, for example, based on elevation. The results of these and other similar studies indicated that SREs have major limitations in reproducing rainfall fields in mountainous regions. Therefore, using additional information, such as ground-based measurement data at different elevations (e.g. Hirpa et al., 2010), may be needed to assess the accuracy of rainfall estimates from SREs.

Therefore, further assessment and comparison studies in different mountainous regions of the country are important to explore the strength and limitations in SREs. The current study is intended to evaluate the performance of CHIRP, CHIRPS, and ARC2 with 37 ground-based measurements across various altitude at Awash River basin at dekedal (10 daily), monthly, and annually temporal scale. These satellite rainfall products are selected because CHIRP and CHIRPS have a relatively high spatial and temporal resolution (5.3 km), and are freely available. The ARC2 is also selected, as it has similar properties as CHIRP, and CHIRPS (TIR-based, relatively high spatial resolution and long time series).

METHOD

Study Area Description

The study focuses on the evaluation of dekadal, monthly, and annual climate data sources for the Awash River basin, Ethiopia. Awash River basin is one of the 12 river basins of Ethiopia and covers a total land area of 110,000 km2. It lies within 7° 53' N and 12° 08' 24" N latitude and longitudes of 37° 57' E and 43° 25'E (Figure 1,at the end of the text). The basin exhibits dramatic variation in elevation, making it suitable to examine satellite rainfall products over different elevations ranging from 4096 meter at highland to 105 meter above sealevel lowland. The movement of the Inter-Tropical Convergence Zone (ITCZ) and the influence of the Indian Monsoon throughout the year, mainly determine the climate pattern of the Awash River Basin (Romilly and Gebremichael, 2010). Based on the movement of ITCZ, the amount of rainfall and the rainfall timing, there are three seasons in the Basin. There are three seasons in the Awash River basin; these three seasons are Kiremt, which is the main rainy season (June-September in a highland and some parts of midland, and July-September over most parts of midland and lowland), the Bega, which is the dry season (October-January), and Belg, the small rainy season (February-May). The mean annual rainfall of the basin varies from about 1,400 mm in the highland northeast of Addis Ababa to 200 mm in the lowland of parts of the basin. The mean annual temperature of the Awash River Basin ranges from 20.8° C in the upper part to 29° C in the lower part.

Data source

Rainfall: Daily rainfall data for 66 independent weather stations located at a different elevation (highland, midland, and lowland) of the Awash River basin were obtained from National

Meteorological Agency (NMA) of Ethiopian (Table 1; at the end of the text after reference). It became apparent that there were missing data points; more than 70% of the data was available from 1987 to 2016 for most of the station included in the study.

Satellite-based Rainfall: Three satellite rainfall estimates are utilized in this study: CHIRP, CHIRPS, and ARC2. A detailed explanation of the CHIRP and CHIRPS products has been provided in Funk et al. (2014; 2015a/b) and for detailed information about ARC2 the reader is rereferred to (Novella et al., 2013). A summary of the explanation of the CHIRP, and CHIRPS algorithm and process is given below, while their summary is provided in (Table 2).

The CHIRP and CHIRPS algorithm combines three main data sources: (a) the Climate Hazards group Precipitation climatology (CHPclim), a global precipitation climatology at 0.05° latitude/longitude resolution estimated for each month based on station data, averaged satellite observations, elevation, latitude and longitude (Funk et al., 2012; 2015b); (b) TIR-based satellite precipitation estimates (IRP); and (c) ground-based rain-gauge measurements. The CHPclim is distinct from other precipitation climatologies in that it uses long-term average satellite rainfall fields as a guide to deriving climatological surfaces. This improves its performance in mountainous countries like Ethiopia (Funk et al., 2015a). The CHIRP/S algorithm involves the following steps (Funk et al., 2015a): (a) derive TIR precipitation estimates (IRP) from quasi-global geostationary satellite observations, which are generated using local regressions between Tropical Rainfall Measuring Mission multi-satellite precipitation analysis pentads (TMPA 3B42: Huffman et al., 2009; 2011) and cold cloud duration (CCD); (b) convert the IRP to percent-age anomalies and multiply by the CHPclim, producing the unbiased precipitation fields. (b) results in the CHIRP product (an unbiased IRP), which is a time series that goes back to 1981 at a spatial resolution of 0.05° latitude/longitude.

Merging of station data with CHIRP is done at pentad (5-day) and monthly time-scales with the pentads later rescaled such that the sum of pentads in a calendar month is equal to the monthly values. A daily version is created from the pentads and monthly fields. The daily CHIRPS use daily CCD percentages to discriminate between rain/no-rain events, and then the corresponding pentad rainfall is partitioned among the daily rain events proportional to the percentage of CCD. Two versions of CHIRPS are produced operationally. The preliminary version uses just GTS stations, which are then updated to the final version with more station data. Preliminary CHIRPS is available 2 days after the end of a pentad, while the final version is generated the third week of the following month. The CHIRP product may also have some inhomogeneity due to missing satellite slots in the early 1980s.

ARC2 is the second version of the ARC and is compatible with the algorithm of the Africa Rainfall Estimation Version 2 (RFE 2.0) (Novella et al., 2013). The product is a composite of 3-hourly geostationary infrared data, which makes it different from RFE, centered over Africa provided by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) and quality-controlled daily rainfall records acquired from the Global Telecommunication System (GTS) gauges (Gebrechorkos et al.,2018). ARC2 is consistent with the historical data sets of the Climate Prediction Center Merged Analysis of Precipitation (CMAP) (Xie and Arkin, 1997) and Global Precipitation Climatology Project (GPCP) (Novella et al., 2013). The data set is updated regularly and it is available at a spatial resolution of 0.1° covering the period from 1983 to present at a daily time-scale (Gebrechorkos et al., 2018). The ARC2 algorithm uses three-hourly thermal infrared (TIR) brightness temperature and a threshold of 235 K for discriminating raining clouds from non-raining ones. This temperature threshold is used to compute CCD from satellite TIR images. Then a simple linear relationship is adopted to convert CCD into rainfall amounts. Rain-gauge data made available through the World Meteorological Organization's GTS are used to adjust the final ARC2 product (Dinku et al., 2018). ARC2 is available at the International Research Institute climate data library (IRI/LDEO, 2016).

Data Analysis

A conventional approach for classifying the Awash River basin based on climatological, physical, socio-economic, agricultural, and water resource characteristics is to divide it into upper, middle, and lower Awash River sub-basins (Bekelle et.al 2017). However, this classification does not show any clear boundary among the sub-basin. Therefore, the current study classifies the basin into three large climatic zones such as highland, midland, and lowland based on the traditional climatic zone of Ethiopia, which mainly relies on altitude and temperature (MoA, 2000; Table 3; at the end of the text).

Quality Control of Rain-gauge Data

Filling of missing value, outlier detection, and homogeneity of data is important steps when analyzing time-series data. Most of the methods that estimate missing data, derive missing values using observations from the neighboring station (Sattari et al., 2016). In the context of the Awash River basin, the number of rain gauge stations with complete records for a longer period is very limited. As a result, filling of the missing values could produce the bias on the comparison. Therefore, the current study ignores filling of the missing data points and if there is station with more than 20% missing value, the station was excluded from the study (e.g. Romilly and Gebremichael, 2011). The approach chosen was to assess the data using a selected normal year, thereby minimizing the random errors in the data. Accordingly, 41 stations with continuous daily rainfall data are selected (Table 1). As there is variation in the years of missing data among stations, the current study selected only normal years i.e a year with Standardized Rain Anomaly (SRA) value between -1 to 1 (Eq. 1). This makes possible the determination of the dry, normal, and wet years in the record (WMO, 2012).

$$SRA = {\left(P-\overline{P}_m\right)}\Big/{\sigma_X} \qquad \text{Equation 1}$$

Where, SRA is standardized rainfall anomaly, P is annual rainfall in year t, $\bar{P}m$ is long-term mean annual rainfall over a period of observation and σx is standard deviation of rainfall over the period of observation. SRA is utilized according to McKee (1993) classification.

Spatial and temporal checks are carried out on the dekadal rainfall values (e.g. Dinku et al., 2018). The spatial quality check procedure compares the values of a given station with corresponding values of the nearby stations if there. The temporal check compares the consistency of a given dekadal value with values for the same dekad of other years at the same station. However, the quality check procedures may not remove all the errors. The homogeneity of the annual total rainfall was investigated by using statistical analysis tool (XLSTAT 2014), with the most widely used homogeneity tests such as; the Pettitt test, the Standard Normal Homogeneity Test (SNHT) for a single break, the Buishand Range test (BRT) and the Von Neumann ratio test (VRNT) to detect the inhomogeneity in the time series (e.g. Firat et al., 2010). According to Wijngaard et al. (2003), this test has been found useful for testing the homogeneity of the climate dataset. According to Wijngaard et al. (2003), the results of the tests are categorized into three classes according to the number of tests rejecting the null hypothesis as follow: Class A (as Useful): the series that rejects one or none of the Ho under the four tests at 5% significance level are considered as homogeneous and can be used for further analysis. Class B (as Doubtful): the series that reject two Ho of the four tests, the series has the inhomogeneous signal and should be critically checked before further analysis. Class C (as Suspect): when there are three or all of the tests reject the null hypothesis at 5% significance level. In this category, the series can be deleted or ignored before further analysis.

Comparison of Satellite Rainfall Etimation against Ground-based Mesurments

The comparison between gridded SREs and ground-based measurement can be made using either grid-to-grid or point-to-grid comparison methods. However, an attempt made to convert point ground observations to gridded interpolated dataset led to poor results due to uneven geospatial distributions of gauge stations (Ayehu et al., 2018). Therefore, in this study point-to-grid comparison approach is used. For each station included in the study, the grid values of satellite rainfall products containing the stations were extracted (e.g. Ayehu et al., 2018). Spatial variability is assessed based on averaged annual rainfall for the considered period derived from both ground-based observation and SREs with the aim of identifying regions where SREs were highly comparable to ground observation. The spatial distribution of rainfall data is directly interpolated by Inverse Distance Weight (IDW) technique with a power of two (e.g. Xianghu, 2014; Eq. 2). The IDW estimator is calculated for the location 0, where no measurements are present as:

$$\hat{z}(u_0) = \frac{\sum_{i=1}^{n} z(u_i) d_{io}^{-r}}{\sum_{i=1}^{n} d_{i0}^{-r}} \quad \text{Equation 2}$$

Where is the distance between the location 0 and each observation and r is an exponent.

Coefficient of variation (Eq. 3) and time series analysis is used to investigate the dekadal, monthly and annual variability of rainfall for both datasets. The time series of the 37-station is classified in to three homogeneous classes (station located at highland, midland, and lowland). The time series analysis is performed by averaging the long year dekadal and monthly data in to homogeneous classes and presented graphically.

$$CV = \left[\frac{S}{\bar{X}}\right] \times 100 \quad \text{Equation 3}$$

Where, CV is the coefficient of variation; $\bar{X}$ is the average long-term rainfall and S is the standard deviation of rainfall.

To evaluate the performance of satellite rainfall product against ground-based measurement four pairwise comparison statistics are used (Table 4, at the end of the text). Pearson correlation coefficient (Eq. 4) is applied to evaluate the agreement of individual products: satellite (S) to ground-based measurement (G) with perfect score of one (Gebrechorkos et al., 2018).

$$r = \frac{\sum_{i=1}^{n} (G_i - \bar{G})(S_i - \bar{S})}{\sqrt{(G - \bar{G})^2 (S - \bar{S})^2}} \quad \text{Equation 4}$$

Where r: is the person correlation coefficient; G: gauge rainfall measurement; $\bar{G}$ - average gauge rainfall measurement; S: Satellite-based rainfall estimate; $\bar{S}$ average Satellite-based rainfall estimate; N: number of data pairs.

Root Mean Square Error, RMSE (Eq. 5), is a frequently used measure of differences between two variables (acquires only positive values) lower RMSE values indicate greater central tendencies and generally smaller extreme errors and have a perfect score of 0 (Toté et al., 2015)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (S_i - G_i)^2}{n}} \quad \text{Equation 5}$$

Nash-Sutcliffe efficiency coefficient, Eff (Eq. 6) shows how well the estimate predicted the observed time series, and it varies from minus infinity to one: negative values mean that the gauge means is better than the satellite-based estimate, zero means that the gauge mean is as good as the estimate, and 1 corresponds to a perfect match between gauge measurements and satellite-based estimates (Toté et al. 2015).

$$Eff = 1 - \frac{\sum_{i=1}^{n} (S_i - G_i)^2}{\sum_{i=1}^{n} (G_i - \bar{G})^2} \quad \text{Equation 6}$$

Bias (Eq. 7) is a measure of how the average satellite rainfall magnitude compares to the ground rainfall observation. A value of 1 is the perfect score. A bias value above (below) 1 indicates an aggregate satellite overestimation (underestimation) of the ground rainfall amounts (Ayehu et al.,2018).

$$Bias = \frac{\sum_{i=1}^{n}(S_i}{\sum_{i=1}^{n} G_i} \quad \text{Equation 7}$$

A detailed description of these statistics can be found in Toté et al. (2015) and Thiemig et al. (2012).

RESULTS AND DISCUSSIONS

Data Processing and Quality Control

Data gaps, errors, and low density of rain gauge networks are the main challenges especially over the lowland parts of the basin. For example, two station from midland (Merssa and Zequala), three station from lowland (Adaitu, Dubty and Gawane), and one station from highland (Sendefa) at which there is at least greater than 15% missing value (Table 1, at the end of the text). As a result, the current study selects only the normal year (the year with SRA value between -1 to 1). Depending on available data 15-25 years (within 1987 –2016), is considered as an ideal period for evaluating climate condition of last three decades. Low density of the rain gauge network and data limitation over the Awash River basin is also noted by Hirpa et al. (2010).

As homogeneity test analysis indicates, 37 stations are found homogeneous whereas four stations (Addis Alem, Awash melkam, Aliyu amba, and Tulu bolo) are found inhomogeneous. 4 out of 41 stations are considered inhomogeneous whose change points are found to be insignificant at 5 % level of significance by the four-homogeneity test (SNHT, Pettitt, BRT, and VNRT). For example, homogeneity test at Addis Ababa Bole station shows station data homogeneity without any changing point (Figure 2). The station found inhomogeneous are excluded from further analysis. For example, all three tests (Pettitt, SHNT, and BRT) detected a change point around 2001 in Addis Alem station (Figure 3). The outcome of the tests is that most tests detected change points at around the same years at almost all stations. The likely reason for this inhomogeneity may be due to non-climatic factors, which make data inhomogeneous such as; location of the stations, type of instrument, observing practices and station environment. Karabörk et al. (2007) noted that simultaneous inhomogeneities may arise from simultaneous changes in observational routines.

Temporal Variability of Rainfall using Ground-based Measurement and SREs

As per the ground-based measurement, the average annual aerial rainfall distribution across the basin varies from 454.8 mm yr-1 in lowland to 1070.1 mm yr-1 highland. The highest variation in annual areal rainfall is observed mainly over the lowland parts of the basin (Table 5, at the end of the text). From the ground-based measurement result, it can be observed that over some parts of midland and most highland parts of basin, the main rainy season began starting from June first and reached its peak in July to August, then the rainfall decreased sharply starting from beginning of September to second of October. The dry season set in the third of October and lasted through January. Meanwhile, the main rainy season in most midland and lowland on-set very fast from the first of July and reach its peak on the third of August, then the rainfall decreased sharply from the first of September to the second of October. The short rainy season start in February and end in June (Figure 4).

In general, the three SREs (CHIRP, CHIRPS, and ARC2) have good agreement with ground-based measurement. CHIRPS and CHIRP described almost similar spatial rainfall distribution with the ground-based measurements over the basin, while ARC2 product shows an underestimation of high rainfall rate and slight overestimation low rainfall value (Table 5). Besides, the ARC2 product showed the highest coefficients variation value in areal annual rainfall than the other SREs. Charles et al. (2005) documented moderate variation in annual rainfall is (CV < 30%), therefore the annual rainfall variation showed by ground-based measurements and SREs is moderate at most parst of the basin.

The basin exhibits tremendous spatial variability in the average annual areal rainfall. All products show that the rainfall pattern is controlled by topography: high-elevation areas receive more rain than do low-elevation areas. This may be due to orographic uplifts since in most case rainfall increases with elevation due to the orographic uplifts (Worqlul et al., 2014). This is in line with other studies (e.g. Romilly & Gebremichael, 2011 and Bekele et al., 2016). As the time series analysis indicated, there is inter-annual rainfall variability in the basin. This might be due to the movement of ITCZ, which causes most of the inter-annual rainfall variability in Ethiopia (Seleshi & Zanke, 2004). Results are in agreement with the other studies (e.g. Degefu, 1987; Edosso et.al., 2010). The high rainfall variation based on the CV is also observed mainly over the lowland parts of the basin. A similar conclusion was arrived by Mersha (1999; 2003) who reported that the rainfall variability is higher in areas of low annual rainfall.

Spatial Variability Between Ground and Satellite-based Rainfall Estimation

The spatial distribution of annual averaged rainfall for selected normal years derived from ground-based measurement, and SREs (CHIRPS, CHIRP and ARC2) are presented in (Figure 5, at the end of the text). As the ground-based measurement indicates, the rainfall varies in relation to the topography from highland to lowland (Figure 5). The largest annual rainfall occurred in the highland part of the basin (with annual rainfall as high as 1,300 mm), while the lowest one is observed in the lowland (about 190 mm). ARC2 showed underestimation of higher rainfall value and slight overestimations of low rainfall rates (Figure 5A). CHIRP and CHIRPS (Figure 5B, & 5C) showed almost similar rainfall distribution with ground-based measurements respectively. In addition, absolute values varied considerably from one dataset to another and their distribution regions differ from that of rain gauge estimates, especially for ARC2. As the result indicates, the inter-annual rainfall variability shown here is similar to the result obtained in temporal rainfall variability analysis. The possible reason and explanation for this variability were given under section 4.2 and will not repeated here. The most likely reason for the poor capacity of ARC2 to estimate the actual rainfall amount might be due to sensor signal unable to penetrate the clouds (Thiemig et al., 2013). Another likely possible reason for underestimations of ARC2 might be attributed to the complex topography of the validation site that may reduce the ability to identify rainy clouds (Dinku et al., 2007; Funk et al., 2015b). This result was consistence with the studies of (Gebremicael et al., 2017; Ayehu et al., 2018). Which indicates that satellite rainfall products have challenges to estimate orographic precipitation in basins with a complex topograph.

Validations of SREs With Ground-based Measurement

SREs Dekadal Validation: The correlation between ground-based measurements and SREs shows wide scatter for all the products at dekadal time scale (Figure 6). The wider scatter is mainly observed over the lowland than the other parts of the basin, and this wider scatter is mainly attributed to ARC2. The CHIRPS products have less scatter compared to the other products. On the other hand, there is no considerable difference between CHIRP and CHIRPS. As the validation result at dekadal scale shows (Figure 7, at the end of the text), all products represent the spatial distribution of rainfall reasonably well. However, there are also differences among the different products. ARC2 has the lowest validation result over lowland (Eff = 0.15), and (CC= 0.63), and highest result over the highland parts of the basin (Eff = 0.69), and (CC = 0.84). It has the highest random error over highland and midland parts of the basin (RMSE = 27 and 28). In terms of bias, it shows under and overestimation of rainfall rate at different elevations of the basin (Figure 7). CHIRPS followed by CHIRP has better validation result (Eff ≥ 0.44), (CC ≥ 0.65) and lowest random error with little or no bias compared to ARC2 products over different elevations of the basin. Overall dekadal validation and comparison indicated that CHIRPS followed by CHIRP has a high level of correspondence with ground-based observations and may have a useful skill for various functions in the study area. As the results scatter plot indicates (Figure 6), ARC2 has wider scatter than the other products may be due to missing of a significant number of rainfall event. ARC2 miss orographic rainfall processes because of the cold temperature thresholds used by ARC2 algorithms (Dinku et al., 2018). This wider scatter (variation) is mainly observed over the lowland than the other parts of the basin. This scatter may be attributed to uncertainty in station locations, and uncertainties associated with some gauge observations. A similar result was arrived by (Romilly & Gebremichael, 2011) in the evaluation of SREs over Ethiopia, and reported that variation in the performance of the SREs is larger at elevations lower than 1500m.

As the results of validation statistics indicate, the two SREs (CHIRPS, CHIRP) shows good agreement with ground-based measurement over the highland and weakest but still good agreements over the lowland parts of the basin (Figure 7), however, ARC2 shows slight overestimations of low rainfall and underestimations of high rainfall amount substantially over the lowland and highland parts of the basin respectively (Figure 7). The worst result of ARC2 may be attributed to two main factors. The first is that ARC2 uses a single rain/no-rain threshold (235 K) for the whole of Africa (Novella and Thiaw, 2013). As a result, ARC2 may miss rainfall from warm cloud processes such as orographic rains. The other factor is that ARC2 uses three-hourly, as opposed to half-hourly, TIR observations. As a result, it may miss some short-lived rainfall events, which are frequent over the Tropics (Dinku et al., 2018). A likely reason for overestimations ARC2 at lowland, ARC2 overestimates rainfall amounts over most the dry and warm regions probably because of sub-cloud evaporation and other factors (Dinku et al., 2018). CHIRP is much better than ARC2. It has a little or no bias and higher validation result than ARC2. In addition to the algorithm itself, CHIRP has one main advantage over ARC. This advantage is the use of carefully generated gauge–satellite climatology, CHPclim (Funk et al., 2015a) to remove mean biases. However, CHIRP, and CHIRPS rainfall fields show some low rainfall values over some areas where ground observations and the other satellite products show zero rainfall values. This artefact may be due to the use of TRMM TMPA as "truth" data in the TIR estimation procedure. Since the TMPA is at 0.25° resolution, training to this data may produce light rain (Dinku et al., 2018). The CHIRPS product has not been discussed separately above because it has been shown that it is very similar to CHIRP. There is no substantial difference observed between these two products. This may be due to the addition of the new station observations for each dekad of each year should have improved CHIRP (Dinku et al., 2018).

SREs Monthly Validation: As expected, the monthly aggregations have reduced the scatters considerably, but wider scatter and underestimation of rainfall amounts by ARC2 stand out in (Figure 8). Similar to dekadal validation the higher scatter is mainly observed over the lowland parts of the basin. CHIRPS and CHIRP products have less scatter. The validation statistics have also improved with higher correlations (CC ≥ 0.89, and 0.81) skill (Eff ≥ 0.79, and 0.65) for CHIRPS and CHIRP products respectively over the different elevation of the basin compared to the dekadal (Figure 9), but shows slight overestimations of rainfall rate over some parts of basin. The performance of ARC2 has also shown improvement over the different elevations of the basin, but ARC2 shows an under and overestimation rainfall rate over different elevations of the basin. However, comparing both topographic features, the magnitude of underestimation is greater than that of overestimation in most products. The lowest RMSE score (21.8 mm month−1) is attributed to CHIRPS, while the highest RMSE score (61.5mm month-1) is attributed to ARC2 (Figure 9). The agreement between the satellite rainfall products and ground-based measurement improved with the increase in time scale (from dekadal to monthly). This may because errors at smaller time scales offset each other when aggregated (Gebremicael et al., 2017). Many studies (e.g., Gebremicael et al., 2017; Gebrechorkos et al., 2018; Ayehu et al., 2018) reported that the performance of satellite rainfall estimates improved as time step increased. In contrast, the performance of these products is not uniform with an increasing spatial scale. All satellite rainfall products show good agreement with ground-based measurement over the highland and weakest agreements over the lowland parts of the basin. This is the contradictory result with studies performed in the Tekeze-Atbara Basin (Gebremicael et al., 2017) and other parts of the country. These differences could be due to factors such as; uncertainty in gauge measurement, data used and gauge density. Since there were more data Gaps, measurement discontinuities and low density of the rain gauge network in lowland than the other parts of the basin. As, a well-designed rain gauge network with a sufficient number of rain gauges, can reflect the spatial and temporal variability of rainfall in a catchment (Yeh et al., 2011).

SREs Annual Validation: Unexpected, low performance is obtained for annual timescale compare to monthly and dekadal timescale. As a result of the scatter plot indicates, the SREs show wide scatter for all products at the annual time scale compared to dekadal and monthly time scale (Figure 10). The wider scatter is mainly observed over the highland than other parts of the basin. This is contradictory with dekadal and monthly analysis. The validation statistics is also considerably reduced, most of the SREs shows slight under and overestimation of rainfall amounts, and they are weakly correlated with ground-based measurement ($r < 0.50$), lower skill and large random errors particularly over the highland parts of the basin (Figure 11). In general, CHIRPS followed by CHIRP shown a satisfactory agreement with the ground-based measurement, and the weakest result is attributed by ARC2 at the annual time scale. In terms of Bias, CHIRPS, and CHIRP presented Bias scores approximately equal to 1, but ARC2 shows an under and overestimation of high and low rainfall rate respectively. The lowest RMSE score (70.4 mm yr−1) over the lowland is attributed to CHIRPS and the highest RMSE score (388.2 mm yr−1) over the midland is attributed to ARC2.

Unexpectedly, the worst performance is obtained at annual timescale compare to dekadal, and monthly. RMSE increase when time step increase, this may be due to increasing rainfall amount when aggregated into a large scale. All SREs showed slight under and overestimations of rainfall rate with low validation statistics particularly over highland parts of the basin. Most validation and comparison studies are conducted at daily, dekadal and monthly temporal scales (e.g., Gebremicael et al., 2017; Gebrechorkos et al., 2018; Ayehu et al., 2018; Tufa Dinku et al., 2018), and reported that the performance of satellite rainfall estimates improved as time step increased. In contrast, the results of this study show that the performance of satellite rainfall products is not uniform with increasing time steps from monthly to yearly. This may be due to the data and location selected by the current study. A similar result was obtained by Zambrano et al., (2017), and reported that the possible sources of error may be due to aggregations of small systematic biases. The result of the topographic features demonstrates that the overall validation result of SREs is better in lowland than in highland areas. This suggests that the satellite product challenges to estimate orographic rainfall in complex topographic areas like the Awash River basin. This result is in agreement with other studies (e.g. Gebremicael et al., 2017; Ayehu et al., 2018; Dinku et al., 2007). In general, CHIRPS followed by CHIRP is the best performing SREs in characterizing basin rainfall characteristics than ARC2 at all temporal and spatial scale. The better performance of CHIRPS can be explained by the fact that it to considers topographic effects and its high spatial resolution (Katsanos et al., 2016). A similar result has been obtained by Dinku et al. (2018); Gebrechorkos et al. (2018) over East of Africa; Ayehu et al. (2018); Bayissa et al. (2017) over Upper Blue Nile Basin, and Duan et al. (2016) over Italy and Toté et al. (2015) over Madagascar are among many other studies that compare CHIRPS with other different SREs and found CHIRPS as best-performing SREs.

CONCLUSIONS

The objective of the study is to assess and compare three widely used high-resolution satellite-based rainfall data (CHIRPS, CHIRP, and ARC2) with 37 rain gauge data located at different elevations at Awash River basin, and investigate their spatial and temporal characteristics at dekadal, monthly and annual rainfall. A point-to-grid-based comparison is carried out at different temporal scale using continuous statistical validation tools. Spatial variability is assessed by interpolating the average annual rainfall for the normal year using IDW with power 2. The coefficient of variation (CV) and time series analysis was used to investigate the temporal rainfall variability.

The data limitation and low density of the rain gauge network were the main challenges especially over the lowland parts of the basin. The basin exhibits tremendous spatial variability in the mean annual rainfall, and both products (ground-based measurement and SREs) show that the rainfall pattern is controlled by topography. The highest rainfall variability is observed mainly over the lowland parts of the basin. Overall, good agreement has existed between the SREs and ground-based measurements in the characterizing of rainfall distribution over the basin. CHIRPS and CHIRP show a consistence rainfall distribution with ground observation over different elevations of the basin. Meanwhile, ARC2 shows an under and overestimation of the rainfall rate over the different elevations of the basin. There were no significant differences between CHIRP and CHIRPS. The agreement between the satellite rainfall products and ground-based measurement improved with an increase in time scale (from dekadal to monthly), but surprisingly the performance of these products is decreased when aggregated into an annual time scale. In addition, all SREs shows good agreement with ground-based measurements over the highland, and weakest agreements over the lowland parts of the basin at dekadal and monthly time scale, but at annual time scale all product shows better performance over the lowland parts of the basin. The reliable performance of CHIRPS at different elevations could make the product more appropriate for various hydrological and rainfall analysis functions for the basin. In addition, the product can be used to check the acceptability of available ground-based measurement or substitute ground observations in the basin where ground station data are not available or accessible in most case. This study suggests the use of satellite data in regions where the spatial network of stations is highly sparse. It seems relevant to use more rain-gauge stations with complete daily data in future studies, and a pixel-to-pixel analysis can be applied to conduct in-depth analysis. Furthermore, studies should be also carried out to understand the reasons why the performance SREs is not uniform with increasing time step over different elevations at the Awash River basin.

References

1. Ayehu G.T., Tadesse T., Gessesse B., Dinku T.. Validation of new satellite rainfall products over the Upper Blue Nile Basin, Ethiopia. 2018; 1921–1936.
2. Bayissa Y., Tadesse T., Demisse G., Shiferaw A.. Evaluation of Satellite-Based Rainfall Estimates and Application to Monitor Meteorological Drought for the Upper Blue Nile Basin, Ethiopia. Remote Sens. 2017; 9(7): 669.
3. Behrangi A., Behnaz K., Tsou C., Amir A., Kuolin Soroosh S., Bacchetta N. Hydrologic evaluation of satellite precipitation products over a mid-size basin. J. Hydrol. 2015; 397:225-237.
4. Bekele D., Alamirew T., Kebede A., Zeleke G., Melesse A.M.. Analysis of rainfall trend and variability or agricultural water management in Awash River Basin, Ethiopia. J. Water Clim. Chang. 2017; 8:127–141.
5. Charles V.J., Ellen M.D., Pamela A.G. Carmen R.. Geospatial indicators of emerging water stress: an application to Africa. AMBIO A J. Hum. Environ. 2005; 34: 230–236.
6. Degefu W. Some aspects of meteorological drought in Ethiopia. In M. H. Glantz (Ed.), Drought and Hunger in Africa. Cambridge: Cambridge University Press. 1987.
7. Dinku T., Block P., Sharoff J., Hailemariam K., Osgood D., Corral J., Thomson M.C.. Bridging critical gaps in climate services and applications in Africa. 2014; 1–13.
8. Dinku T., Ceccato P., Grover-Kopec E., Lemma M., Connor S.J., Ropelewski C.F.. Validation of satellite rainfall products over East Africa's complex topography. International Journal of Remote Sensing. 2007; 28(7): 1503–1526.
9. Dinku T., Funk C., Peterson P., Maidment R., Tadesse T., Gadain H., Ceccato P.. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society. 2018.
10. Duan Z., Liu J., Chiogna G., Disse M. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. Science of theTotalEnvironment. 2016; 573:1536–1553.