

Assessing the skill of gridded satellite and reanalysis precipitation products over in East and Southern Africa

Francis Kamau MUTHONI^{1*}, Francis Michael MSANGI¹ and Exavery KIGOSI¹

¹*International Institute of Tropical Agriculture (IITA), P.O. Box 10, Duluti, Arusha, Tanzania.*

*Corresponding author: Francis Kamau Muthoni, fkmuthoni@gmail.com

Received: March 11 2022, Accepted: October 24 2022

RESUMEN

La validación de los productos de precipitación en cuadrícula (GPP) aumenta la confianza de los usuarios y destaca las posibles mejoras de los algoritmos para manejar procesos complejos de formación de lluvia. Evaluamos la habilidad de tres GPP (CHIRPS-v2, CHELSA y TerraClimate) para estimar las observaciones de pluviómetros y comparamos las tendencias de precipitación derivadas de estos productos en las regiones del este y sur de África (ESA). Se usaron diagramas de Taylor y Eficiencia de Kling-Gupta (KGE) para evaluar la precisión. Se utilizaron las pruebas de Mann-Kendal modificada y el estimador de pendiente de Sen para determinar la importancia y la magnitud de las tendencias, respectivamente. Los tres GPP tuvieron un desempeño variado en rangos temporales y altitudinales. La habilidad de los tres GPP a escala mensual fue en general alta, pero mostró un rendimiento inferior en elevaciones superiores a 1500 m s. n. m., especialmente durante la temporada de Octubre-Noviembre-Diciembre (OND). Los tres GPP se desempeñaron igualmente bien entre el rango de elevación de 1001 a 1500 m s. n. m. CHELSA- v2.1 fue más preciso a 0-500 m s. n. m. pero tuvo la habilidad más baja en los rangos de elevación de 501 - 1000, y arriba de 1500 m s. n. m., lo que ocasionó una sobreestimación de las tendencias de precipitación anual y estacional sobre los terrenos montañosos y las grandes masas de agua continentales. Las tendencias de precipitación cuantificadas revelaron una alta variabilidad espacio-temporal. En general, las tendencias de precipitación y la habilidad, derivadas de los datos de CHIRPS-v2 y TC, mostraron una convergencia sustancial excepto en Tanzania. Nuestros resultados enfatizan la importancia de la validación de los conjuntos de datos climáticos para evitar la propagación de errores en diferentes modelos y aplicaciones. Así mismo demuestran que los datos de precipitación nuevos o de mayor resolución no siempre son los más precisos, ya que una actualización de los algoritmos puede introducir artefactos o sesgos.

ABSTRACT

Validation of gridded precipitation products (GPP) increases the users' confidence and highlights possible improvements in the algorithms to handle complex rain-forming processes. We evaluated the skill of three GPPs (CHIRPS-v2, CHELSA, and TerraClimate) in estimating the rain gauge observations and compared the precipitation trends derived from these products across the East and Southern Africa (ESA) region. We used Taylor diagrams and Kling-Gupta Efficiency (KGE) to assess the accuracy. A modified Mann-Kendal test and a Sen's slope estimator were utilized to determine the trends' significance and magnitude, respectively. The three GPPs had varied performance over temporal and altitudinal ranges. The skill of the three GPPs, at a monthly scale, was generally high but showed lower performance at elevations over 1500 masl, especially during the October-November-December (OND) season. The three GPPs performed equally well between the 1001 – 1500 masl elevation range. CHELSA-v2.1 was most accurate at 0-500 masl but had the lowest skill in both 501 – 1000 and above 1500 masl elevations, which caused over-estimation of the annual and seasonal precipitation trends over mountainous terrain and large inland water bodies. The quantified precipitation trends revealed high spatial-temporal variability. Generally, the skill and precipitation trends derived

from CHIRPS-v2 and TC data showed substantial convergence except in Tanzania. Our results emphasize the importance of validating climate datasets to avoid error propagation in different models and applications. Moreover, we demonstrate that new or higher-resolution precipitation data are not always accurate since an algorithm update can introduce artifacts or biases.

Keywords: Climate change and variability, satellite time series, trend analysis, CHIRPS-v2, CHELSA, TerraClimate.

1. Introduction

Precipitation variability is the primary driver of agricultural production in the predominantly rain-fed system in Sub-Sahara Africa (Adhikari et al., 2015). Smallholder subsistence farmers practicing rain-fed agriculture in Sub-Sahara Africa (SSA) are the most vulnerable to the impacts of climate change and variability (Cairns et al., 2013). Over the recent decades, the precipitation patterns across Africa have experienced significant changes in the amount (Cattani et al., 2018; Muthoni et al., 2019) and temporal shifts (Haghtalab et al., 2019; Atiah et al., 2021). In the East and Southern Africa (ESA) region, changes in precipitation amount have a direct impact on crop yields (Omoyo et al., 2015; Mkonda and He, 2018), the shift in crop suitability (Mumo et al., 2021) and the outbreak of crop pest and diseases (Kimunye et al., 2020; Niassy et al., 2021). Therefore, spatially explicit information on the magnitude of changes in precipitation over time and space is needed to support the design of appropriate adaptation measures.

Monitoring the spatial and temporal changes of climatic variables across Africa is challenging due to the limited availability of reliable rain gauge data. Available rain gauge networks are sparse, and their records are characterized by many gaps (Contractor et al., 2020; Dinku, 2019). Recently, there has been increased availability of time series gridded precipitation products (GPP) from three broad categories: (1) rain gauge only (GO), recorded solely from rain gauge data (e.g., Contractor et al., 2020), (2) model reanalysis (MRA), based on a numerical weather prediction models or data assimilation (Karger et al., 2017; Abatzoglou et al., 2018; Hersbach et al., 2020); (3) the satellite-based precipitation estimates (SPE); generated from satellites data only or blending of satellite, rain gauge and numerical weather models (e.g., Funk et al., 2015). The Rainfall Estimates on a Gridded Network (REGEN; Contractor et al., 2020) is an example of rain gauge-only gridded data. Existing

MRA data includes the TerraClimate (Abatzoglou et al., 2018), CHELSA (Karger et al., 2017), and ERA-5 (Hersbach et al., 2020). MRA precipitation data is produced by combining forecast model estimates with observations via data assimilation to generate optimized global estimates of climate data without spatial or temporal gaps (Gleixner et al., 2020). The blended SPEs mostly applied in Africa include the Climate Hazards Group InfraRed Precipitation with Station data version two (CHIRPS-v2; Funk et al., 2015) and the Africa Rainfall Estimate Climatology (ARC-v2.0; Novella and Thiaw, 2013). The data from GPPs are applied across different sectors to resolve weather-related issues such as drought/flood monitoring, early warning systems, agro-advisory, water management, and climate change analysis. These data are also applied to formulate evidenced-based climate change adaptation strategies. The choice of a GPP significantly influences the accuracy in a specific application area (Bobrowski et al., 2021).

Recent validation studies in the ESA region demonstrated that the monthly to annual aggregates of the SPE data have adequate skill to estimate rain gauge observations (Dinku et al., 2018; Muthoni et al., 2019; Muthoni, 2020) and, therefore, can complement the rain gauge data in data-scarce regions. However, the GPPs has their strength and weakness at different temporal scales and biophysical conditions. The main challenge is their ability to represent rain gauge observations over areas with complex meteorological patterns, such as mountainous terrain, inland water bodies, and coastlines (Dinku et al., 2007; Kimani et al., 2017). In mountainous terrain, different GPPs can yield significant differences in precipitation estimates over short distances (Henn et al., 2018). Rigorous evaluation and intercomparison of the GPPs are needed to assess their reliability under varied environmental contexts.

Previous evaluation of MRA precipitation datasets over the ESA region showed that the accuracy of

ERA-5 (25 Km) and ERA-Land (9 Km) had substantial spatial variation (Gleixner et al., 2020). These MRAs captured the precipitation seasonality well but exhibited low temporal and spatial correlation with rain gauge observations over the ESA region (Lemma et al., 2019; Gleixner et al., 2020). Although ERA-5 (25 Km) has shown tremendous improvements in capturing rainfall in East Africa compared to ERA-Land (9 km), both have a low spatial resolution. In contrast, the gridded SPEs that incorporate satellite and rain gauge data in their retrieval algorithm, such as CHIRPS-v2, are extensively evaluated over the ESA region. Reports show that CHIRPS-v2 has a better skill for reproducing observed rainfall (Dinku et al., 2018; Muthoni et al., 2019). However, comparing the SPE and MRA precipitation products has received less attention. Muthoni (2020) showed that CHIRPS-v2 performed better than TerraClimate rainfall in West Africa. Gleixner et al. (2020) and Lemma et al. (2019) showed that the CHIRPS-v2 product represented rainfall trends in East Africa better than the coarser spatial resolution ERA-Interim and ERA-5 datasets. Recent advances have produced newer MRA precipitation data with a higher spatial resolution (1 – 4 Km), such as CHELSA (1 Km) and TerraClimate (4 Km). Although the skill of the two MRA datasets is evaluated globally, an in-depth verification in regions with a critical need for reliable GPPs, like ESA, is lacking. Finer resolution or newer versions of climate data may improve the accuracy, e.g., ERA-5 improved skill for capturing precipitation in Africa compared to its predecessor, the ERA-interim (Gleixner et al., 2020). Nonetheless, this is not always the case since updates of the algorithms can introduce other artifacts or biases (Awange et al., 2019; Bobrowski et al., 2021). Therefore, there is a need to ascertain if the two newly available MRA datasets with a finer spatial resolution are more reliable than the legacy CHIRPS-v2 product.

Analysis of long-term precipitation trends is critical for unraveling the temporal progression of climatology to inform climate adaptation policies. Due to the paucity of rain gauge observation networks, the GPPs provide the only plausible data for monitoring the long-term trends of precipitation for locations without rain gauge stations. Nevertheless, examining the congruence of precipitation trends generated from GPPs that emphasize different aspects

of the rainfall regime is essential. Gridded time series data have been applied to monitor rainfall trends in the ESA region. However, the suitability of MRA data for identifying long-term trends is debatable (Bengtsson et al., 2004; Thorne and Vose, 2010; Dee et al., 2011). This setback is because observational datasets applied in the assimilation system are temporary and spatially inhomogeneous, which can cause jumps in the data that affect the retrieval of long-term trends. The ERA-interim and ERA-5 data in the ESA region showed inconsistent precipitation trends compared to CHIRPS-v2 (Gleixner et al., 2020). The assimilation systems of the atmosphere circulation models are continuously improved; therefore, the newly produced MRA data may improve the monitoring of long-term trends.

This paper assessed the skill of the CHIRPS-v2, and two MRA products, i.e., TerraClimate (TC) and CHELSA-v2.1, in reproducing the rain gauge observations at different temporal and altitudinal ranges over the ESA region. The study further investigates whether the higher resolution from CHELSA-v2.1 data (1km) improves the estimation of long-term precipitation trends compared to the coarser resolution of TC (4 km) and CHIRPS-v2 (5 Km). Moreover, the ability of the MRA dataset to capture long-term annual and seasonal precipitation trends was compared against satellite-based CHIRPS-v2 data.

2. Materials and Methods

2.1 Study Area

The study area covers approximately 2 million Km² encompassing seven countries in the East and Southern Africa (ESA) region, i.e., Tanzania, Kenya, Uganda, Rwanda, Burundi, Zambia, and Malawi (Fig. 1). The study area traverses a complex agro-ecological gradient characterized by high variability in topography, precipitation, temperature, and vegetation cover. The climate variability is primarily influenced by the seasonal movement of the intertropical convergence zone (ITCZ) and the warming of the Indian Ocean (Diem et al., 2014). Zambia, Malawi, and West - Central – Southern Tanzania experience unimodal precipitation, while northern Tanzania, Kenya, Uganda, Rwanda, and Burundi experience bimodal rainfall seasons (Seregina et al., 2019). Annual rainfall ranges between 250 – 2500 mm, while average temperature

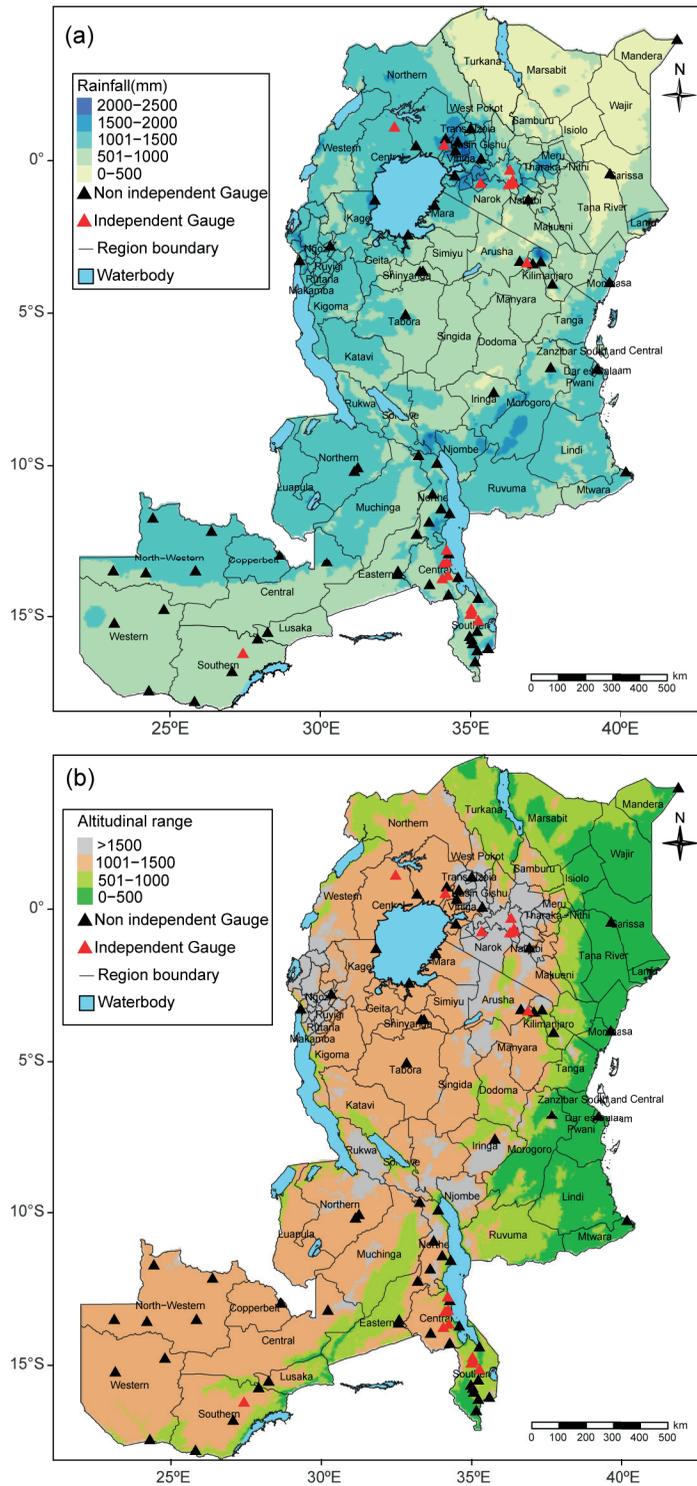


Fig. 1. Study area covering seven countries in the East and Southern African (ESA) region overlaid on the average annual rainfall (a) and a 30 m digital elevation model (DEM) from Shuttle Radar Topography Mission (SRTM) (METI and NASA, 2011). The location of rain gauge stations used to evaluate the accuracy of gridded climatic layers is shown.

ranges between 15°C – 40°C. Savannah vegetation dominates the region, although tropical rainforest extends to some parts of Uganda and Rwanda. North – Eastern part of Kenya experiences arid conditions, while Central Tanzania is semi-arid.

2.2 Rain gauge observation and satellite data

Data from 82- rain gauge stations (Table SI) distributed across the ESA region (Fig. 1) with records from 1981 to 2018 was collated from national meteorological services and private institutions, supplemented with data from the global summary of the month (GSOM; Lawrimore et al., 2016). Most of the stations were ingested in the algorithms for retrieving the three GPPs either for bias correction, downscaling, or interpolation. Therefore, re-using them to evaluate the accuracy of the GPPs is not independent. However, we recognize the scarcity and the very restrictive accessibility of data from the few existing stations in Africa (Dinku, 2019). The data from the synoptic and global transmission system (GTS) stations are mostly ingested in the original algorithms for retrieving the GPPs. The independent datasets are mainly from stations maintained by volunteers, and therefore their data quality may be low compared to those maintained by professional meteorologists (Dinku et al., 2018). Considering the above limitations, we undertook two evaluations: (a) with all datasets (including the non-independent) and (b) with only the independent stations to assess the reliability of existing independent stations. Although using only the independent stations is a good practice, the spatial coverage of the eligible rain gauge stations was inevitably reduced to 18, all located above 500 m above sea level.

Moreover, stations with low-quality data or less than 60 monthly observations were dropped. A total of 40779 monthly observations were collated and

applied for validating the GPPs. Table SI shows the information on the rain gauge station data used for validation.

The GPPs data was downloaded from the Climate Hazards Group Infrared Precipitation with Stations version two database (CHIRPS-v2; Funk et al., 2015), and the two MRA datasets acquired from the climatologies at high-resolution for the earth's land surface areas version 2.1 (CHELSA-v2.1; Karger et al., 2017) and the TerraClimate (TC; Abatzoglou et al., 2018) databases (Table I). The CHIRPS-v2 covers between 50°S – 50°N and 180°E – 108°W with 0.05° resolution (~5.5km) from 1981 to the near present (Funk et al., 2015). The CHIRPS-v2 data was created by blending the monthly precipitation climatology, the quasi-global geostationary thermal infrared (IR) satellite observations from two NOAA Climate Prediction Centre (CPC), the Tropical Rainfall Measuring Mission (TRMM) 3B42 product from NASA, the atmospheric model rainfall fields from the NOAA Climate Forecast System version 2 (CFSv2) and in situ precipitation observations (Funk et al., 2015).

The TC database provides monthly climatic data from 1981 to the near present with a 4 Km spatial resolution at the equator (Table I). The TC datasets were produced using a climatically aided interpolation combining high spatial resolution climatological normals from the WorldClim datasets (Fick and Hijmans, 2017) with coarser resolution but time-varying data from Climate Research Unit time series data version 4.0 (CRU Ts4.0) and the Japanese 55 – year Reanalysis (JRA55). The procedure applies interpolated time-varying anomalies from CRU Ts4.0/JRA55 to the high-spatial-resolution climatology of WorldClim to create a fine-grained dataset that covers a broader temporal record. CHELSA-v2.1 database provides downscaled model output precipitation estimates of the ERA-Interim climatic reanalysis at a

Table I. Characteristics of the three gridded precipitation products (GPPs).

Data	Resolution			Source
	Coverage	Spatial	Temporal	
CHELSA-v2.1	Global	~1km	1979– 2018	(Karger et al., 2017)
CHIRPS-v2.0	Quasi-global (50° N–S)	~5km	1981-Present	(Funk et al., 2015)
TerraClimate	Global	~4km	1958– Present	(Abatzoglou et al., 2018)

very high resolution of 30 arc seconds (approximately 1 Km at the equator) with global coverage (Karger et al., 2017). The precipitation algorithm incorporates orographic predictors, including wind fields, valley exposition, and boundary layer height, with some bias correction. CHELSA provides monthly precipitation data from 1981 to the near present.

2.3 Validation of Gridded Data

A point-to-pixel approach was applied to compare the values of the three GPPs and the rain gauge observations. The values of the GPPs were extracted at the original resolution of each product. Two approaches were employed to assess the skill. First, the Taylor diagram was used to evaluate the correlation, the standard deviation, and the bias of the satellite products compared to the rain gauge data. Taylor diagrams provide a brief statistical summary of how well the patterns match each other in terms of their correlation (r), the centered root-mean-square difference (cRMSE), and the standard deviation (Taylor, 2001). Secondly, the skill of the three GPPs was further assessed using a modified Kling-Gupta Efficiency (KGE; Kling et al., 2012). The KGE was decomposed into three elements: the r representing Person's coefficient of correlation; the β that assesses the bias of the data and γ assesses dispersion of the time series data. The KGE analysis was accomplished using the 'HydroGOF' R package (Zambrano-Bigiarini, 2020). The r value assesses the linear correlation between variables over time (temporal agreement). The β value estimates the ratio between the mean of simulated and observed variables. Therefore $\beta > 1$ indicates an overestimation bias while $\beta < 1$ signifies underestimation bias compared to the reference rain gauge data. The γ values show the variability ratio computed using the standard deviation or the coefficient of variation between simulated and observed variables (variance).

The study area exhibits complex and heterogeneous topography that may affect the accuracy of the GPPs generated from different rain-forming systems (Kimani et al., 2017). Therefore, we assessed the skill of the GPPs at monthly and seasonal temporal scales, i.e., January, February, and March (JFM), April, May, and June (AMJ), July, August, and September (JAS), and October, November, and December (OND). We evaluated the monthly and seasonally aggregated

data over four altitudinal ranges (0 - 500, 501 - 1000, 1001 - 1500, >1500 meters (m) above sea level (a.s.l)). We retrieved elevation data from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) with a 30-meter spatial resolution (METI and NASA, 2011).

The KGE and its three components were also compared for all months (monthly data), separated by four seasons (JFM, AMJ, JAS, OND, and altitudinal ranges using the 'raincloud' plots (Allen et al., 2021). The raincloud plot revealed the variability of KGE and its three decomposition values for the three SPE products over different temporal and altitudinal ranges. The raincloud plots provided a multi-platform tool for robust data visualization that simultaneously presented a combination of the (jittered) raw data points, violin/boxplot, and density plots. The raincloud plots enable a comprehensive visualization of variability in the dataset by providing an overview of raw data, probability distribution, and statistical inference immediately via medians and confidence intervals. They provide users with information on both individual observations and general patterns.

2.4 Trend Analysis

The monthly dataset from the three GPPs was aggregated into four seasons (JFM, AMJ, JAS, and OND) and 38 annual time series. A modified Mann-Kendall statistic (Hamed and Ramachandra Rao, 1998) was used to test the significance ($p < 0.1$) of linear trends in every pixel of the seasonal and annual gridded time series data. A modified Mann-Kendall statistic was selected because it accounts for serial autocorrelation in the time series data. The magnitude of the trend was quantified using Theil-Sen's median slope estimator (Sen, 1968). Trend analysis was accomplished using the 'eco.theilsen' function from the 'EcoGenetics' R package (Roser et al., 2017). The trend analysis for each input seasonal and annual time series produced gridded maps representing Theil-Sen's slope and Mann-Kendall significance test. A similar approach was used to calculate the annual and seasonal precipitation trends from two-rain gauge stations with no data gap, i.e., Morogoro Maji in Tanzania and Msekera in Zambia. This enabled a comparison of trends derived from the rain gauge stations and the three GPPs at annual and seasonal scales.

3. Results

3.1 Validation of gridded data with rain gauge station

All the three GPPs showed high skill in estimating the monthly rain gauge observations ($r > 0.85$, $cRMSE < 0.6$) except at elevations above 1500 masl where accuracy was lower ($r < 0.8$, $cRMSE > 0.6$; Fig. 2a). CHELSA-v2.1 data had the highest skill at 0 – 500 masl altitudinal range ($r = 0.94$, $cRMSE = 0.38$) but the lowest at an altitude above 1500 masl ($r = 0.73$, $cRMSE = 0.76$; Fig. 2a). At seasonal scale, the three

GPPs had a higher skill compared to the monthly scale ($r > 0.95$, $cRMSE < 0.38$), but still the accuracy was lowest for the altitude above 1500 masl (Fig. 2c). Both monthly and seasonally aggregated precipitation data showed that the three GPPs had the lowest skill of estimating precipitation at mountainous terrain above 1500 masl (Fig. 2 - 3). When the monthly precipitation was validated per seasonal blocks, all GPPs showed exceptionally lower skill during OND season for the altitudes above 1500 masl ($0.6 < r < 0.65$; Fig. 3).

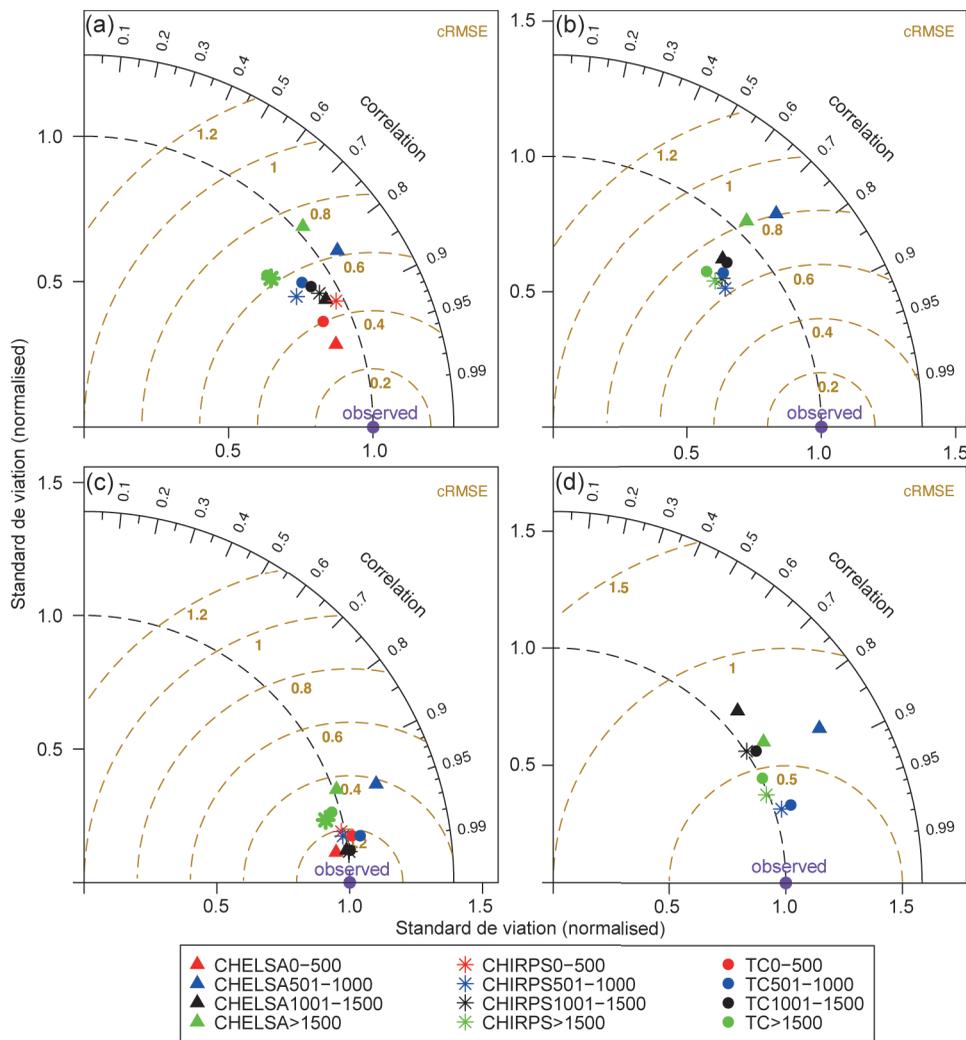


Fig. 2. The skill of monthly (a, b) and seasonally aggregated (c, d) gridded precipitation products with the left and right columns representing the validation with all and with only the independent rain gauge stations. The legend labels represent a combination of the gridded precipitation products from TerraClimate, CHELSA-v2.1, and CHIRPS-v2 databases at four altitudinal ranges (masl) in the East and Southern Africa (ESA) region.

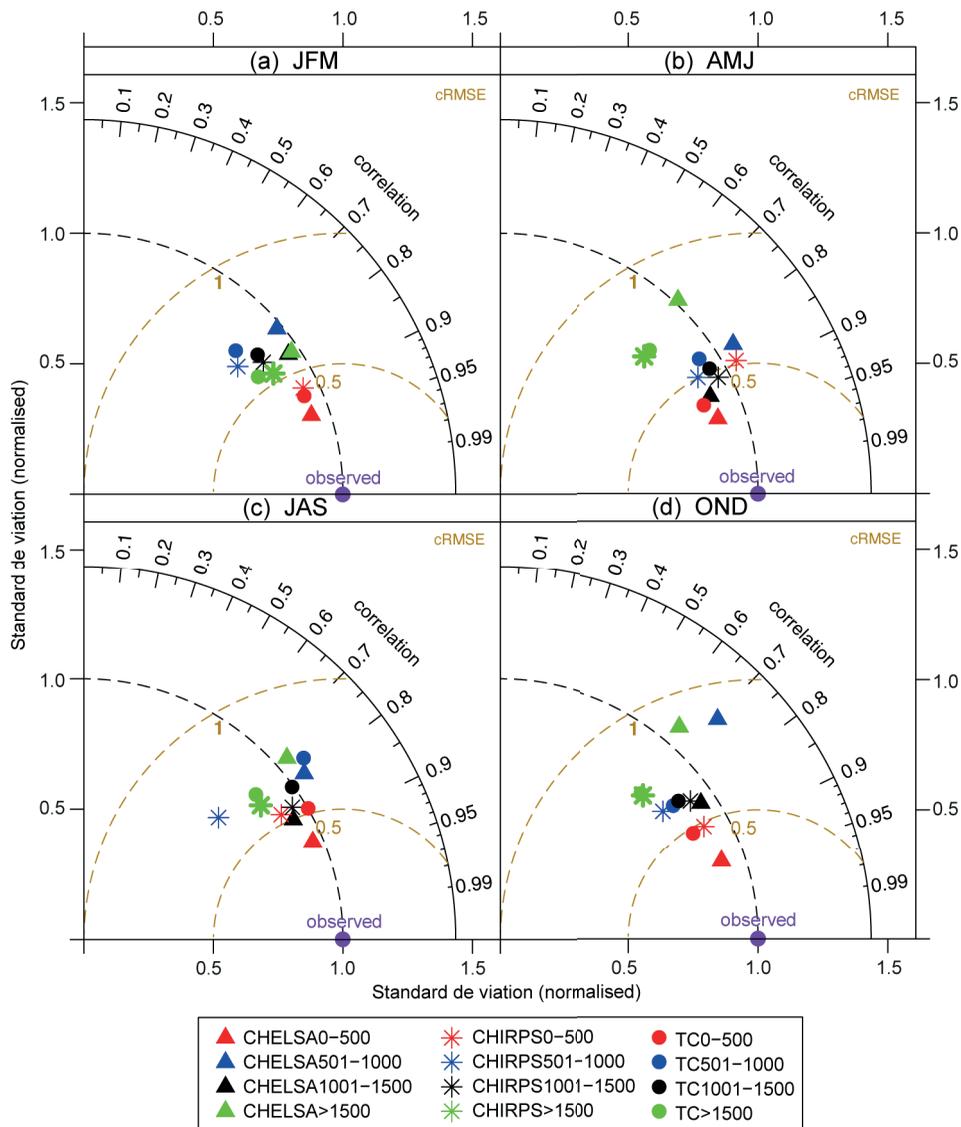


Fig. 3. The skill of gridded precipitation estimates (GPE) from CHELSA-v2.1, CHIRPS-v2, and TerraClimate databases at four altitudinal levels (0-500, 501-1000, 1001-1500, and above 1500 m) during JFM (a), AMJ (b), JAS (c) and OND (d) seasons. The legend labels represent a combination of one satellite product and the altitudinal range in masl. All rain gauge stations are used.

The skill of the three GPPs reduced when validated with the independent stations for both monthly ($r > 0.7$, $cRMSE < 0.6$; Fig. 2b) and seasonal scales ($r > 0.75$, $cRMSE > 0.4$; Fig. 2d), but the CHELSA-v2.1 generally showed the lowest accuracy above 1500 masl.

Raincloud plots revealed the variability of KGE and its three decomposition values for the three GPPs at different temporal (monthly and seasonal) and

altitudinal scales when observations from all stations (Fig. 4) and only 18 independent rain gauge stations (Fig. 5) are used. The raincloud plots for evaluation with all 82 rain gauge stations showed the median KGE for the three products was above 0.8 except for 501-1000 masl and above 1500 masl altitude. The KGE was remarkably lower during the JAS and OND seasons at 501-1000 masl and above 1500 masl altitudes (Fig. 4a). Considering the KGE and its

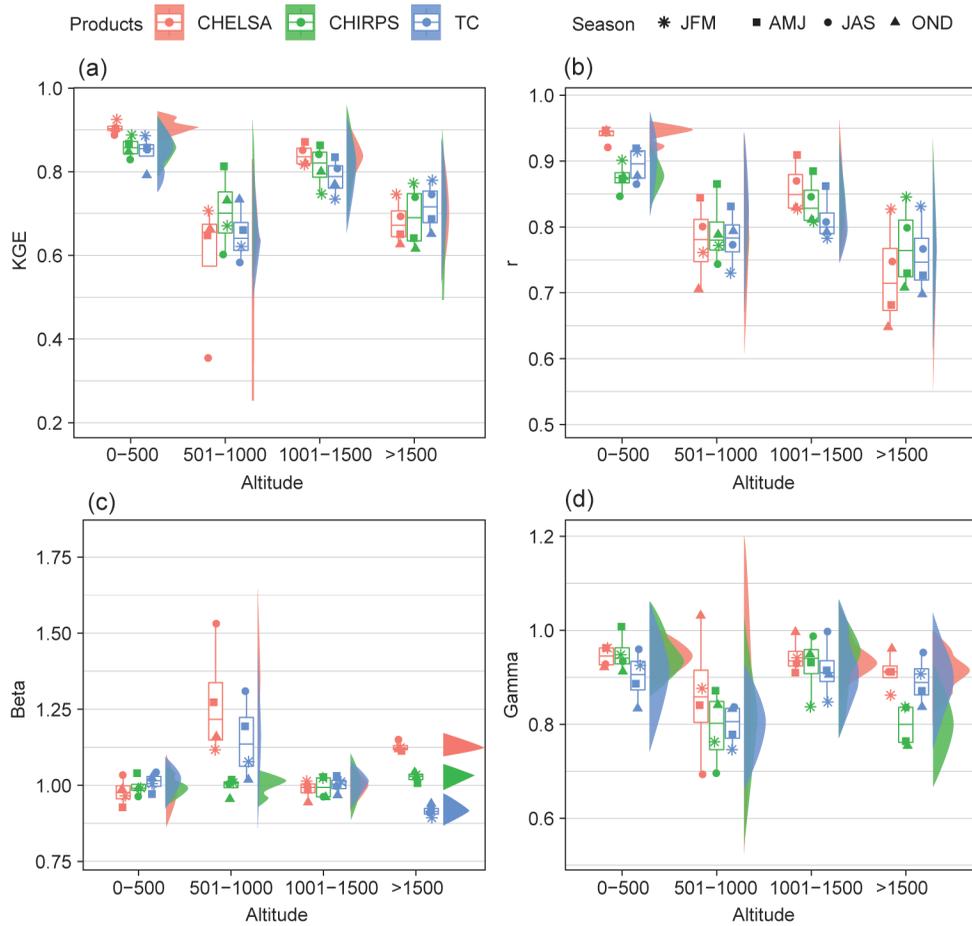


Fig. 4. Raincloud plots showing the Kling-Gupta Efficiency (KGE) values and its three decomposition values (correlation (r), bias (β), and dispersion (γ), which reflect the skill of the monthly gridded datasets from CHELSA, CHIRPS-v2 and TerraClimate (TC) to estimate all the rain gauge observations at four altitudinal levels (0 - 500, 501-1000, 1001-1500, and above 1500 masl). The satellites and seasons are labeled with different colors and shapes, respectively.

three decomposition values, the three GPPs showed relatively similar skill at 1001 – 1500 masl altitudinal range (Fig. 4a - d. Focusing on the individual KGE components, the temporal agreement (r) for all three GPPs was the lowest at the altitudinal range above 1500 masl ($0.72 > r < 0.76$) but distinctively much lower during the OND season (Fig. 4b), signifying higher temporal mismatch compared to the observation rain gauge. This indicates that lower temporal agreement during the OND season at elevations above 1500 masl (Fig. 4b) had the strongest influence on overall KGE (Fig. 4a). The CHELSA-v2.1 data showed the highest median over-estimation bias at

the 501 – 1000 ($\beta= 1.24$); Fig. 4c) and at >1500 masl altitudinal ranges ($\beta= 1.125$); Fig. 4c), and this bias was highest during the JAS season. At altitude > 1500 masl, the individual KGE accuracy measures showed that CHELSA-v2.1 had the lowest median temporal agreement ($r = 0.72$; Fig. 4b), highest overestimation bias ($\beta = 1.125$; Fig. 4c) though it estimated the dispersion more accurately ($\alpha = 0.92$; Fig. 4d) compared to the other two products.

None of the 18 independent stations were located below 500 masl of altitude (Table SI), so the performance at that range could not be independently verified. The KGE values derived from the evaluation

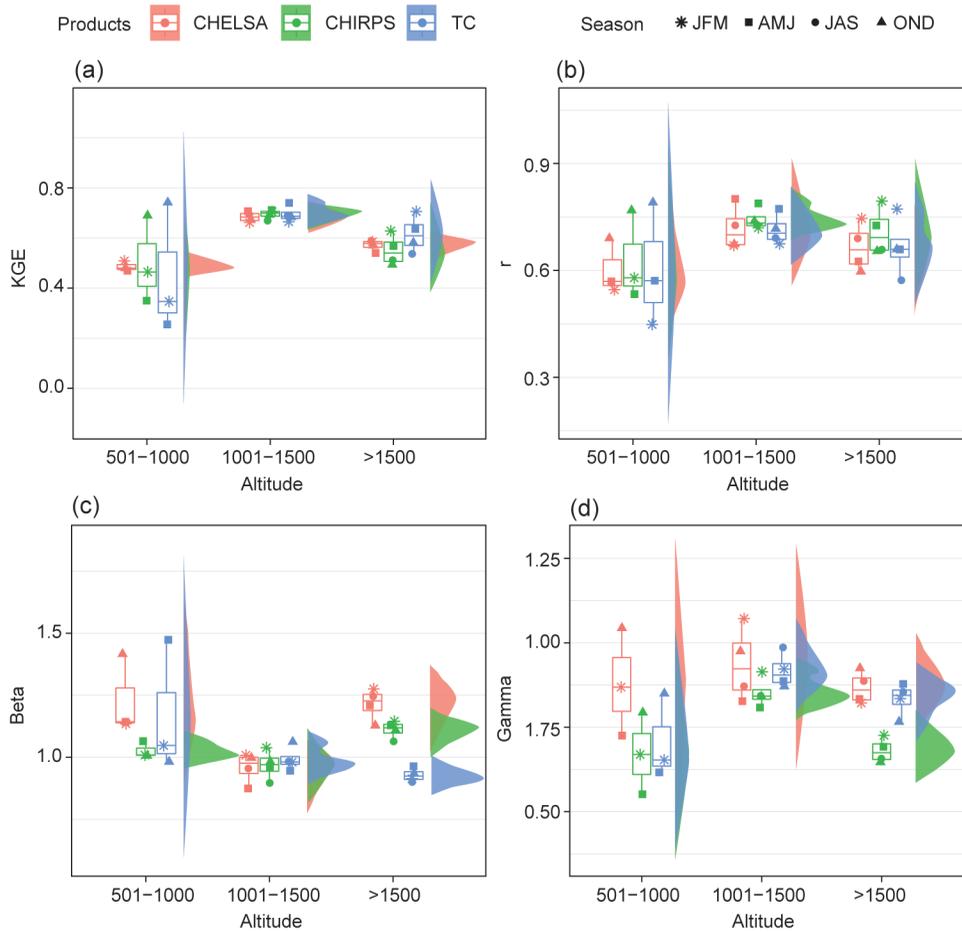


Fig. 5. Raincloud plots showing the Kling-Gupta Efficiency (KGE) values and its three decomposition values (correlation (r), bias (β), and dispersion (γ), which reflects the skill of the monthly gridded datasets from CHELSA, CHIRPS-v2 and TerraClimate (TC) to estimate observations from 18 independent rain gauge stations at four altitudinal levels (0 - 500, 501-1000, 1001-1500, and above 1500 masl). The satellites and seasons are labeled with different colors and shapes, respectively.

with the independent stations only (Fig 5) were consistently lower than when using all the rain gauge stations (Fig 4). The skill of the three GPPs, revealed by KGE and its decomposition values, was consistently the lowest on the 501-1000 masl altitudinal interval. The three GPPs showed relatively similar temporal agreement (Fig 5b) across the altitudinal ranges. Like results with all the stations, the independent observations showed that the CHELSA-v2.1 had the highest over-estimation bias at 501-1000 masl and >1500 masl altitudinal ranges (Fig 5c). Still, CHELSA-v2 outperformed the other two products in replicating the dispersion of rainfall over space and time (Fig 5d).

3.2 Annual precipitation trends

The CHELSA-v2.1 dataset overestimated the annual and seasonal precipitation trends over mountainous terrain over 1500 m.a.s.l and large inland waterbodies, e.g., Lake Victoria and Malawi (Fig. 6a, Fig. S1). Even at lower altitudes, the precipitation trends estimated from CHELSA-v2.1 were the highest compared with the CHIRPS-V2 and TC data. The precipitation trends derived from CHELSA-v2.1 data are overestimated. All three GPP products showed a significant increase in annual precipitation across southern and western Zambia, with differing magnitude (1 – 14 mm year⁻¹)

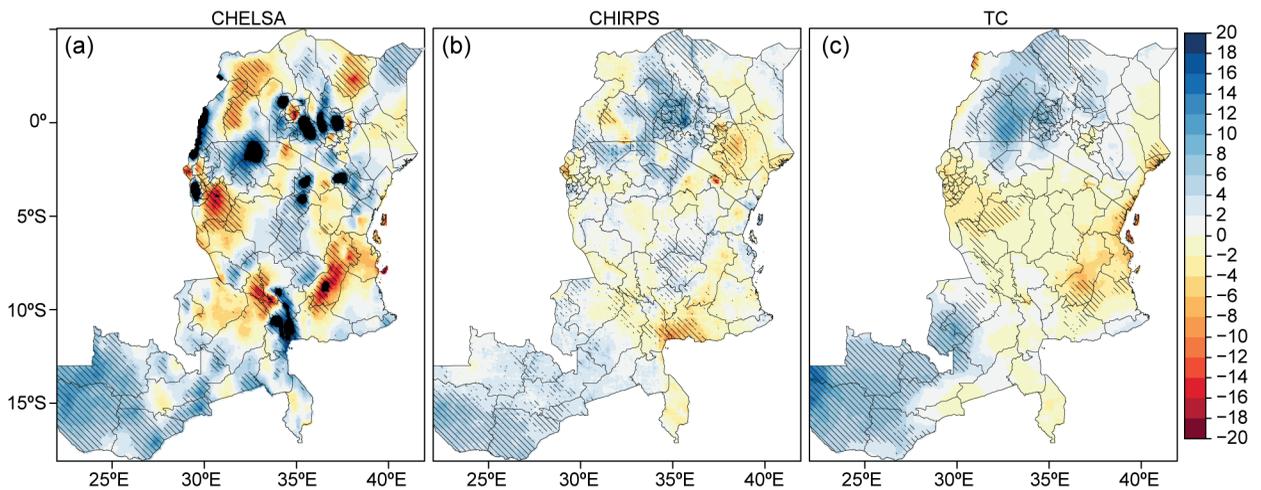


Fig. 6. The annual precipitation trends (mm year⁻¹) over 38 years (1981 – 2018) derived from (a) CHELSA-v2.1, (b) CHIRPS-v2, and (c) TerraClimate (TC) products over Eastern and Southern Africa (ESA) region. The hatched lines represent areas with significant trends. The black-shaded areas over the trends derived from the CHELSA dataset (a) mask out areas with unrealistic values, presented separately in Fig. S1.

and extent (Fig. 6). The CHIRPS-v2 and TC dataset revealed a significant annual precipitation increase in transboundary Lake Victoria North basin in Western Kenya and Eastern Uganda. Still, TC data showed a larger area with a significant wetting that extended further to Central Uganda and North-western Kenya. Generally, the trends derived from CHIRPS-v2 and TC showed substantial convergence except in Tanzania (Fig. 6b-c). In Tanzania, CHIRPS-v2 showed a significant wetting trend covering parts of Dodoma, Iringa, Singinda, Kagera, Mara, and Arusha regions (Fig. 6b). However, the TC revealed a drying trend across Tanzania that was significant in a portion of Tabora, Katavi, Morogoro, Lindi, Ruvuma, Pwani and Tanga regions (Fig. 6c).

CHIRPS-v2 and TC datasets showed a drying trend in southern Lamu and Garissa counties along the Kenyan coast (Fig. 6b - c). CHIRPS-v2 revealed a drying trend over south-eastern Kenya (Kitui, Makueni, Machakos, Tharaka-Nithi counties), but the TC data returned contrasting results in that area (Fig. 6b-c). CHIRPS-v2 data showed a high magnitude drying trend (14 – 20 mm year⁻¹) over Mount Kilimanjaro in Tanzania (Fig. 6b) that could be linked to systematic biases over the mountainous terrain, presented in Figures 2 – 4.

3.3 Seasonal precipitation trends

Seasonal analysis revealed that all the SPEs recorded a significant wetting trend (2 – 8 mm Season⁻¹ year⁻¹) in southwestern Zambia during the JFM (Fig. 7a, e, i) and OND seasons (Fig. 7d, h, l). Similarly, during the OND season, all the SPE products showed a low-magnitude wetting trend with a differing spatial extent in Kenya (Fig. 7d, h, l). But still, CHELSA data returned unrealistic high trends (Fig. 7d). A drying trend dominated in the AMJ and JAS seasons across the region but with varying spatial extent (Fig. 7b, c, f, g, j). The CHIRPS-v2 data showed a drying trend covering almost the entire ESA region during the JAS season (Fig. 7g). TC data captured a peak drying trend (0 – 6 mm Season⁻¹ year⁻¹) during the AMJ season along the southern Lake Victoria basin in Tanzania, Rwanda, and Burundi and along the east African coastline (Fig. 7j).

3.4 Comparing precipitation trends derived from the rain gauge stations and gridded products

Comparing the precipitation trends derived from the rain gauge stations and GPPs revealed diverse performance over space (stations) and time (seasons; Table II). At Morogoro Maji station, all the GPPs overestimated the annual drying trend compared to the trend derived from the rain gauge stations (Table II). At Msekera station, all the GPPs returned

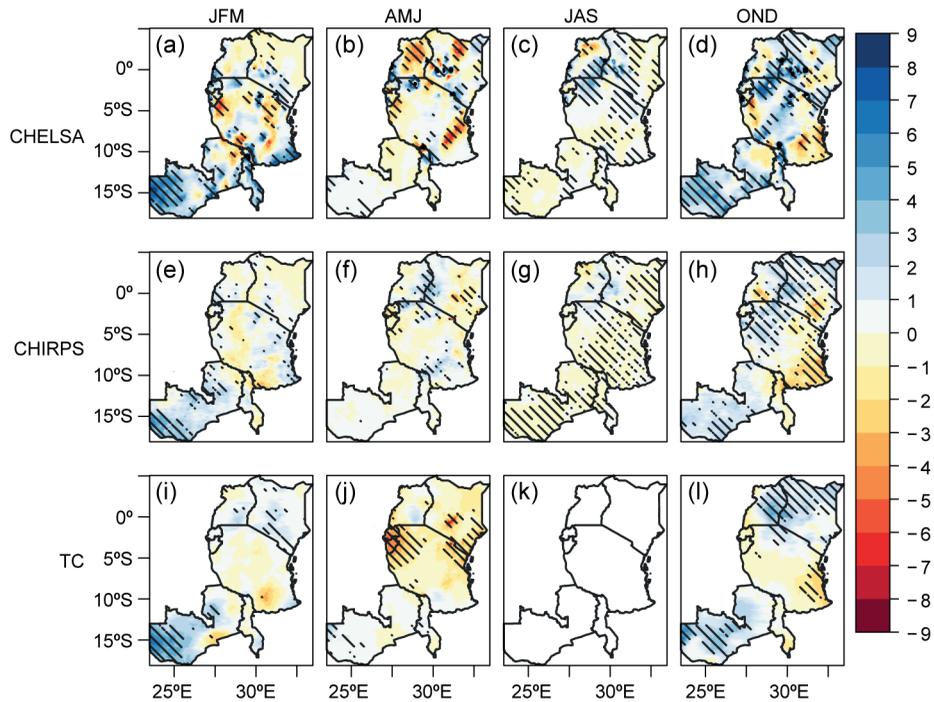


Fig. 7. Seasonal precipitation trends ($\text{mm Season}^{-1} \text{ year}^{-1}$) for 38 years (1981 – 2018) derived from CHELSA-v2.1 (a-d), CHIRPS-v2 (e - h) and TerraClimate (TC; i - l) products over Eastern and Southern Africa (ESA) region. The hatched lines represent areas with significant trends.

Table II. Comparison of annual and seasonal precipitation trends derived from rain gauge stations and the gridded precipitation product.

Station	Season	Rain Gauge	CHELSA-v2.1	CHIRPS-v2	TC
Morogoro Maji	Annual	-0.100	-1.952	-1.830	-3.317
	JFM	1.771	2.350	0.737	0.052
	AMJ	0.092	-2.249	-0.733	-1.501
	JAS	-0.013	0.075	-0.186	
	OND	-2.188	0.604	-1.863	-1.355
Msekera	Annual	5.589	8.672	2.465	1.065
	JFM	6.875	4.423	0.729	0.646
	AMJ	0.243	-0.103	-0.075	0.207
	JAS	0.000	-0.003	-0.004	
	OND	0.840	5.505	0.583	0.029

a positive annual trend of different magnitudes. In both stations during the AMJ season, all the GPPs generally showed the opposite direction of the trend compared to the one from the rain gauge stations. During OND season in both stations, CHELSA-v2.1 returned inconsistent trends compared to the rain gauge trend, but the other two GPPs showed substantial agreement.

4. Discussion

4.1 Validation of gridded precipitation products

Validation results revealed that the three GPPs had varied performance over temporal and altitudinal ranges. The three monthly GPPs data skill was generally high but performed less at elevations over 1500 masl, especially during the OND season. At mountainous terrain above 1500 masl, CHELSA-v2.1

had lower performance than the other two products but outperformed others at 0-500 masl elevation. The KGE results showed that at altitudes over 1500 masl, CHELSA-v2.1 had the lowest temporal agreement with the rain gauge (r), which could result from temporal changes in the density of observation data. At mountainous terrain above 1500 masl, the GPPs could have failed to capture the orographic rain-forming process due to poor calibration since precipitation gauge networks are sparser in high elevations with limited accessibility (Lundquist et al., 2019). This phenomenon significantly affected CHELSA-v2.1 data, suggesting that its reanalysis model failed to capture orographic processes and air masses movements over mountainous terrain and inland lakes. Even when the CHELSA-v2.1 dataset was expected to perform better over mountainous landscapes since the precipitation algorithm incorporated orographic predictors such as wind fields, valley exposition, and boundary layer height with subsequent bias correction (Karger et al., 2017). Similarly, Bobrowski et al. (2021) reported that CHELSA-v2.1 data presented distorted precipitation amounts over the Himalayas mountains in Nepal, possibly due to interpolation of station biases in areas with low density of stations in windward or leeward side (Bobrowski et al., 2021). Kimani et al. (2017) reported underestimation bias from CHIRPS-v2 at high elevations in East Africa, especially during the OND season, attributed to challenges in capturing orographic precipitation due to poor rain gauge distribution. Estimating orographic rainfall is a substantial challenge for many satellite products (Diem et al., 2014). Therefore, adding more station data can improve the rainfall estimates in high elevations.

Although the station data are not a direct input to the downscaling model for generating the TC product, the product is derived from the WorldClim data (Fick and Hijmans, 2017) that was generated by interpolating the station data. The biases of the TC data at elevations above 1500 masl could be error propagation from input data to the retrieval algorithm. Our results emphasize the importance of validating climate datasets to avoid error propagation in different models and applications. Our results reveal that the three algorithms, particularly the CHELSA-v2.1, need further calibration over the ESA region's mountainous and inland water bodies. The CHELSA-v2.1

data could be improved by replacing ERA-Interim with ERA-5 in the downscaling algorithm, as the former has shown a substantial reduction of precipitation bias in the ESA region (Gleixner et al., 2020).

The independent evaluation showed slightly lower accuracy than all stations, including those ingested in the original algorithm. Interestingly, both the non-independent and the independent assessments consistently revealed that CHELSA-v2.1 had the highest over-estimation bias but outperformed the other two GPPs in capturing the dispersion of rainfall. The lower skill in the independent evaluation could be linked to several factors. First, biases were significantly reduced at the locations where data from the non-independent stations were ingested in the retrieval algorithms of the GPPs, resulting in an improved agreement between observations and the GPPs. Also, it could result from little or low data quality in the independent stations. As noted by Dinku et al. (2018), most of the independent rain gauge stations are maintained by volunteers; therefore, their quality may not be as good as those collected by professional meteorologists. The evaluation with independent rain gauge stations is inconclusive, given the few stations. Further independent assessment is recommended once more observation datasets are available.

The paucity of rain gauge data in Africa can be addressed by increased investments in automatic weather stations like the initiative promoted by the Trans-African Hydro-Meteorological Observatory (TAHMO; van de Giesen et al., 2014) that plan to install over 20000 stations across sub-Saharan Africa. Moreover, the Enhancing National Climate Services (ENACTS; Dinku, 2019) initiative has dedicated efforts to creating capacity and tools for data quality control.

4.2 Trend Analysis of Climatic Variables

The CHELSA-v2.1 dataset showed unrealistic trends in mountainous terrain located over 1500 masl and large inland waterbodies due to the overestimation biases observed in the KGE accuracy assessment. Therefore, annual precipitation trends derived from CHELSA-v2.1 data are unreliable and should be interpreted cautiously. Generally, the skill and the precipitation trends derived from the CHIRPS-v2 and TC datasets in the ESA region closely matched (except in Tanzania), revealing the potential of the statistical

downscaling method in estimating precipitation over areas with low density of rain gauge stations like the ESA region. However, in West Africa, the monthly CHIRPS-v2 was more accurate than TC (Muthoni, 2020). Therefore, the skill of the GPPs varies over space and time.

The coarser resolution ERA-Interim and ERA-5 reportedly produced inconsistent precipitation trends in East Africa compared to CHIRPS-v2 data (Gleixner et al., 2020). This led to the conclusion that reanalysis data is unsuitable for trend analysis. The reanalysis data inherits characteristics of the parent dataset, i.e., ERA-Interim and CRUTs4.0 for CHELSA-v2.1 and TC, respectively. Therefore they may not capture the temporal variability in orographic precipitation ratios and inversions at finer scales than their parent datasets (Abatzoglou et al., 2018). This limitation can partly explain the difference between the trends derived from the rain gauge stations and GPPs over space and time. However, in a recent evaluation of the three GPPs over mountainous terrain in Java, the TC emerged as the most accurate and reliable in representing the temporal dynamics of the precipitation compared to the CHIRPS-v2 and CHELSA-v2.1 (Dumont et al., 2022). Similarly, our results suggest remarkable improvements in TC's statistical downscaling. Nonetheless, the trends generated by TC and CHIRPS-v2 over Tanzania showed substantial differences. TC data showed a drying trend over Tanzania that was significant in small portions, but CHIRPS-v2 showed a low magnitude but statistically significant wetting trend over the Kagera, Mara, Arusha, Shinyanga, Dodoma, Singinda, and Iringa regions of Tanzania. The opposite direction of the trends between the two products in Tanzania warrants further explanation.

Generally, CHIRPS-V2 produced low-magnitude trends compared to the MRA products. This could be linked to the CHIRP climatology, which has low bias. The CHIRP data set is calculated as a percent anomaly multiplied by the climatological mean. When this mean is low, the precipitation estimated by the CHIRPS-v2 algorithm will almost always be low (Harrison et al., 2019). This bias correction is known to reduce the rainfall variance in the CHIRPS, especially in drier locations, and this may explain the low magnitude trends from CHIRPS-v2 than in

the other two products. Considering this aspect, the trends estimated from CHIRPS-V2, especially in drylands, could be regarded as conservative.

Our results show a convergence of evidence suggesting wetting trends over Southern and Western Zambia for annual and seasonal precipitation. These wetting trends occur largely in rangelands with annual rainfall ranging between 500 – 1000 mm (Fig 1a). Muthoni et al. (2019) also observed a similar annual precipitation trend in the same area using the CHIRPS-v2 dataset. The significant wetting in Zambia was observed during OND and JFM seasons that coincide with the growing season, suggesting reduced soil moisture stress in the area that can boast crop yields. The wetting trends in Western and Southern regions of Zambia occur in rangelands with annual rainfall ranging between 500 – 1000 mm (Fig 1a); therefore, the wetting trends can reduce moisture stress. However, a large part of the Lake Victoria north basin experiences annual average rainfall over 1500 mm (Fig 1a); hence, further increase in moisture can exacerbate flooding.

5. Conclusions

Our analysis evaluated the skill of three gridded precipitation products (GPP) with varying spatial resolution in estimating the rain gauge station network observations and compared the long-term precipitation trends derived from these products. Validation results revealed that the three GPPs had varied performance over temporal and altitudinal ranges. CHELSA-v2.1 had a lower skill than the other two products. Our results demonstrate that estimating orographic rainfall remains challenging for the GPPs. The importance of validating climate datasets is emphasized to avoid error propagation in different models and applications. Our results further demonstrate that new or higher-resolution precipitation data is not always the most accurate since an algorithm update can introduce artifacts or biases. Results elucidate the strengths and shortcomings of the three gridded precipitation products over time and space to guide their application in different contexts and sectors. There is a convergence of evidence on decreasing moisture stress in Zambia that can boost crop productivity in the predominantly rainfed farming system.

Acknowledgments

USAID funded this study through grant number: AID-BFS-G-11-00002 under the Feed the Future initiative to support the Africa RISING program—special thanks to many farmers, ranchers, and volunteers that provided the rainfall data.

References

- Abatzoglou JT, Dobrowski SZ, Parks SA, Hegewisch KC. 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data* 5: 170191. <https://doi.org/10.1038/sdata.2017.191>
- Adhikari U, Nejadhashemi AP, Woznicki SA. 2015. Climate change and eastern Africa: a review of impact on major crops. *Food and Energy Security* 4: 110-132. <https://doi.org/10.1002/fes3.61>
- Allen M, Poggiali D, Whitaker K, Marshall TR, van Langen J, Kievit RA. 2021. Raincloud plots: a multi-platform tool for robust data visualization [version 2; peer review: 2 approved]. *Wellcome Open Research* 4:63. <https://doi.org/10.12688/wellcomeopenres.15191.2>
- Atiah WA, Muthoni FK, Kotu B, Kizito F, Amekudzi LK. 2021. Trends of Rainfall Onset, Cessation, and Length of Growing Season in Northern Ghana: Comparing the Rain Gauge, Satellite, and Farmer's Perceptions. *Atmosphere* 12: 1674. <https://doi.org/10.3390/atmos12121674>
- Awange JL, Hu KX, Khaki M. 2019. The newly merged satellite remotely sensed, gauge and reanalysis-based Multi-Source Weighted-Ensemble Precipitation: Evaluation over Australia and Africa (1981–2016). *Science of The Total Environment* 670: 448-465. <https://doi.org/10.1016/j.scitotenv.2019.03.148>
- Bengtsson L, Hagemann S, Hodges KI. 2004. Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research: Atmospheres* 109: D11111. <https://doi.org/10.1029/2004JD004536>
- Bobrowski M, Weidinger J, Schickhoff U. 2021. Is New Always Better? *Frontiers in Global Climate Datasets for Modeling Treeline Species in the Himalayas*. *Atmosphere* 12: 543. <https://doi.org/10.3390/atmos12050543>
- Cairns JE, Hellin J, Sonder K, Araus JL, MacRobert JF, Thierfelder C, Prasanna BM. 2013. Adapting maize production to climate change in sub-Saharan Africa. *Food Security* 5: 345-360. <https://doi.org/10.1007/s12571-013-0256-x>
- Cattani E, Merino A, Guijarro J, Levizzani V. 2018. East Africa Rainfall Trends and Variability 1983–2015 Using Three Long-Term Satellite Products. *Remote Sensing* 10: 931. <https://doi.org/10.3390/rs10060931>
- Contractor S, Donat MG, Alexander LV, Ziese M, Meyer-Christoffer A, Schneider U, Rustemeier E, Becker A, Durre I, Vose RS. 2020. Rainfall Estimates on a Gridded Network (REGEN)—a global land-based gridded dataset of daily precipitation from 1950 to 2016. *Hydrology and Earth System Sciences* 24: 919-943. <https://doi.org/10.5194/hess-24-919-2020>
- Dee DP, Källén E, Simmons AJ, Haimberger L. 2011. Comments on “Reanalyses Suitable for Characterizing Long-Term Trends”. *Bulletin of the American Meteorological Society* 92: 65-70. <https://doi.org/10.1175/2010BAMS3070.1>
- Diem JE, Ryan SJ, Hartter J, Palace MW. 2014. Satellite-based rainfall data reveal a recent drying trend in central equatorial Africa. *Climatic Change* 126: 263-272. <https://doi.org/10.1007/s10584-014-1217-x>
- Dinku T. 2019. Challenges with availability and quality of climate data in Africa. In: Melesse AM, Abtew W, Senay G, eds. *Extreme Hydrology and Climate Variability: Monitoring, Modelling, Adaptation and Mitigation*. Elsevier, pp. 71-80, <https://doi.org/10.1016/B978-0-12-815998-9.00007-5>
- Dinku T, Ceccato P, Grover-Kopec E, Lemma M, Connor SJ, Ropelewski CF. 2007. Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing* 28: 1503-1526. <https://doi.org/10.1080/01431160600954688>
- Dinku T, Funk C, Peterson P, Maidment R, Tadesse T, Gadain H, Ceccato P. 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society* 144: 292-312. <https://doi.org/10.1002/qj.3244>
- Dumont M, Saadi M, Oudin L, Lachassagne P, Nugraha B, Fadillah A, Bonjour JL, Muhammad A, Hendarmawan, Dörfli N, Plagnes V. 2022. Assessing rainfall global products reliability for water resource management in a tropical volcanic mountainous catchment. *Journal of Hydrology: Regional Studies* 40: 101037. <https://doi.org/10.1016/j.ejrh.2022.101037>
- Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37: 4302-4315. <https://doi.org/10.1002/joc.5086>

- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J. 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* 2: 150066. <https://doi.org/10.1038/sdata.2015.66>
- Gleixner S, Demissie T, Diro GT. 2020. Did ERA5 Improve Temperature and Precipitation Reanalysis over East Africa? *Atmosphere* 11: 9. <https://doi.org/10.3390/atmos11090996>
- Haghtalab N, Moore N, Ngongondo C. 2019. Spatio-temporal analysis of rainfall variability and seasonality in Malawi. *Regional Environmental Change* 19: 2041–2054. <https://doi.org/10.1007/s10113-019-01535-2>
- Hamed KH, Ramachandra Rao A. 1998. A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology* 204: 182-196. [https://doi.org/10.1016/S0022-1694\(97\)00125-X](https://doi.org/10.1016/S0022-1694(97)00125-X)
- Harrison L, Funk C, Peterson P. 2019. Identifying changing precipitation extremes in Sub-Saharan Africa with gauge and satellite products. *Environmental Research Letters* 14: 085007. <https://doi.org/10.1088/1748-9326/ab2cae>
- Henn B, Newman AJ, Livneh B, Daly C, Lundquist JD. 2018. An assessment of differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology* 556: 1205-1219. <https://doi.org/10.1016/j.jhydrol.2017.03.008>
- Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut J-N. 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* 146: 1999-2049. <https://doi.org/10.1002/qj.3803>
- Karger DN, Conrad O, Böhrner J, Kawohl T, Kreft H, Soria-Auza RW, Zimmermann NE, Linder HP, Kessler M. 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific Data* 4: 170122. <https://doi.org/10.1038/sdata.2017.122>
- Kimani WM, Hoedjes JCB, Su Z. 2017. An Assessment of Satellite-Derived Rainfall Products Relative to Ground Observations over East Africa. *Remote Sensing* 9: 430. <https://doi.org/10.3390/rs9050430>
- Kimunye JN, Were E, Mussa F, Tazuba A, Jomanga K, Viljoen A, Swennen R, Muthoni FK, Mahuku G. 2020. Distribution of *Pseudocercospora* species causing Sigatoka leaf diseases of banana in Uganda and Tanzania. *Plant Pathology* 69: 50-59. <https://doi.org/10.1111/ppa.13105>
- Kling H, Fuchs M, Paulin M. 2012. Runoff conditions in the upper danube basin under an ensemble of climate change scenarios. *Journal of Hydrology* 424-425: 264-277. <https://doi.org/10.1016/j.jhydrol.2012.01.011>
- Lawrimore JH, Ray R, Applequist S, Korzeniewski B, Menne MJ. 2016. Global Summary of the Month (GSOM), Version 1. Available at <https://www.ncdc.noaa.gov/cdo-web/search?datasetid=GSOM> (accessed 2019 July 15)
- Lemma E, Upadhyaya S, Ramsankaran R. 2019. Investigating the performance of satellite and reanalysis rainfall products at monthly timescales across different rainfall regimes of Ethiopia. *International Journal of Remote Sensing* 40: 4019-4042. <https://doi.org/10.1080/01431161.2018.1558373>
- Lundquist J, Hughes M, Gutmann E, Kapnick S. 2019. Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. *Bulletin of the American Meteorological Society* 100: 2473-2490. <https://doi.org/10.1175/BAMS-D-19-0001.1>
- METI, NASA. 2011. ASTER Global Digital Elevation Model (ASTER GDEM) version 2. Ministry of Economy, Trade, and Industry, United States National Aeronautics and Space Administration. Available at <http://www.jspacesystems.or.jp/ersdac/GDEM/E/4.html> (accessed 2019 October 29)
- Mkonda MY, He X. 2018. Climate variability and crop yields synergies in Tanzania's semiarid agroecological zone. *Ecosystem Health and Sustainability* 4: 59-72. <https://doi.org/10.1080/20964129.2018.1459868>
- Mumo L, Yu J, Ojara M, Lukorito C, Kerandi N. 2021. Assessing changes in climate suitability and yields of maize and sorghum crops over Kenya in the twenty-first century. *Theoretical and Applied Climatology* 146: 381-394. <https://doi.org/10.1007/s00704-021-03718-6>
- Muthoni FK. 2020. Spatial-Temporal Trends of Rainfall, Maximum and Minimum Temperatures Over West Africa. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13: 2960-2973 <https://doi.org/10.1109/JSTARS.2020.2997075>

- Muthoni FK, Odongo VO, Ochieng J, Mugalavai EM, Mourice SK, Hoesche-Zeledon I, Mwila M, Bekunda M. 2019. Long-term spatial-temporal trends and variability of rainfall over Eastern and Southern Africa. *Theoretical and Applied Climatology* 137: 1869-1882. <https://doi.org/10.1007/s00704-018-2712-1>
- Niassy S, Agbodzavu MK, Kimathi E, Mutune B, Abdel-Rahman EFM, Salifu D, Hailu G, Belayneh YT, Felege E, Tonnang HEZ, Ekesi S, Subramanian S. 2021. Bioecology of fall armyworm *Spodoptera frugiperda* (J. E. Smith), its management and potential patterns of seasonal spread in Africa. *PLoS ONE* 16: e0249042. <https://doi.org/10.1371/journal.pone.0249042>
- Novella NS, Thiaw WM. 2013. African Rainfall Climatology Version 2 for Famine Early Warning Systems. *Journal of Applied Meteorology and Climatology* 52: 588-606. <https://doi.org/10.1175/JAMC-D-11-0238.1>
- Omoyo NN, Wakhungu J, Oteng'i S. 2015. Effects of climate variability on maize yield in the arid and semi arid lands of lower eastern Kenya. *Agriculture & Food Security* 4: 8. <https://doi.org/10.1186/s40066-015-0028-2>
- Roser LG, Ferreyra LI, Saidman BO, Vilardi JC. 2017. EcoGenetics: An R package for the management and exploratory analysis of spatial data in landscape genetics. *Molecular Ecology Resources* 17: e241-e250. <https://doi.org/10.1111/1755-0998.12697>
- Sen PK. 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association* 63: 1379 -1389. <https://doi.org/10.1080/01621459.1968.10480934>
- Seregina LS, Fink AH, van der Linden R, Elagib NA, Pinto JG. 2019. A new and flexible rainy season definition: Validation for the Greater Horn of Africa and application to rainfall trends. *International Journal of Climatology* 39: 989-1012. <https://doi.org/10.1002/joc.5856>
- Taylor KE. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres* 106: 7183-7192. <https://doi.org/10.1029/2000JD900719>
- Thorne PW, Vose RS. 2010. Reanalyses Suitable for Characterizing Long-Term Trends. *Bulletin of the American Meteorological Society* 91: 353-362. <https://doi.org/10.1175/2009BAMS2858.1>
- van de Giesen N, Hut R, Selker J. 2014. The Trans-African Hydro-Meteorological Observatory (TAHMO). *WIREs Water* 1: 341-348. <https://doi.org/10.1002/wat2.1034>
- Zambrano-Bigiarini M. 2020. hydroGOF: Goodness-of-Fit Functions for Comparison of Simulated and Observed Hydrological Time Series. R Package Version 0.4-0. <https://doi.org/10.5281/zenodo.839854>

SUPPLEMENTARY MATERIAL

Table SI. List of rain gauge stations and data gaps. The bolded rows represent the independent stations not ingested in initial algorithms for gridded products.

ID	Country	Station	Lat	Long	Altitude	Start	End	Counts (Months)	Data Gap (%)
1		Alupe	0.470	34.120	1179	1981	2004	288	36.8
2		Bomet	-0.780	35.330	1926	1981	2000	240	47.4
3		Bungoma Water Service	0.580	34.570	1409	1981	2004	288	36.8
4		Butere	0.280	34.500	1294	1981	1991	132	71.1
5		Crescent Island Naivasha	-0.775	36.407	1895	1981	2010	360	21.1
6		Garissa	-0.467	39.633	136	1981	2003	188	58.8
7		Gilgil Kwetu farm	-0.344	36.303	2376	1981	2010	357	21.7
8	Kenya	Homabay	-0.530	34.470	1195	1981	2004	288	36.8
9		Jomo Kenyatta Inter. Airport	-1.317	36.917	1622	1981	2018	209	54.2
10		KARI Naivasha	-0.690	36.402	1904	1981	2010	360	21.1
11		Kijabe Naivasha	-0.817	36.267	1883	1983	2010	336	26.3
12		Kitale	1.016	35.000	1882	1981	2018	117	74.3
13		Mandera	3.933	41.867	223	1981	1998	203	55.5
14		Mombasa	-4.033	39.617	56	1981	2008	215	52.9
15		Tinderet Tea	0.020	35.350	1775	1981	2004	288	36.8
16		Bolero	-10.967	33.733	1107	1981	1990	110	75.9
17		Bvumbwe	-15.917	35.067	1147	1981	1990	112	75.4
18		Chichiri	-15.783	35.033	1099	1981	1990	109	76.1
19		Chileka	-15.683	34.967	762	1981	2017	118	74.1
20		Chinguluwe	-13.690	34.240	647	2008	2018	82	82
21		Chipeni	-13.791	34.056	1077	2005	2018	91	80
22		Chitedze	-13.983	33.633	1152	1981	1990	113	75.2
23		Chitipa	-9.700	33.267	1279	1981	1990	108	76.3
24		Dedza	-14.317	34.267	1687	1981	1990	112	75.4
25		Herbert	-14.886	35.036	642	2007	2018	83	81.8
26		Karonga	-9.950	33.883	539	1981	1990	111	75.7
27		Lemu	-14.785	35.024	669	2006	2018	97	78.7
28		Linga	-12.800	34.200	531	2008	2018	79	82.7
29	Malawi	Makanga	-16.517	35.150	47	1981	1990	100	78.1
30		Makoka	-15.517	35.217	1027	1981	1990	110	75.9
31		Malula	-14.958	34.985	610	2005	2018	105	77
32		Mangochi	-14.433	35.250	474	1981	1990	111	75.7
33		Matandika	-15.167	35.257	684	2006	2018	98	78.5
34		Mimosa	-16.082	35.583	617	1981	1990	96	78.9
35		Mwansambo	-13.278	34.111	653	2008	2018	83	81.8
36		Mzimba	-11.883	33.617	1329	1981	1990	112	75.4
37		Mzuzu	-11.450	34.017	1256	1981	1990	112	75.4
38		Nkhata Bay	-11.600	34.300	497	1981	1990	111	75.7
39		Nkhota Kota	-12.917	34.267	483	1981	1990	111	75.7
40		Salima	-13.750	34.583	508	1981	1990	111	75.7
41		Thyolo	-16.149	35.217	631	1981	1990	109	76.1
42			Zidyana	-13.231	34.214	567	2008	2018	81

Table SI. List of rain gauge stations and data gaps. The bolded rows represent the independent stations not ingested in initial algorithms for gridded products.

ID	Country	Station	Lat	Long	Altitude	Start	End	Counts (Months)	Data Gap (%)
43		Arusha	-3.333	36.633	1467	1981	1989	108	76.3
44		Bukoba	-1.333	31.817	1142	1981	2012	380	16.7
45		Dar es salaam Airport	-6.867	39.200	55	1981	2017	60	86.8
46		Iringa	-7.633	35.767	1364	1981	1990	111	75.7
47		Kilimanjaro Airport	-3.417	37.067	898	1981	1990	110	75.9
48		Morogoro Maji	-6.818	37.660	512	1981	2018	456	0
49		Moshi	-3.350	37.333	856	1981	1989	108	76.3
50	Tanzania	Mtwara	-10.267	40.183	20	1981	2017	116	74.6
51		Musoma	-1.500	33.800	1147	1981	2014	408	10.5
52		Mwanza	-2.467	32.917	1150	1981	2015	420	7.9
53		Same	-4.083	37.717	899	1981	1990	109	76.1
54		Shinyanga Maji	-3.668	33.319	1190	1981	2010	354	22.4
55		Shinyanga Met	-3.661	33.413	1137	1985	2012	327	28.3
56		Tabora Airport	-5.083	32.833	1178	1981	2014	119	73.9
57		Tengeru	-3.383	36.867	1133	2005	2018	161	64.7
58		Chipata	-13.550	32.583	999	1981	1990	120	73.7
59		Choma	-16.833	27.067	1272	1981	1990	119	73.9
60		Kabompo	-13.600	24.200	1099	1981	1990	112	75.4
61		Kafue Polder	-15.767	27.917	977	1981	1990	120	73.7
62		Kaoma	-14.800	24.800	1162	1981	1990	119	73.9
63		Kasama	-10.217	31.133	1396	1981	1990	120	73.7
64		Kasempa	-13.533	25.850	1181	1981	1990	120	73.7
65		Livingstone	-17.817	25.817	995	1981	1990	120	73.7
66		Lundazi	-12.283	33.200	1149	1981	1990	118	74.1
67	Zambia	Malende	-16.240	27.430	1104	2005	2018	106	76.8
68		Misamfu	-10.100	31.250	1372	1981	2018	311	31.8
69		Mongu	-15.250	23.150	1053	1981	1990	120	73.7
70		Mount Makulu	-15.550	28.250	1227	1981	1990	120	73.7
71		Msekera	-13.650	32.570	1026	1981	2018	453	0
72		Mwinilunga	-11.750	24.433	1321	1981	1990	120	73.7
73		Ndola	-13.000	28.650	1263	1981	1990	120	73.7
74		Serenje	-13.233	30.217	1406	1981	1990	113	75.2
75		Sesheke	-17.467	24.300	958	1982	1990	107	76.5
76		Solwezi	-12.183	26.383	1373	1981	1990	120	73.7
77		Zambezi	-13.533	23.117	1076	1981	1990	118	74.1
78		Jinja	0.450	33.183	1123	1981	1986	60	86.8
79	Uganda	Luweero	1.067	32.467	1073	1981	2016	396	13.2
80		Tororo	0.683	34.167	1176	1981	1986	60	86.8
81	Burundi	Bujumbura	-3.317	29.317	778	1981	1989	108	76.3
82		Muyinga	-2.833	30.333	1681	1981	1989	106	76.8

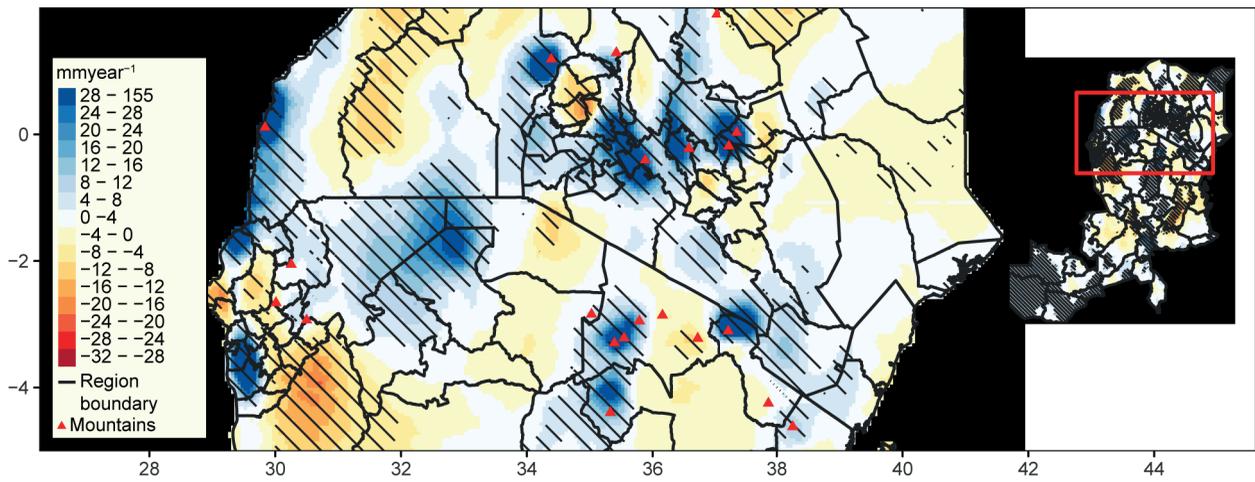


Fig. S1. A zoom-in of the precipitation trends estimated from CHELSA-v2.1 shows an overestimation over the ESA region's mountainous and inland water bodies.