Comparison of satellite remote sensing derived precipitation estimates and observed data in Kenya

Joseph M. Macharia⁠a,⁎, Felix K. Ngetichb, Chris A. Shisanyaa

a Department of Geography, Kenyatta University, P.O. Box 43844-00100, Nairobi, Kenya
b Department of Land and Water Management, University of Embu, P.O. Box 6-60100, Embu, Kenya

ARTICLE INFO
Keywords:
Rainfall
Kenya
CHIRPS 2.0
TRMM-TMPA 3B42 V7
GPCC
MERRA

ABSTRACT
This study evaluated the accuracy of four satellite remote sensing (SRS) based products in predicting rainfall (amounts and spatial distribution) over Kenya between 1998 and 2013. The four SRS products used include; two satellite products (Climate Hazards Group InfraRed Precipitation with Station (CHIRPS 2.0) and Tropical Rainfall Measuring Mission - Multi-satellite Precipitation Analysis (TMPA) 3B42 version 7 (TRMM)), one gauge-interpolated product (Global Precipitation Climatology Centre (GPCC)) and one re-analysis product (Modern-Era Retrospective Analysis for Research and Application (MERRA)). The monthly precipitation data were evaluated for completeness, converted to individual raster files, projected to the World Geodetic System (WGS) 1984 - Universal Transverse Mercator (UTM) zone 37 N ensuring a similar processing extent, rescaled to a common resolution and reclassified following defined uniform intervals for ease of comparison. Thereafter, they were subjected to five different metrics based on eight agro-ecological zones (AEZs) of Kenya, in reference to observed rainfall data obtained from Kenya meteorological department (KMD). Results show that all SRS products both overestimated or underestimated rainfall amounts on a pixel to pixel comparison. Based on point to point proportion of variance evaluation (r²), TRMM best-estimated rainfall in the tropical cool humid (r² = 0.64), tropical warm humid (r² = 0.58) and tropical cool subhumid (r² = 0.39) zones and can be used for agricultural advisory services. The GPCC product best-estimated rainfall in the tropical warm semiarid (r² = 0.46) and warm tropical sub-humid (r² = 0.21), while CHIRPS 2.0 best-estimated rainfall in the tropical warm arid (r² = 0.33) and therefore the two products could be best used to predict rainfall in the ASALs and drought-related studies, with potential for irrigation. The MERRA product best-estimated rainfall in tropical cool arid (r² = 0.97) and tropical cool semiarid (r² = 0.53) and could, therefore, be best used for high elevation and drought-related studies. These results demonstrate the promising potential of the satellite remote sensed data in complementing the existing meteorological observed data which are often marred by inconsistency and scarcity, and hence unreliable in the existing agricultural advisory and other climate-based applications in Kenya, and sub-Saharan Africa at large. However, given the observed AEZ dependant variations in the satellite estimates, it is advisable to choose the most suitable SRS product for specific activities per AEZ and calibrate before utilisation.

1. Introduction

In many regions of the world, Kenya included, high quality measured weather data (MWD) are not readily available because of sparse rain gauge distribution (Dinku et al., 2011; Satgé et al., 2016) and due to many weather stations having relatively short and mostly incomplete or inconsistent historical records of observations (Zambrano et al., 2017). The available data from these stations are sometimes inconsistent with many gaps and do not spatially represent the whole country thus being unreliable (Mourtzinis et al., 2017). To complement the available measured weather data, precipitation estimates can be derived from satellite sensors (Ioannidou et al., 2016; Satgé et al., 2016).

Satellite datasets are becoming increasingly essential to fill in the spatial and temporal data gaps for climate-based applications in areas for which rain gauge data are unavailable or unreliable (As-Syakur et al., 2016; Shrestha et al., 2017; Zambrano et al., 2017; Bai et al., 2018). Satellite-based precipitation products have advantages over ground-based observations in terms of spatial and temporal resolution and areal coverage. They provide a potential alternative as the data source for data-sparse or ungauged areas (Meng et al., 2014; As-Syakur et al., 2016). The satellite estimates are, however, limited since

⁎ Corresponding author.
E-mail address: machariamjoseph@gmail.com (J.M. Macharia).

https://doi.org/10.1016/j.agrformet.2019.107875
Received 3 April 2019; Received in revised form 16 December 2019; Accepted 18 December 2019
0168-1923/ © 2019 Elsevier B.V. All rights reserved.
their sensors use one or several proxy variables to indirectly predict the relationship between observations and precipitation (Toté et al., 2015). Algorithms for satellite rainfall predictions are based either on thermal infrared (TIR) bands (from which cloud-top temperature can be inferred) or on passive microwave (PMW) sensors (Paredes-Trejo et al., 2016). Several global remotely sensed datasets, with historical records spanning 18 years or more, appropriate for climate-related studies are freely available (Funk et al., 2015; Zambrano et al., 2017).

Many gridded precipitation products, with varying temporal and spatial resolutions, have been generated and are freely available at a global scale (Duan et al., 2016). They are broadly categorised into four classes. The first category is the gauge-only products which are based on observations from rain gauge stations only, using different interpolation methods and often available at a coarser spatial resolution, more than 0.5° (Duan et al., 2016). They have been used widely for various applications including the Global Precipitation Climatology Centre (GPCC) monthly precipitation product (Schneider et al., 2014), the Climate Prediction Centre (CPC) unified gauge-based analysis of global daily precipitation (Chen et al., 2008) and the Climatic Research Unit (CRU) monthly precipitation (Harris et al., 2014). The second category of the precipitation products uses atmospheric models that combine satellite and in-situ observations of various atmospheric properties as inputs. These products include National Centers for Environmental Prediction-National Centre for Atmospheric Research (NCEP-NCAR) (Kalnay et al., 1996) and European Centre for Medium-Range Weather Forecasts (ECMWF) (Balsamo et al., 2015). The third category of precipitation product uses satellite-only products and uses either the infrared (IR) information frequently measured from geostationary satellites, the microwave (MV) information less commonly measured from low earth-orbiting satellites, or a combination of IR and MV information (Duan et al., 2016). The fourth category of precipitation products is the satellite-gauge products that combine two individual (gauge-only and satellite-only) products through different bias correction or blending procedures. They include TRMM (Tropical Rainfall Measuring Mission) multi-satellite precipitation analysis (TMPA) (Huffman et al., 2007), the CMORPH (CPC MOPPHing technique) (Joyce et al., 2004), PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) (Hsu et al., 1997) available at a 0.25°, or finer, spatial resolution and the recently released “satellite-gauge” type CHIRPS product (Climate Hazards Group InfraRed Precipitation with Station data) which provides precipitation at the finest spatial resolution of 0.05° (Funk et al., 2015).

These precipitation products have been evaluated for their accuracy at different spatial levels including, global, continental, regional, country and basin levels. Based on these previous evaluations, it is evident that the available products have a global orientation, and their performance varies from region to region and from one product to the other. Thus, evaluating the accuracy of these precipitation products based on independent measurements before use is of inevitable importance (Duan et al., 2016). Therefore, the objective of this study was to evaluate the accuracy of four selected monthly long-term satellite remote sensing-based precipitation products, compared to observed data over Kenya for the period between 1998 and 2013, as a potential alternative source of weather variable/indicator for agricultural advisory and other climate-related applications. Consequently, this study evaluated the accuracy of TRMM-TMPA 3B42 V7 (Huffman et al., 2007), CHIRPS 2.0 (Funk et al., 2015), GPCC (Schneider et al., 2014) and MERRA product (Rienecker et al., 2011) for precipitation amount estimation across Kenya’s agro-ecological zones.

2. Materials and methods

2.1. Study area

Kenya lies along the equator in East Africa (Fig. 1a) covering an estimated area of 582,646 km². About 2% of the Kenyan land is covered by water, and 80% is arid to semi-arid (ASALs) zones while approximately 17% is considered as high potential agricultural land (FAO, 1996). This high potential agricultural land sustains about 75% of the Kenyan population (Hugo and Mugulavai, 2010). Kenya lies between longitudes 34° E to 42° E, and latitudes 5° S to 5° N and borders Uganda to the west, Somalia to the east, Ethiopia and Sudan to the North and Tanzania to the south (Fig. 1a). Kenya is subdivided into 47 counties administratively, lying at an altitude between sea level and the peak of Mt. Kenya at 5199 m above sea level. It is classified into eight major agro-ecological zones (AEZ) (FAO, 1996). The zones are: tropical cool arid (TCA), tropical cool humid (TCH), tropical cool semi-arid (TCSA), tropical cool sub-humid (TCSH), tropical warm arid (TWA), tropical warm humid (TWH), tropical warm semi-arid (TWSA) and tropical warm sub-humid (TWSH) (Fig. 1b). These zones are associated with corresponding temperature variations ranging from below 0 °C freezing temperatures to 40 °C hot (Otolo and Wakhungu, 2013).

Most parts of the country experience bimodal rainfall regime with the "long rains" season being experienced from March to May (MAM) while the “short rains” occurring from October to December (OND) (Yang et al., 2015). According to Anyah and Semazzi (2006), these rainy seasons particularly coincide with periods of the year when the Inter-tropical convergence zone (ITCZ) is overhead the equator. The average annual rainfall is 680 mm with a variation from less than 250 mm in northern Kenya to over 2000 mm in Western Kenya. The climate is influenced by the ITCZ and relief, ranging from permanent snow above 4600 metres on top of Mt. Kenya to the driest parts like the Chalbi desert in Marsabit county in the northern part of the country.

2.2. Data sets sources and acquisition

Monthly precipitation (mm month⁻¹) estimates from four satellite products and the ground-based KMD monthly precipitation data, between 1st January 1998 and 31st December 2013, were acquired. The first satellite product was TRMM-TMPA 3B42 version 7 at 0.25° by 0.25° spatial resolution (Rienecker et al., 2011). The second satellite product was TRMM-TMPA 3B42 version 7 at 0.25° by 0.25° spatial resolution (Huffman et al., 2007). The third satellite product was GPCC at 0.5° by 0.5° spatial resolutions (Schneider et al., 2014) while the fourth satellite product was the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) V2.0 at a 0.05° × 0.05° resolution (Funk et al., 2015). The accuracy for the four satellite products was carried out in reference to observed data from Kenya Meteorological Department (KMD), the institution mandated to collect and store climate data in Kenya (GoK, 2012). For this study, data from 32 surface climate observing stations countrywide (Synoptic Stations) were used (Fig. 1c).

2.3. Processing and standardisation of precipitation data

The KMD and MERRA datasets were first assessed for completeness and, where data were missing, data reconstruction was done through interpolation. The datasets were managed in Microsoft Excel® spreadsheet, saved in comma separated version (.csv) file and uploaded to ArcGIS® 10.4 where they were preprocessed, converted to point layer and georeferenced accordingly. The KMD and MERRA datasets were interpolated to raster files using Kriging technique (Fig. 2). Kriging technique was chosen for this study due to the sparseness of the rain gauges across the country and in line with other studies carried out using sparsely distributed rain gauge data (Dinku et al., 2010; Kisaoka et al., 2015; Webbe et al., 2017). The TRMM-TMPA, GPCC and CHIRPS 2.0 products were already aggregated in monthly time-series rasters and did not require additional temporal modifications before
analysis. However, before further analysis, all the five products were transformed to World Geodetic System (WGS) 1984, Universal Transverse Mercator (UTM) zone 37 N projected coordinate system to ensure a similar processing extent for all products. They were further rescaled from their original spatial resolutions to that of the observed rainfall layer (0.0318° by 0.0318°) for uniformity. Lastly, they were reclassified following defined uniform intervals for ease of comparisons. A simplified flowchart showing the procedure used for processing and standardization of the SRS-based precipitation data is shown in Fig. 2.

2.4. Satellite products evaluation

Data from the time-series of the monthly precipitations were compared between satellite products and the KMD monthly data. Due to the uniqueness in topographical systems over Kenya, and to evaluate the potential regional variations particularly in line with agricultural productivity, the country was divided into eight agricultural zones as classified by (FAO, 1996). Annual differences between satellite data and KMD interpolated data were calculated from the spatially aggregated data to determine the level of accuracy amongst the datasets. For comparative evaluations, the datasets were subjected to five metrics: Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Multiplicable bias and Pearson’s correlation coefficient (CC). Analysis for ME, MAE, RMSE and Multiplicable bias was pixel to pixel and carried out spatially. For Pearson’s correlation, the rasters were first extracted to eight agroecological zones and thereafter converted to point data and analysed using R Studio version 1.0.143.

3. Results

3.1. Annual average rainfall based on monthly totals

Based on the pixel-level summary of the five dataset layers of multi-year mean monthly rainfall totals, it was established that KMD data (gauge) recorded the highest minimum rainfall amount (19.4 mm) in a pixel while the GPCC recorded the least minimum amount of data (4.0 mm). The TRMM product, on the other hand, recorded the highest amounts of maximum rainfall (217.9 mm) in a pixel while MERRA product recorded the lowest amount of maximum rainfall (159.7 mm) in a pixel. The average highest amounts recorded per pixel was obtained from the gauge data (69.0 mm) while the lowest average recorded per pixel was from the GPCC data source at 30.2 mm (Fig. 3).

3.2. Statistical analysis between satellite-derived products and gauge data

3.2.1. Mean error (ME)

Results indicate that all the satellite sources both overestimated and underestimated the rainfall amounts across the Kenyan agro-ecological zones (Fig. 4). Based on pixel to pixel analysis, CHIRPS 2.0 dataset had the highest margin of error (196 mm yr⁻¹) comprising of the highest
overestimation (118 mm yr\(^{-1}\)) but the lowest underestimation (78 mm yr\(^{-1}\)) while MERRA dataset had the highest overestimation (124 mm yr\(^{-1}\)) but the lowest underestimation (41 mm yr\(^{-1}\)) (Fig. 4).

There existed a distinct similarity in the prediction of rainfall for CHIRPS 2.0, GPCC and TRMM where the three satellite products tended to mainly overestimate rainfall in the high altitude zones of Kenya i.e. tropical cool humid (TCH), tropical cool arid (TCA), tropical cool sub-humid (TCSH) and tropical cool sub-humid (TCSH) zones of Kenya and which are zones that generally received high rainfall amounts. The three satellite products underestimated rainfall in the tropical warm semi-arid (TWSA), tropical warm sub-humid (TWSH) and the tropical cool semi-arid (TCSA) zones which are low altitude zones and generally received low rainfall amounts (Fig. 4). However, MERRA dataset showcased an inverse relationship, compared to the three others, in the prediction of rainfall (Fig. 4c). MERRA dataset underestimated rainfall in the high altitude zones but overestimated rainfall amounts in the low altitude zones of Kenya (Fig. 4).

3.2.2. Bias

Results indicate that all the four SRS products used for this study had both positive and negative bias (Fig. 5). On a point to point comparison, all the datasets scored relatively the same in reference to positive bias (underestimation) in the range of (0.91–0.93) (Fig. 5).

However, in terms of the negative bias (overestimation) the SRS products performed differently with CHIRPS 2.0, GPCC and TRMM exhibiting a similarity on negative bias in the high altitude zones of Kenya including the tropical cool humid (TCH), tropical warm humid (TWH) and tropical cool arid (TCA) (Fig. 5). MERRA product had the highest negative bias of 3.55, mainly exhibited in the low altitude zones (Tropical warm arid and tropical warm semi-arid (TWSA), while GPCC had the lowest negative bias of 0.38 (Fig. 5).

3.2.3. Mean absolute error (MAE)

CHIRPS 2.0 dataset displayed the lowest mean absolute error (MAE) for both the high altitude and low altitude zones of Kenya (Fig. 6a). The GPCC and TRMM followed similar projections with the highest MAE recorded in the low altitude zones such as tropical cool semi-arid (TCSA), tropical warm semi-arid (TWSA) and tropical warm arid (TWA) while the lowest MAE was recorded in the high altitude zones such as tropical cool humid (TWH) and tropical warm humid (TWH) (Fig. 6b and d). However, MERRA dataset had the highest MAE of 19.4 mm
across the SRS products and mainly recorded in the high altitude zones while TRMM had the lowest mean absolute error of 9.1 mm across SRS products (Fig. 6c).

3.2.4. Root mean square error (RMSE)

Results indicate that the four SRS products generally performed the same in the prediction of rainfall over Kenya (Fig. 7). CHIRPS 2.0 and MERRA data had the highest RMSE (11 mm), followed by TRMM (10 mm) while GPCC had the lowest RMSE (9 mm) across the Kenyan
agro-ecological zones (Fig. 7). It is worth noting that the accuracy for the four SRS products differed across the zones (Fig. 7). The accuracy for GPCC and TRMM was highest in the high altitude zones, MERRA’s was highest in low altitude areas while that of CHIRPS 2.0 dataset had a crosscutting effect in both high and low altitude zones though with a higher emphasis on the low altitude zones (Fig. 7).

3.2.5. Pearson's correlation

Correlation analysis based on the various Kenyan agro-ecological zones reveals discrete differences amongst the satellite products (Table 1). The CHIRPS 2.0 dataset had the highest, though weak, correlation coefficient (0.33) in predicting rainfall in the tropical warm arid (TWA), which is an arid zone receiving the least amounts of rainfall.
in Kenya. The MERRA satellite product had the highest correlation coefficient (0.97) in predicting rainfall in the tropical cool arid (TCA) zone, which is the highest altitude zone at the peak of Mt. Kenya, as well as the tropical cool semi-arid (TCSA) with a correlation coefficient of 0.53. The GPCC products had the highest correlation coefficient for the tropical warm semi-arid (TWSA) (0.46) and tropical warm sub-humid (TWSH) (0.21). The TRMM satellite product had the highest correlation coefficient in predicting rainfall in the tropical cool humid (TCH) (0.64), tropical cool sub-humid (TCSH) (0.39) and tropical warm humid (TWH) (0.58) (Table 1).

Fig. 6. Mean annual absolute error based on monthly rainfall totals (mm) over Kenya from 1998 to 2013 for (a) CHIRPS 2.0 data; (b) GPCC data; (c) MERRA data; and (d) from TRMM data.
4. Discussion

Overall the four SRS products captured most of the Kenyan regional variations in terms of rainfall distributions in the different agro-ecological zones and were reasonably comparable to the ground-based rainfall observations (Fig. 3). The accuracy in estimating rainfall for most of the satellite products (apart from MERRA) tended to be high in agro-ecological zones that received high amounts of rainfall as compared to the zones that generally received low amounts of rainfall. These agro-ecological zones that received the highest amounts of rainfall are the humid and sub-humid zones in the high altitude areas (tropical cool humid (TCH), tropical warm humid (TWH), tropical warm sub-humid (TWSH) and tropical cool sub-humid (TCSH)) while those that received lowest amounts of rainfall were the arid and

Fig. 7. Mean annual root mean square error based on monthly rainfall totals (mm) over Kenya from 1983 to 2013 for (a) CHIRPS 2.0 data; (b) GPCC data; (c) MERRA data; and (d) TRMM data.

ological data and satellite products across the various agroecosystems of Kenya semiarid (TWSA) and tropical cool semiarid (TCSA). These results are

The discrepancies observed in estimations of rainfall over Kenya as a result of overestimation and underestimations by the SRS products, could be as a result of the satellites having errors between the observation and precipitation estimates. As observed in previous studies (Toté et al., 2015; Rossi et al., 2017; Shrestha et al., 2017), the satellite-based rainfall estimates have random errors and non-negligible bias due to lack of direct relation between observation and estimated precipitation as well as due to inadequacy in sampling and algorithm imperfections. It could also be attributed to the error during collecting and digitising gauge data, especially while using manual methods which introduce uncertainties in the observed data as was noted by Satgé et al. (2016).

The relatively high variations in amounts of rainfall predicted by the SRS products particularly in the high agricultural potential areas (highlands) and low potential areas (arid and semiarid) areas (Fig. 4c) could be attributed to factors related to altitude. It is worth noting that the accuracy of the SRS products tended to be high in the low lying zones of L. Victoria basin compared to the zones at high elevation (e.g. Mt. Kenya), agreeing with other studies that assessed accuracy of SRS products across elevation (Toté et al., 2015; Paredes-Trejo et al., 2016; Shrestha et al., 2017). As was observed by Ochoa et al. (2014), the critical challenges in the estimation of precipitation from satellites estimate arise from the process scheme for the microwave (MW) and infrared (IR) data with the major problem with IR data processing being that the global algorithms do not consider the altitude of the agrome-
teor. The underestimation over the mountainous region (Fig. 4) could be ascribed to the warm orographic rain process, while the over-
estimations over the dry regions (Fig. 4) may be because of sub-cloud evaporation (Dinku et al., 2011). The variations in rainfall estimations in the drier agro-ecological zones of Kenya (Fig. 4) could be attributed to rainfall suppression by desert aerosols and surface effects amongst other error sources that severely affect satellite rainfall estimation in drier parts of the country as was noted by Awange et al. (2016).

Spatially, CHIRPS 2.0 data lacks the capacity to accurately predict rainfall amounts in the high elevation to rugged mountainous zones (e.g., Mt. Kenya) but accurately predicts rainfall amounts in regions which are relatively low-lying to flat regions (e.g., Lake Victoria basin). This agrees with other studies (Toté et al., 2015; Paredes-Trejo et al., 2016; Shrestha et al., 2017) who identified the accuracy of the CHIRPS 2.0 data being high in predicting rainfall in flat low-lying to open flat region but poor in predicting rainfall in the high elevation regions. The underestimation of rainfall by CHIRPS 2.0 datasets particularly in the high elevation zones (e.g. Mt. Kenya, tropical cool arid) could be attributed to complexities in orography as a result of great elevation differences (Fig. 4a). The accuracy in predicting rainfall amount for CHIRPS 2.0 data also tended to be higher in regions that generally received low amounts of rainfall and which are characterised by arid and semiarid conditions e.g. Tropical Warm Arid (TWA). These results however, disagree with Zambrano et al. (2017) and Bai et al. (2018) who noted that CHIRPS 2.0 product has low capacity in estimating rainfall amounts in areas that receives low rainfall amounts.

The GPCC product tended to accurately predict rainfall across the agro-ecological zones (Fig. 4b) with a relatively high level of agreement with the ground-based gauge observation across the zones (lowest ME, Bias and RMSE, Figs. 4b, 5b and 7b respectively). This could be due to the formulation (gauge-interpolation) of satellite estimates through an integration of a vast dataset based on several gauge stations observations across the world and which goes through a sophisticated method of quality control (Schneider et al., 2014). These results are consistent with other studies carried out using GPCC data and which identified a substantial agreement between GPCC satellite product and the observed gauge observation (Manzanas et al., 2014; Asfaw et al., 2017; Baidu et al., 2017). There was, however, a slight overestimation of the rainfall amounts in some sections of the tropical cool humid (TCH), though the average rainfall in the zone was lower than observed rainfall (Fig. 4b). These high altitude areas around the Lake Victoria basin and which generally received high amounts of rainfall, portrayed a strong correlation between satellite estimates and observed rainfall data ($r^2 = 0.61$, Table 1). This agrees with Jury (2017) in a study carried out in the Upper Zambezi catchment and whose findings identified a strong correlation between GPCC satellite data and observed rainfall. The GPCC product, however, underestimated (slightly) the prediction of rainfall amounts in low altitude zones (arid and semi-arid zones) that generally received low amounts of rainfall (ME = 58 mm, Fig. 4b). This slight disparity (pixel-pixel) agrees with Baidu et al. (2017) who identified that the GPCC data tends to underestimate rainfall amounts in most parts of Ghana, which is generally low-lying. Spatially, GPCC data was not accurate in predicting rainfall amounts in both the lowlands and the Kenyan highlands since it tended to underestimate the rainfall amounts in these zones but accurately predicted rainfall around L. Victoria region which receives high rainfall amounts (Fig. 4b).

The MERRA product tended to extremely underestimate rainfall amounts in the regions such as Lake Victoria basin which are high alti-
tude and generally received high rainfall amounts (Fig. 4c). The MERRA product had the least level of accuracy as indicated by the highest bias ($-3.65 \pm 0.91$), MAE (19.4 mm) and RMSE (19.4 mm), Figs. 5c, 6c and 7c respectively. This observation is inverse to the predictions by the three other SRS products, evaluated in this study, which had a high correlation in predicting rainfall in the zones that received high amounts of rainfall. These discrepancies could be attributed to a lack of influence from gauge data as the output is purely controlled by model parameterisations (Rienecker et al., 2011). However, as was noted by Gelaro et al. (2017), some of these deficiencies in MERRA data have been addressed by MERRA version 2 (MERRA-2) which assimilates precipitation observations to force the land surface. Spatially, MERRA data tended to closely predict rainfall amounts in the high elevation to rugged mountainous zones (e.g., Mt. Kenya) ($r^2 = 0.97$), though it inaccurately predicted the rainfall amounts in regions which are relatively low-lying to flat regions (e.g. Lake Victoria basin, TCH ($r^2 = 0.05$) and along the Kenyan coastal line, TCSH

<table>
<thead>
<tr>
<th>AEZ</th>
<th>CHIRPS 2.0</th>
<th>GPCC</th>
<th>MERRA</th>
<th>TRMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TCA 0.04</td>
<td>0.52</td>
<td>0.97</td>
<td>*</td>
</tr>
<tr>
<td>2</td>
<td>TCH 0.40</td>
<td>0.61</td>
<td>0.05</td>
<td>0.64</td>
</tr>
<tr>
<td>3</td>
<td>TCSA 0.01</td>
<td>0.30</td>
<td>0.53</td>
<td>0.35</td>
</tr>
<tr>
<td>4</td>
<td>TCSH 0.15</td>
<td>0.25</td>
<td>0.01</td>
<td>0.39</td>
</tr>
<tr>
<td>5</td>
<td>TWA 0.33</td>
<td>0.25</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>TWH 0.08</td>
<td>0.31</td>
<td>0.38</td>
<td>0.58</td>
</tr>
<tr>
<td>7</td>
<td>TWSA 0.32</td>
<td>0.46</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>8</td>
<td>TWISH 0.05</td>
<td>0.21</td>
<td>0.16</td>
<td>0.08</td>
</tr>
</tbody>
</table>

* Could not be computed due to a similarity in the rainfall amounts in the point data under TRMM satellite product.
* AEZ (Agro-ecological zones): TCA, tropical cool arid; TCH, tropical cool humid; TCSA, tropical cool semiarid; TCSA, tropical cool sub-humid; TWA, tropical warm arid; TWH, tropical warm humid; TWSA, tropical warm semiarid; TWISH, tropical warm sub-humid.
* CHIRPS 2.0, Climate Hazards Group Infrared Precipitation with Station.
* GPCC, Global Precipitation Climatology Centre.
* MERRA, National Aeronautics and Space Administration: The Modern-Era Retrospective analysis for Research and Application.
* TRMM, Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B42 (version 7).
The TRMM data accurately and qualitatively predicted rainfall over Kenya and relative agreement with the observed gauge data (lowest Bias, MAE and RMSE, Figs. 5d, 6d and 7d respectively). This could be due to its ability to provide a calibration-based sequential scheme for combining precipitation estimates from multiple satellites, as well as gauge analyses where feasible at fine scales (0.25° × 0.25° and 3 hourly) (Huffman et al., 2007). In the current study, the TRMM product at 0.5° × 0.5° had the best correlation coefficient (r²) for most (TCH = 0.64, TWH = 0.58 and TCSH = 0.39) of the agro-ecological zones in Kenya (Table 1). These results agree with Meng et al. (2014) who noted that the monthly TRMM data have a much better linear relationship with the gauge rainfall data with a high determination coefficient. Spatially, TRMM data accurately predicted rainfall amounts in areas that generally received high amounts of rainfall (e.g., Lake Victoria basin and particularly the tropical cool humid (TCH) zone) with a strong correlation (r = 0.64, Table 1) and an overestimation of rainfall in these zones (ME = 102 mm, Fig. 4d). These results disagree with Katsanos et al. (2016) and Ouatiki et al. (2017) who identified TRMM product tending to, in general, overestimate rainfall in dry regions and underestimate rainfall in humid regions. However, according to Shrivastava et al. (2014), the TRMM data captures rainfall distribution qualitatively, but there exist differences in the quantities of predicted rainfall with respect to the ground observations.

Just like the GPCC data, TRMM underestimated the prediction of rainfall amounts (Fig. 4d) mainly in the low altitude zones, mainly arid and semiarid agro-ecological zones that received lower rainfall amounts and which forms 80% of the Kenyan agro-ecological zones (tropical warm arid (TWA), tropical cool arid (TCA), tropical warm semiarid (TWSA), tropical cool semiarid (TCSH) and tropical warm sub-humid (TWSH)) (FAO, 1996). These results agree with Shrivastava et al. (2014), Meng et al. (2014) and Dhib et al. (2017) who identified that the TRMM prediction on rainfall was slightly lower (underestimates) than the observed gauge data in most dry parts of the study area. However, TRMM data was not accurate in predicting rainfall amounts in the Kenyan highlands/mountainous zones due to the similarity in the data resulting from the coarse spatial resolution (0.5° × 0.5°), for the small area covered by the tropical cool arid (TCA) zone (*, Table 1). According to Rossi et al. (2017), TRMM data fails to accurately capture rainfall amounts in the mountainous areas with a high level of underestimations. As was further noted by Rossi et al. (2017), the instruments used to create the TMPA datasets also struggle with resolving shallow orographic rainfall, as do most satellite-based rainfall sensors, and as such, these tend to be biased in mountainous areas.

5. Conclusion

Based on the results from this study, we conclude that all the satellite products overestimate and also underestimate rainfall amounts based on the pixel to pixel differences. There was general agreement between SRS products and ground observations of rainfall across the various AEZs of Kenya. The accuracy of the SRS products in estimating rainfall was high in AEZs of Kenya that generally received high amounts of rainfall. There exists a clear difference in the capability of each satellite product in predicting rainfall in the various AEZs of Kenya and therefore careful choice of SRS product per AEZ needs to be carried out. The TRMM product was identified as the best-suited satellite product to predict rainfall in the high altitude humid zones and therefore could be best used to predict rainfall for general agricultural advisory services. The GPCC and CHIRPS datasets were found best suited for predicting rainfall in the low altitude zones and could therefore be best suited to predict rainfall in semiarid zones of Kenya, which are characterised with low agricultural production with potential for irrigation. Lastly, MERRA product was found best suited for predicting rainfall in the cool arid zones and could therefore be used for both mountainous terrains and drought-related studies in Kenya. These findings further indicate that SRS products have the potential to complement the lack of reliable and accurate data for agricultural advisory services and other applications such as climate-related studies. However, given observed AEZ-dependent variations in the suitability of the satellite estimates, it is advisable to not only choose the most suitable satellite product for any given AEZ in Kenya but also to calibrate the data before using.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgements

This research was partially supported by System for Land-based Emission Estimation in Kenya (SLEEK) and Kenya National Research Fund (NRF). We would wish to pass our gratitude to Kenya Meteorological Department for availing monthly data over Kenya for the period of study. We also thank Kenyatta University and University of Embu for their contribution towards this research.

References


