Evaluating TMPA Rainfall over the Sparsely Gauged East African Rift

ELISE MONSIEURS,[a,b,c] DALIA BACH KIRSCHBAUM,[c] JACKSON TAN,[d,e] JEAN-CLAUDE MAKI MATESO,[f] LIESBET JACOBS,[g] PIERRE-DENIS PLISNIER,[h] WIM THIERY,[i,j] AUGUSTA UMUTONI,[k] DIDACE MUSONI,[l] TOUSSAINT MUGARUKA BIBENTYO,[m] GLOIRE BAMULEZI GANZA,[m] GUY ILOMBE MAWE,[m] LUC BAGALWA,[f] CLAIRIA KANKURIZE,[n] CAROLINE MICHELLIER,[a] THOMAS STANLEY,[d,o,c] FRANCOIS KERVYN,[a] MATTHIEU KERVYN,[p] ALAIN DEMOULIN,[b] AND OLIVIER DEWITTE[a]

[a] *Department of Earth Sciences, Royal Museum for Central Africa, Tervuren, Belgium*
[b] *Department of Geography, University of Liège, Liège, Belgium*
[c] *Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[d] *Universities Space Research Association, Columbia, Maryland*
[e] *Climate and Radiation Laboratory, NASA Goddard Space Flight Center, Greenbelt, Maryland*
[f] *Department of Geophysics, Centre de Recherche en Sciences Naturelles, Lwiro, Democratic Republic of the Congo*
[g] *Department of Geography and Tourism, KU Leuven, Leuven, Belgium*
[h] *G. L. Eco, Grez-Doiceau, Belgium*
[i] *Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium*
[j] *Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland*
[k] *Ministry of Infrastructure, Kigali, Rwanda*
[l] *Rwanda Meteorology Agency, Kigali, Rwanda*
[m] *Department of Geology, Université Officielle de Bukavu, Bukavu, Democratic Republic of the Congo*
[n] *Faculté des Sciences de l'Environnement, Université Polytechnique de Gitega, Gitega, Burundi*
[o] *NASA Goddard Earth Sciences Technology and Research, Columbia, Maryland*
[p] *Department of Geography, Earth System Science, Vrije Universiteit Brussel, Brussels, Belgium*

(Manuscript received 15 May 2018, in final form 24 July 2018)

ABSTRACT

Accurate precipitation data are fundamental for understanding and mitigating the disastrous effects of many natural hazards in mountainous areas. Floods and landslides, in particular, are potentially deadly events that can be mitigated with advanced warning, but accurate forecasts require timely estimation of precipitation, which is problematic in regions such as tropical Africa with limited gauge measurements. Satellite rainfall estimates (SREs) are of great value in such areas, but rigorous validation is required to identify the uncertainties linked to SREs for hazard applications. This paper presents results of an unprecedented record of gauge data in the western branch of the East African Rift, with temporal resolutions ranging from 30 min to 24 h and records from 1998 to 2018. These data were used to validate the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) research version and near-real-time products for 3-hourly, daily, and monthly rainfall accumulations, over multiple spatial scales. Results indicate that there are at least two factors that led to the underestimation of TMPA at the regional level: complex topography and high rainfall intensities. The TMPA near-real-time product shows overall stronger rainfall underestimations but lower absolute errors and a better performance at higher rainfall intensities compared to the research version. We found area-averaged TMPA rainfall estimates relatively more suitable in order to move toward regional hazard assessment, compared to data from scarcely distributed gauges with limited representativeness in the context of high rainfall variability.

1. Introduction

Hydrometeorological hazards triggered by extreme rainfall, such as floods and rainfall-initiated landslides, pose a serious socioeconomic threat in many parts of the world and more particularly in mountainous areas (Kjekstad and Highland 2009; Jacobs et al. 2016a;

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-18-0103.s1.

Corresponding author: Elise Monsieurs, elise.monsieurs@africamuseum.be

DOI: 10.1175/JHM-D-18-0103.1

Alfieri et al. 2017; Kumar et al. 2017). Moreover, in the context of ongoing climate change, it is estimated that rainfall extremes may intensify, particularly in the tropics (IPCC 2013; Gariano and Guzzetti 2016; Souverijns et al. 2016; Thiery et al. 2016), which would also increase the vulnerability of the population (Washington et al. 2006). Accurate rainfall data are fundamental to better characterizing extremes as well as ultimately mitigating the weather-related hazards in areas such as tropical Africa. Hydrometeorological hazards (herein referring to both hydrological and hydrologically triggered hazards such as landslides) can be related to high-intensity, short-duration events and/or long-duration rainfall (Gariano and Guzzetti 2016; Sidle and Bogaard 2016). High-resolution rainfall data are therefore fundamental for accurate hazard assessment (Brunetti et al. 2018; Kirschbaum and Stanley 2018).

Ground-based rainfall measurements in tropical Africa are either sparse or nonexistent and often include erroneous data or large gaps (Serrat-Capdevila et al. 2016; Dezfuli et al. 2017). The main reasons are the lack of African states' means and political will or interest to support such data collection, as well as the relatively high cost of establishing and maintaining infrastructure in these areas (Washington et al. 2006; Monsieurs et al. 2017). This problem is even more prominent in mountainous areas, where rainfall presents large spatial variability due to strong topographic transitions (Dinku et al. 2008) and the susceptibility to hydrometeorological hazards is generally high (Sidle and Bogaard 2016). Spatial and temporal discontinuities in rainfall data from gauges limit the ability to study regional extremes over a long record. In addition to the constraints on data availability in space and time, the use of rain gauges comprise a range of error sources for rainfall measurements such as, for example, rainfall undercatch and gauge malfunctioning (Sevruk et al. 2009; Grimaldi et al. 2018, and references herein). Finally, the latency for data availability generally hampers the development of hazard early warning systems (Gebregiorgis et al. 2017).

Satellite rainfall estimates (SREs) with high spatial and temporal resolution and large areal coverage provide an opportunity for regional rainfall data acquisition in remote areas. Yet, SREs remain an ongoing challenge (Hobouchian et al. 2017; Rossi et al. 2017; Brunetti et al. 2018). Visible and infrared sensors on board geostationary satellites infer surface rainfall based on cloud albedo and cloud-top temperature, respectively (Kidd 2001). The dominant associated uncertainties emanate from warm-rain processes in the development stage of deep convection in tropical areas or orographic enhancement of rainfall (Dinku et al. 2008, 2010) and subcloud evaporation (Mashingia et al. 2014; Hobouchian et al. 2017). Passive microwave sensors on low-Earth-orbiting satellites allow more accurate estimates of instantaneous rainfall by observing the precipitation signal within or beneath the cloud (Kidd and Huffman 2011). However, cold surfaces and ice cover may be wrongly interpreted as rainy scenes, and the low observation frequency is problematic (Dinku et al. 2010; Kummerow et al. 2015). Better results for accurate rainfall estimation can be achieved by products derived from the combination of microwave observations (high-quality rainfall observation) and infrared observations (higher spatiotemporal resolution and continuous sampling; Ebert 2007; Kidd and Huffman 2011; Salio et al. 2015; Gebregiorgis et al. 2017; Poméon et al. 2017). Still, these combined products suffer strong uncertainties in topographically complex terrain (Derin and Yilmaz 2014; Zambrano-Bigiarini et al. 2017) and over areas comprising inland water bodies, inducing a complicated microwave signal due to the cold water and the warm land (Tian and Peters-Lidard 2007).

This paper focuses on the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) product for six reasons: 1) it is currently among the most widely recognized multisatellite rainfall products (Gebregiorgis et al. 2017; Hobouchian et al. 2017); 2) it was designed to improve tropical rainfall observations by combining microwave and infrared rainfall estimates, at a spatial resolution of 0.25° × 0.25° and 3-hourly temporal resolution (Huffman et al. 2007); 3) it provides one of the longest consistent records (from 1998 to present) of freely available, spatially homogeneous SRE products over the tropics; 4) it has been validated with satisfactory results in many parts of the world (Dinku et al. 2008; Habib et al. 2009; Islam et al. 2012; El Kenawy et al. 2015; Munzimi et al. 2015); 5) it was proven successful in several hydrometeorological hazard applications (e.g., Li et al. 2009; Kirschbaum et al. 2015; Yaduvanshi et al. 2015; Avalon Cullen et al. 2016; Abdelkareem 2017; Kumar et al. 2017; Thiery et al. 2017; Kirschbaum and Stanley 2018); and 6) it serves as important input data for high-resolution, satellite-based rainfall estimates such as Climate Hazards Group Infrared Precipitation with Stations (CHIRPS; 0.05°; Funk et al. 2015). While the TRMM satellite is no longer operating, the multisatellite TMPA product will continue to be produced until the IMERG (from 2014 to present, 0.1°, half-hourly) retrospective reprocessing is completed (expected by late 2018). This reprocessing will provide a higher spatial and temporal resolution from 2000 to present and will extend the latitudinal coverage to 65°N–65°S (Huffman et al. 2015).

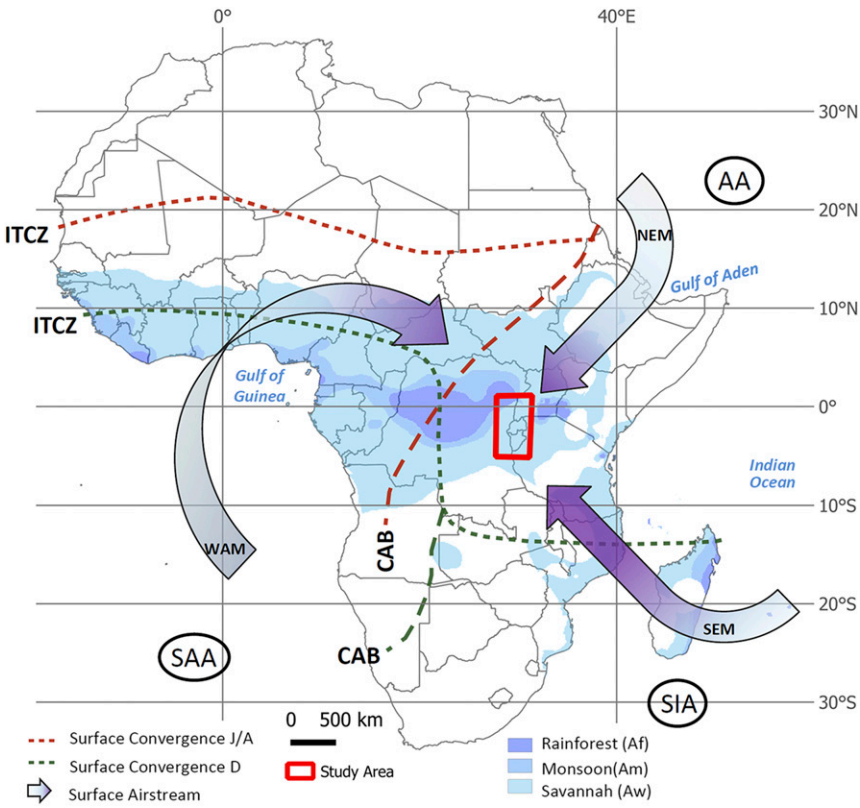

FIG. 1. Surface pressure patterns, airstreams, and convergence zones that affect the climate in the study area (modified after Nicholson 1996; Gasse et al. 2008). SAA = South Atlantic anticyclone, AA = Arabian anticyclone, SIA = South Indian anticyclone, WAM = West African monsoon, NEM = northerly East African monsoon, SEM = southerly East African monsoon, CAB = Congo Air Boundary (red line: July/August, green line: December). Tropical climate types are according to the Köppen–Geiger classification (Peel et al. 2007).

Rigorous validation is necessary to characterize uncertainties in the SRE records in order to more effectively use these data for hydrometeorological hazard modeling and support local authorities in risk management (Mashingia et al. 2014). Even though many current SRE data are freely available, little validation has been done on a regional scale in tropical Africa, particularly at daily and subdaily time scales and in mountainous regions (Cattani et al. 2016). In this paper, we therefore aim to evaluate the TMPA product over a data-scarce region in tropical Africa with complex topography. We focus on the western branch of the East African Rift, a region known for being highly susceptible to hydrometeorological hazards (Maki Mateso and Dewitte 2014; Jacobs et al. 2016a, 2017; Monsieurs et al. 2017, 2018; Stanley and Kirschbaum 2017). The paper is organized as follows: section 2 describes the validation region and the data used for analysis, section 3 outlines the validation approach, section 4 provides results of this approach, section 5 discusses the findings, and the final section presents the conclusions. A list with acronyms is given in appendix A.

2. Setting and data description

a. Study area

The validation region stretches over the western branch of the East African Rift from Lake Tanganyika up to the Rwenzori Mountains (Fig. 1). The regional climate is driven by three principal surface airstreams and two major surface convergence zones (Fig. 1). A comprehensive description of the complex system of climate drivers in equatorial Africa is given by Dezfuli (2017). The region has a bimodal rainfall regime (Fig. 2), with the first rainy season starting in September (Monsieurs et al. 2018). While December–February is considered a "dry" season (Nicholson 1996), results from the average monthly TMPA data in Fig. 2 show that there are still over 100 mm of average rainfall, whereas the period from June to August can clearly be distinguished as a dry season.

Rainfall anomalies are related to the strong interannual variability of the ITCZ (Souverijns et al. 2016), El Niño–Southern Oscillation, and Indian Ocean

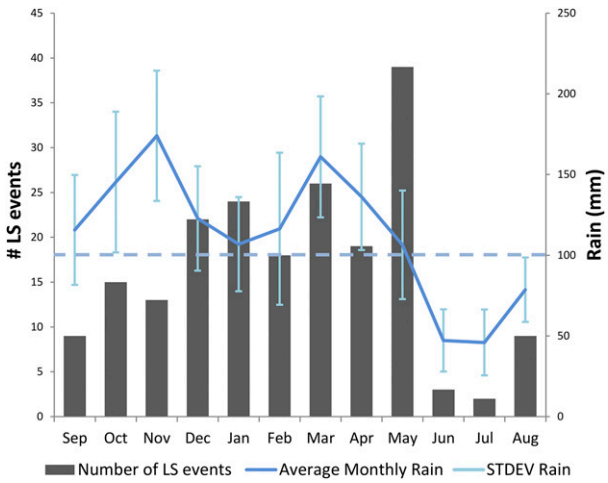

FIG. 2. Rainfall seasonality and distribution of landslide events (LS) in the study area. The dashed line highlights the 100-mm monthly average rainfall threshold that was used to group the analyses in wet and dry months. The presented mean monthly rainfall is based on 20 years (1998–2018) of TMPA 3B42 daily data. Modified after Monsieurs et al. (2018).

dipole events (Behera and Yamagata 2001; Shaaban and Roundy 2017). This is reflected by the large standard deviations of monthly rainfall in Fig. 2. In addition, the local climate is strongly modulated by the complex rift topography (Jacobs et al. 2016a; Smets et al. 2016), and the presence of large lakes (Thiery et al. 2015; Docquier et al. 2016). Consequently, there is large rainfall variability across the study area. The complex topography, presence of lakes, and variable rainfall regimes within the study area represent many of the typical features that pose difficulties for estimating rainfall from satellite observations.

Several types of natural hazards threaten the densely populated areas in the study region, including floods, earthquakes, volcanic eruptions, landslides, and their interactions (Jacobs et al. 2016b, 2017; Michellier et al. 2016; Delvaux et al. 2017; Thiery et al. 2017; Nobile et al. 2018). Similar to rain gauge data collection, collecting information on hydrometeorological hazards is tedious in the study area (Monsieurs et al. 2017, 2018) and beyond the scope of the present study. Figure 2 provides a summary of the most extensive landslide hazard inventory currently available in the region, which is an updated version of the inventory compiled by Monsieurs et al. (2018) comprising 199 landslide events with known location and date over a time span of 50 years (1968–2018). There is a clearly observable signal related to the wetting of the soil throughout the two rainy seasons and subsequent landslide occurrences, which reach a peak in May.

b. Data description

1) RAIN GAUGE DATA

Gauge data with a sampling frequency of one day or better were collected for 24 gauges from a variety of sources, including international research projects; universities; and local research, religious, and governmental institutions. In addition, a gauge network was installed and maintained as part of the Remote Sensing and In Situ Detection and Tracking of Geohazards (RESIST) and AfReSlide projects since 2014 (http://resist.africamuseum.be/; http://afreslide.africamuseum.be/), which comprises 10 gauges in the Rwenzori Mountains (Uganda) and 12 on the Rift flanks in DR Congo. These gauges have a temporal resolution ranging between 30 min and 1 h. The 46 gauges in total (herein referred to as the gauge network) cover the latitudinal range of the study area; however, an optimal distribution of gauges in the longitude is mainly hampered by inaccessibility due to low levels of security west of Lake Kivu (Fig. 3). Minimum and maximum rain gauge elevations in this network are 664 and 2435 m, with a mean of 1600 m and a standard deviation of 428 m. The temporal extent of the gauge data over the TMPA record used in this study (from 1 January 1998 to 1 January 2018) is presented in Fig. 4. This gauge network currently represents the most extensive data compilation from gauges with a temporal resolution of 24 h or better over the study area. Moreover, validation of SRE in equatorial Africa has, to the authors' knowledge, never been performed with a comparable rain gauge density at the present spatiotemporal resolution. However, continuous time series are almost nonexistent because of frequent power outages and political instabilities in the region. Hence, it is not feasible to limit the validation to only the period when all gauge data temporally overlap. We therefore consider all data presented in Fig. 4 for further analyses, keeping in mind that gauges with long records may have a stronger influence on the results. An extended description of the gauge network installation, maintenance, and data preparation is given in the online supplemental material.

Six gauges coincide with gauges in the GPCC network used for TMPA calibration. Gauges used for satellite rainfall calibration should mostly be avoided for validation (Dinku et al. 2010), although this has not always been the case (Su et al. 2008; Poméon et al. 2017). However, we choose here to include these six gauges for validation, because of the limited data availability in this region, but also because GPCC gauge data are only used to adjust the monthly bias (Huffman and Bolvin 2014), and the day-to-day relative variations in rain are driven entirely by the satellite data.

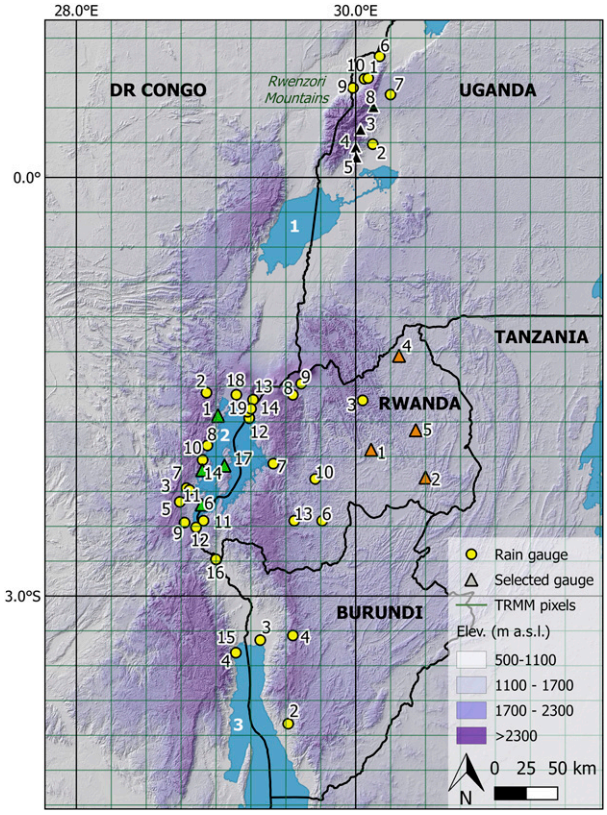

FIG. 3. Spatial distribution of rain gauges in the study area with a temporal resolution of 24 h or better. Gauges represented as triangles refer to the selected gauges for the validation in distinct contexts on the ground: Black = topographic complex terrain, green = presence of large water bodies, orange = relative low-altitude continental environment. Numbers and colors refer to gauge codes in Fig. 4. Numbers in the lakes: 1 = Lake Edward, 2 = Lake Kivu, 3 = Lake Tanganyika. Background hill shading from SRTM (90 m).

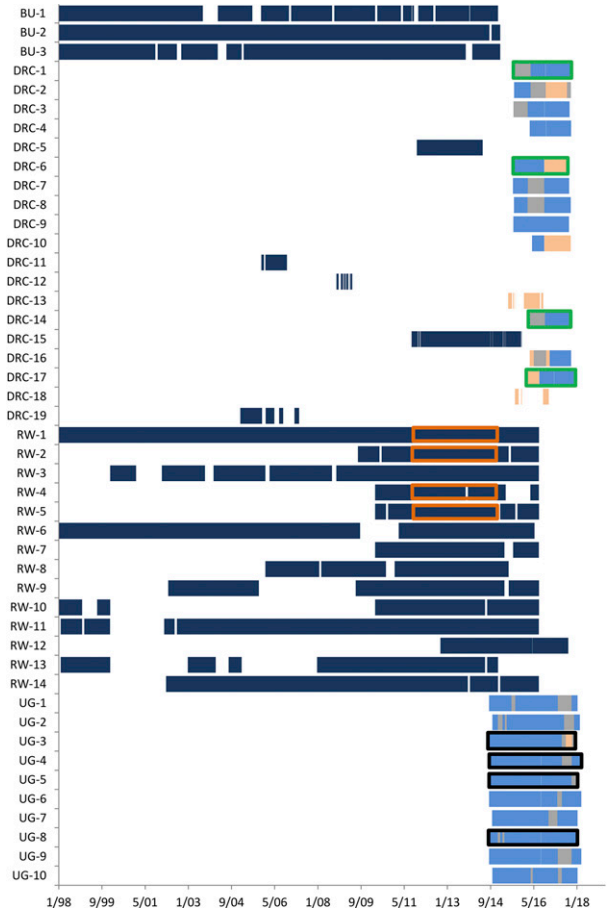

FIG. 4. Temporal characterization of rain gauge data from 1 Jan 1998 to 1 Jan 2018. Gauge codes refer to the numbers on the map in Fig. 3 with an additional land code: BU = Burundi, DRC = DR Congo, RW = Rwanda, UG = Uganda. Gauges installed and maintained by the authors are light blue; gauge data collected from other sources are dark blue. Periods that contain suspicious data are shaded light orange, and data gaps are shaded gray. The colored boxes refer to the selected periods and gauges used for the validation in distinct contexts on the ground: Black = topographic complex terrain, green = presence of large water bodies, orange = relative low continental environment.

The 46 gauges are located in 31 different TMPA pixels. Eight TMPA pixels contain more than one gauge, with a total amount of 23 gauges. Gauge data within the same TMPA pixel are not averaged so that all (TMPA, gauge) pairs are evaluated in a consistent way. In addition, averaging of gauge data within each TMPA pixel would decrease the magnitude of extreme rainfall events, which are of high relevance for hydrometeorological hazard assessment. Instead, we treat the multiple gauge data within the same TMPA pixel as independent comparisons, each of which are assigned the same TMPA value when temporally overlapping. The gauge network comprises 20 gauges that are located in TMPA pixels containing inland water bodies (Fig. 3).

2) TMPA DATA

The TMPA algorithm uses instruments on board TRMM to serve as the calibration standard for a network of passive microwave sensors. Gaps in the high-quality microwave observations are filled in using microwave-calibrated infrared data, thus allowing TMPA to provide coverage from 50°N to 50°S every 3 h (Huffman et al. 2007, 2017). The average revisit time of the microwave observations changes with latitude and available sensors (Hou et al. 2014; Nelkin 2017). TMPA represents a snapshot at some point during the 3-h window, that is, not the actual 3-hourly average rain rate (Villarini and Krajewski 2007) over an area of ~775 km^2 (at the equator). TMPA data are available as a variety of products (Huffman et al. 2007). The research version (3B42, version 7, herein referred to as

TMPA ResV) is available from 1 January 1998 to present, over 50°N–50°S, with a 0.25° × 0.25°, 3-hourly spatiotemporal resolution. This version uses Precipitation Radar (PR) and TRMM Microwave Imager (TMI) for calibration of all passive microwave inputs, and is also calibrated against the gauge-based GPCC rainfall data on a monthly basis, though the correction ratio is limited to between 0.2 and 3.0 mm h^{-1} (Huffman et al. 2007). Because of this postprocessing, TMPA ResV is only available after 2 months (Huffman and Bolvin 2014). The near-real-time product (3B42RT, version 7, herein referred to as TMPA RT) has the same spatiotemporal resolution and is available from 1 March 2000 to present, with postprocessing limited to TMI-only calibration (Huffman et al. 2010). On the other hand, the latency of TMPA RT is only eight hours and it has a spatial extent from 50°N to 50°S with experimental data currently extending from 50°–60°N to 50°–60°S. Further details can be found in Kummerow et al. (2000, 2015) and Huffman et al. (2007, 2017).

TMPA ResV and RT were downloaded from NASA Goddard Earth Sciences Data and Information Services Center (https://disc-beta.gsfc.nasa.gov/). Accumulations are computed taking into account the local time zone and gauge accumulation period. We only use TMPA data when corresponding gauge data were available. The equivalent time series for TMPA ResV and RT comprise 92 941 and 87 357 days, respectively.

3. Methodology

Taking the spatiotemporal constraints of the gauge dataset into account (Figs. 3, 4), we adopted a validation scheme that checks multiple validation approaches and spatiotemporal scales, providing a broad picture of TMPA's error characteristics. We focus on illustrating the differences between the two TMPA products toward providing insights for their use in hydrometeorological hazard assessment. However, assessing the hazards themselves is out of the scope of this work. Namely, the identification of rainfall thresholds for triggering hazards will be done in a separate study.

a. Validation statistics

In this work three validation approaches are used. All statistical tests are described in Table 1, based on the overview of standard and diagnostic validation methods by Wilks (2006) and Ebert (2007). The entire validation was performed using the R open-source software, release 3.3.2 (http://www.r-project.org/). The first approach comprises visual comparison of the rainfall value distributions, using quantile–quantile (QQ) plots. The second validation approach measures TMPA's accuracy for continuous variables including rainfall amount and intensity, while normalizing by the gauge-measured rainfall rate of the respective period. This approach relies on a combination of accuracy metrics to evaluate scatter, or random error, by use of the Pearson correlation coefficient (COR); error direction, by normalized mean error (NME; also known as mean bias); error magnitude, by normalized mean absolute error (NMAE); and importance of extremes, by normalized root-mean-square error (NRMSE). Last, TMPA's accuracy in rainfall detection is evaluated based on a contingency table (Table 2). A comprehensive statistic summarizing the contingency table is the Heidke skill score (HSS; Heidke 1926; Wilks 2006). HSS measures how well estimates perform compared to random chance, with negative values indicating a worse performance than random chance, positive values indicating a better-than-random performance, and a value of one corresponding to perfect skill. For the purpose of evaluating SRE for hydrometeorological hazard applications, additional highly relevant metrics are the probability of detection (POD) and the probability of false alarm (POFA; Martelloni et al. 2012; Rossi et al. 2017; Thiery et al. 2017). The use of the term POFA is preferred above *false alarm ratio* (Wilks 2006) because of its common confusion with *false alarm rate*, which is a different metric (Barnes et al. 2009). POD refers to TMPA's ability to correctly identify rain occurrence, that is, separate wet and dry days or, if rain thresholds are imposed to evaluate TMPA's efficiency in identifying heavy rains, separating days drier or wetter than the threshold. POFA works the same way with respect to thresholds.

b. Spatiotemporal validation scales

Validation has been conducted for three temporal rainfall accumulation periods: 3-hourly (TMPA native resolution), daily, and monthly. We focus mostly on the daily time scale to utilize the entire gauge dataset, since 35% of the gauges have no subdaily information. Furthermore, daily resolution is highly relevant for regional hazard model calibration due to the dearth of information on the exact time of the occurrence of hydrometeorological hazards (e.g., Kirschbaum et al. 2015), especially in the context of central Africa (Jacobs et al. 2016a; Monsieurs et al. 2018). Mountainous environments, however, are characterized by high rainfall variability and induce accelerations of streamflow volume concentration that might cause hydrometeorological hazards such as flash floods (Devrani et al. 2015). Therefore, we validated TMPA on a 3-hourly resolution for the rainy seasons in 2016, that is, March–May and September–November, using 13 gauges from the network we maintain with subdaily temporal resolution and which have data over

TABLE 1. Validation metrics for the evaluation of SREs. Variable G refers to gauge observations and S to satellite observations. Variables $\overline{G}$ and $\overline{S}$ are the means and δ_G and δ_S are the standard deviations of G and S, respectively. Variable N is the total amount of observations. The terms in HSS, POD, and POFA are explained in Table 2. Normalized metrics are expressed in percent; all metrics are unitless.

Statistical measure	Equation	Range	Optimum score	Description
Pearson correlation coefficient	$\mathrm{COR} = \dfrac{\sum_i^N (G_i - \overline{G})(S_i - \overline{S})}{\sqrt{\sum_i^N (G_i - \overline{G})^2}\sqrt{\sum_i^N (S_i - \overline{S})^2}}$	$[-1, 1]$	1	Reflects the degree of linear correlation between SRE and gauge observations.
Normalized mean error (%)	$\mathrm{NME} = \dfrac{\sum_i^N (S_i - G_i)}{\sum_i^N G_i} \times 100$	$[-100, +\infty]$	0	Average difference between SRE and gauge observations. Overestimation is represented as positive bias.
Normalized mean absolute error (%)	$\mathrm{NMAE} = \dfrac{\sum_i^N \lvert S_i - G_i \rvert}{\sum_i^N G_i} \times 100$	$[0, +\infty]$	0	Mean magnitude of the errors without considering their direction.
Normalized root-mean-square error (%)	$\mathrm{NRMSE} = \dfrac{\sqrt{\frac{1}{N}\sum_i^N (S_i - G_i)^2}}{\frac{1}{N}\sum_i^N G_i} \times 100$	$[0, +\infty]$	0	Mean magnitude of the errors, giving greater weight to the larger errors.
Heidke skill score	$\mathrm{HSS} = \dfrac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}$	$[-\infty, 1]$	1	Fractional improvement of SRE over random forecast.
Probability of detection	$\mathrm{POD} = \dfrac{\mathrm{hit}}{\mathrm{hit} + \mathrm{miss}}$	$[0, 1]$	1	Likelihood that the event would be forecasted, given that it occurred.
Probability of false alarm	$\mathrm{POFA} = \dfrac{\mathrm{false\ alarm}}{\mathrm{hit} + \mathrm{false\ alarm}}$	$[0, 1]$	0	Proportion of forecasted events that fail to materialize.

TABLE 2. Contingency table for categorical validation metrics, where *a*–*d* are as in Table 1 for HSS.

	Gauge > 0	Gauge = 0
SRE > 0	Hit (*a*)	False alarm (*c*)
SRE = 0	Miss (*b*)	Correct negative (*d*)

the selected period. Last, TMPA performance is tested on a monthly scale, because accumulated antecedent rainfall conditions are proven to play a crucial role for hazards such as landslides (Segoni et al. 2018). However, monthly TMPA accumulations are strongly influenced by GPCC. Only continuous metrics are computed at the monthly scale as detection of rainfall for monthly accumulations is irrelevant to SRE skill evaluation. Months comprising gaps have been excluded from this analysis.

We evaluated the different temporal rainfall accumulations by season, ground conditions, and rainfall characteristics. These factors are known to affect the errors in SREs (Kidd and Huffman 2011). Two elevation classes (500–1700 and 1700–2500 m MSL) are chosen such that the amount of gauge data in each class is approximately equal. A threshold of 100-mm average monthly rainfall was chosen to define wet versus dry periods and test the impact of seasonality on TMPA's performance (Fig. 2). To evaluate TMPA's sensitivity to rainfall intensity, data were grouped into 10 intensity classes as measured by the gauges and adapted from Kim et al. (2017).

In addition to the spatial grouping by elevation, we selected four gauges from three contrasting environments (further referred to as "context"): 1) the Lake Kivu area, with TMPA pixels having more than 35% of their surface covered by water; 2) the region above 1700 m in the Rwenzori Mountains, characterized by very complex topography; and 3) the comparatively low-altitude continental environment of eastern Rwanda, with an average altitude of 1480 (±80) m MSL for the four gauges (Fig. 3). The first two contexts include exclusively gauges that we maintain, of which all available data are included for this analysis. No information on the data quality is available in the third context group, for which data from 2012 to 2015 have been selected in order to evaluate SREs over a similarly long time period with almost no data gaps. The selected gauges and time periods are highlighted in Fig. 4. Analyses are performed at daily and monthly temporal resolution, elucidating the impact of these different ground conditions to the performance of TMPA.

c. *Point-to-pixel approach*

The sparse and temporally and spatially heterogeneous gauge coverage does not allow for an accurate rainfall interpolation, especially given the strong topographic gradients within the domain. Therefore, gauge data are not extrapolated and aggregated to TMPA's resolution, as recommended by Chen and Knutson (2008). Instead, we apply the point-to-pixel validation approach (Islam et al. 2010; Thiemig et al. 2012; El Kenawy et al. 2015). A severe limitation to this approach is the discrepancy between local-scale gauge data and spatially averaged TMPA data in the context of local convective storms and orographic rainfall.

Indeed, local rainfall variability introduces uncertainty into the gauge data used as a reference to validate TMPA. This source of uncertainty is especially relevant for extreme rainfall analyses (Chen and Knutson 2008; Sun and Barros 2010). To evaluate the gauge data uncertainty, we study the correlation between the time series of independently measuring gauges in a pixel and the daily standard deviations between their respective measurements. Five TMPA pixels contain two gauges and one pixel contains four gauges with overlapping time series. There are a total of 3599 (ResV) and 3651 (RT) overlapping days between intrapixel gauges over these six pixels, with an average overlap per pixel of 604 days. The average distance between the gauges in one pixel is 18.72 (±7.28) km. To estimate the relative effects of gauge and TMPA data uncertainties on the validation process, the measures of local rainfall variability within a pixel are compared with the mean residual standard error (ResSE) of TMPA data. This error (mm day^{-1}) is calculated as the square root of the mean squared residual of the linear regression of TMPA against gauge data.

d. *Analyses of extremes*

A final analysis considers TMPA's performance in capturing extreme rainfall events. The hundred highest daily rainfall events for TMPA and gauge data over the same spatial and temporal domain are evaluated, of which the 10 highest records are related to reported hydrometeorological hazards. Information on hydrometeorological hazards is collected using the following principal sources: Monsieurs et al. (2018, their Fig. 2), http://reliefweb.int, http://floodlist.com, http://www.emdat.be, http://www.glidenumber.net, Vandecasteele et al. (2010), and Jacobs et al. (2016a). Analyses are performed for the time period from March 2000 to January 2018 in order to be able to compare for each event the three datasets, that is, gauge, TMPA ResV, and TMPA RT.

4. Results

Validation results relevant for using TMPA in hydrometeorological hazard assessment are first presented for the daily temporal resolution, allowing a broad

picture of TMPA's performance over multiple spatial scales. This is followed by the 3-hourly and monthly validation results for the selected gauges and time periods. Last, we show the analyses on daily extremes related to hydrometeorological hazards for the entire gauge network. Results for TMPA ResV and RT data products are compared to analyze the trade-offs for the short latency of TMPA RT data, as compared to the supposedly improved ResV product.

a. Daily resolution

1) REGIONAL PERFORMANCE

Overall, we found that TMPA RT underestimates rainfall (negative NME) on average more severely (40.35%) at the regional scale compared to the ResV (15.15%; Table 3, category A). The QQ plot (Fig. 5) confirms that the TMPA distributions significantly differ from that of gauge data, from which RT deviates more severely. The correlation between both datasets at a daily scale is moderately low (COR; Table 3, category A). While TMPA ResV has lower rainfall underestimations compared to TMPA RT, the mean absolute error is higher (NMAE; Table 3), which implies that NMAE is largely driven by random errors that cannot be reduced by gauge correction. The regional detection skills are relatively good (Table 3, category A). However, these metrics have a relatively high standard deviation, indicating a significant variation in performance among the gauges.

When grouping the validation by elevation, we are able to reveal an improvement of the correlation between TMPA RT and the gauges, which outperformed TMPA ResV for both elevation categories. Average daily TMPA rainfall underestimation is lower for low-altitude sites (Table 3, category B). The other statistics do not significantly vary for the two elevation categories (Table 3, category B), pointing to no obvious topographic control on the correlation and detection skills on a regional scale.

Rainfall underestimation by TMPA, error magnitude, and importance of outliers tend to increase during the drier months of the year (Table 3, category C). Even though HSS is similar for wet and dry months, the POD is higher and POFA lower for the wet months (Table 3, category C). The higher correlation for the relatively dry months is related to the many zero-rain days.

When examining the performance of TMPA as a function of daily rainfall intensities, we found that all rainfall intensities above 5 mm day^{-1} are underestimated by TMPA with increasing magnitudes for higher rainfall intensities (Table 3, category D). We note that at the lowest intensities, disproportionally high normalized errors are due to the normalization against low gauge averages. In terms of rain detection, the HSS decreases for higher rainfall intensities but remains positive at all rain intensities, indicating a declining performance in rain detection but still better than random chance (Table 3, category D). The POD has relatively high values at 0 and 1 mm day^{-1}, but declines with increasing rainfall intensities. The probability of false alarms increases for higher rainfall intensities, but drops to zero for the threshold above 60 mm day^{-1} in the RT version just because no values appear above this threshold due to strong underestimation (Table 3, category D).

2) CONTEXT-BASED EVALUATION

TMPA was evaluated for three contrasting contexts in order to quantify the extent to which the validation results are affected by water–land mixed pixels (mixed) and complex topographical environments at higher altitudes (complex topography) as compared to a reference lower-elevation continental context (low). The correlation between TMPA and gauge data within the complex topography context is lowest compared to that in mixed and low-topography contexts (Table 4, category A), with TMPA ResV consistently showing lower values. TMPA's performance for pixels characterized with a complex topography is also significantly lower (Table 4, category A) with respect to average underestimation, error magnitude, importance of outliers, and rainfall detection skill (HSS). However, POD and POFA results do not show the same trend, these two statistics being lowest in mixed pixels (Table 4, category A).

Separating wet and dry months results in TMPA performance significantly lower in mixed pixels than complex topography and low contexts during dry months, according to COR, HSS, POD, and POFA (Table 4, category B). Regarding the rainfall intensity control, HSS and POD are found to degrade most severely with increasing intensities in complex topographic conditions (Table 4, category C).

Looking at the complete picture of validation results over different ground contexts in different seasons and grouped by rainfall intensities, we found TMPA overall encounters the greatest difficulties in correctly estimating rainfall in complex topographical contexts compared to mixed and low contexts (Table 4). We studied another facet of the orographic control, namely, the effect of leeward versus windward mountain side, on the daily TMPA performance by testing separately two gauges from each flank of the Rwenzori Mountains (northwest: UG6, UG9; southeast: UG3, UG4; Fig. 3), knowing that the southeast flank is the rain shadow side (Jacobs et al. 2016a). Validation results for COR, NME, HSS, and POFA are significantly better for the rainy side of the

TABLE 3. Summary of validation results for daily TMPA ResV and RT products (both version 7). The number of days used for analyses in each category is given under "Days." Category A is regionally calculated metrics with their standard deviations (Std dev), category B is elevation groups (m MSL), category C is groups based on average monthly rainfall (Fig. 2), and category D is intensity groups (all in mm day^{-1}; bold numbers refer to the thresholds used for the calculation of HSS, POD, and POFA). The results in parentheses for the highest intensity group refer to the same outcome as the second highest intensity group, as the upper threshold for the calculations is the same.

		COR		NME		NMAE		NRMSE		HSS		POD		POFA		Days	
		ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT
A	Regional	0.38	0.37	−15.15	−40.35	108.29	97.58	231.10	217.33	0.36	0.36	0.64	0.63	0.30	0.30	92 941	87 357
	Std dev	0.10	0.11	21.36	15.50	14.70	9.80	35.62	31.29	0.10	0.09	0.07	0.06	0.09	0.09		
B	<1700 m	0.39	0.47	−13.12	−38.17	108.26	97.45	239.91	225.89	0.39	0.38	0.66	0.65	0.34	0.34	49 226	47 264
	>1700 m	0.34	0.48	−28.32	−49.46	102.82	94.67	220.34	208.93	0.41	0.41	0.64	0.63	0.30	0.30	43 715	40 093
C	>100 mm month^{-1}	0.34	0.35	−20.57	−43.90	104.76	95.29	208.59	196.91	0.34	0.34	0.66	0.65	0.29	0.28	69 592	65 450
	<100 mm month^{-1}	0.42	0.43	−26.67	−44.42	112.40	103.92	387.13	375.31	0.37	0.36	0.57	0.56	0.56	0.57	23 349	21 907
D	**0** mm day^{-1}	—	—	—	—	—	—	—	—	0.40	0.39	0.65	0.64	0.32	0.32	51 160	47 901
	<**1** mm day^{-1}	0.06	0.05	473.30	308.84	575.68	413.34	1213.07	873.19	0.40	0.36	0.61	0.56	0.39	0.50	9566	9157
	1–**5** mm day^{-1}	0.08	0.08	45.62	2.05	146.01	111.09	242.43	166.11	0.33	0.19	0.42	0.24	0.51	0.64	13 980	13 224
	5–**10** mm day^{-1}	0.07	0.07	−27.81	−48.93	82.65	75.18	103.64	87.13	0.26	0.10	0.29	0.10	0.60	0.74	7529	7069
	10–**15** mm day^{-1}	0.04	0.03	−46.81	−62.06	73.87	72.95	84.72	79.78	0.20	0.05	0.20	0.05	0.69	0.81	4327	4023
	15–**20** mm day^{-1}	0.02	0.02	−58.48	−70.34	71.80	74.44	79.20	79.61	0.15	0.03	0.14	0.03	0.75	0.85	2345	2195
	20–**25** mm day^{-1}	0.06	0.07	−59.74	−71.32	69.09	73.58	75.45	78.30	0.11	0.02	0.10	0.01	0.81	0.90	1510	1424
	25–**30** mm day^{-1}	0.04	0.05	−66.29	−76.41	72.40	77.54	78.03	81.23	0.09	0.01	0.07	0.01	0.84	0.94	944	887
	30–**60** mm day^{-1}	0.05	0.08	−72.66	−81.29	75.59	81.57	81.36	85.68	0.03	0.00	0.02	0.00	0.94	0.00	1426	1339
	>**60** mm day^{-1}	0.09	0.13	−82.78	−87.51	82.78	87.51	87.28	90.51	(0.03	0.00)	(0.02	0.00)	(0.94	0.00)	154	138

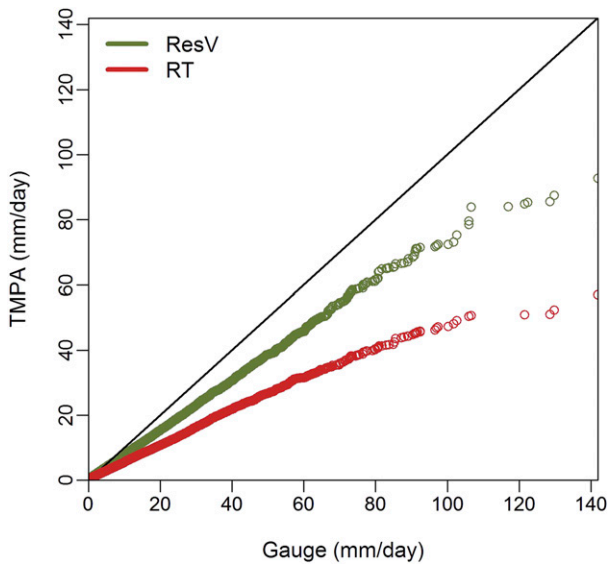

FIG. 5. QQ plot for daily rainfall measured by the gauges and the equivalent TMPA databased on the availability of rain gauge data in that pixel for the entire study area. The solid black line presents the 1:1 line. TMPA distributions significantly differ from that of gauge data, of which the TMPA RT product deviates more than the TMPA ResV product.

mountain (northwest). Rainfall is more severely underestimated on the rain shadow (southeast) side, with NME = −39.23% (ResV), −49.47% (RT), compared to the northwest flank where NME = −14.91% (ResV), −26.73% (RT).

3) INTRAPIXEL VARIABILITY

The most meaningful evidence for estimating gauge data uncertainty is found in the pixel with four gauges, which allows a more robust estimation of the intergauge correlation and the standard deviation of the daily samples. With about 700 days of overlapping gauge data, analyzing this single pixel is statistically significant from the temporal point of view. We observe that while the average correlation between the gauges is only ~0.29 (ResV: 0.293; RT: 0.296), the average gauge-TMPA correlation is higher ~0.34 (ResV: 0.335; RT: 0.336; Table 5). The problem of this high intergauge variability, highlighted in an example of stacked 1-month excerpts of their time series (Fig. 6), is confirmed by the much improved gauge–TMPA correlation when first averaging the gauge data: ~0.47 (ResV: 0.472; RT: 0.470). This shows that TMPA validation is strongly biased by the poor representation at the pixel scale of the field evidence we use as control data.

We quantify the gauge data uncertainty by the average daily standard deviation between gauge measurements, which is calculated to be ~4 mm day^{-1} (ResV: 4.03; RT: 3.92) for the pixel containing four gauges. Once the underestimation trend is removed, the ResSE error on TMPA RT data with respect to ground "truth" (3.86 mm day^{-1}) is slightly lower than the uncertainty in the gauge data (Table 5). Taking into account that this error is affected by the uncertainty of both TMPA and gauge data, we can conclude that most of the TMPA residual error is linked to the uncertainty on gauge data and TMPA would thus appear as a good performing indicator of relative rainfall at pixel scale in the study area.

Using TMPA ResV data adds uncertainty to the recognized gauge data uncertainty, with a higher 4.74 mm day^{-1} value of TMPA ResSE compared to the average standard deviation in the gauge data (Table 5). This is in line with our earlier findings, namely, that the gauge-calibrated TMPA ResV product has a lower precision in the study area compared to TMPA RT. Owing to this, we propose that TMPA RT data are a better rainfall proxy than TMPA ResV data in this area.

However, with one single pixel, we cannot illustrate a possible spatial variability in gauge uncertainty. Therefore, we also present results for the five pixels containing two gauges, but with a prior call for caution in discussing them because (n = 2) samples are far from ideal to get realistic standard deviation estimates and tend to underestimate them on average. Results show values of ~2.4 mm day^{-1} (ResV: 2.44; RT: 2.42), which are somewhat smaller compared to the standard error on residuals for TMPA in the same pixels (ResV: 5.22; RT: 3.28; Table 5). Besides confirming that a large part of the uncertainty in TMPA is related to the gauge uncertainty and that TMPA ResV introduces its own contribution to uncertainty compared to TMPA RT, the figures obtained for the two-gauge pixels also suggest that TMPA data uncertainty is marginally affected by local topography.

Overall, these results highlight the high local rainfall variability over the study area, which implies that multiple rain gauges are needed to represent the rainfall in a 25 km × 25 km area (TMPA pixel size). Hence, the limited quality of SRE validation at the TMPA resolution is partly due to the inadequacy of using individual rain gauges as control data. This also highlights the relevance of TMPA area-averaged rainfall estimates for regional hydrometeorological hazard analyses in this area, as point observations represent only a small area in the proximity of the gauge.

b. 3-hourly and monthly resolution

We also tested TMPA's performance at its 3-hourly native resolution during the rainy seasons of 2016. Validation results in Table 6 are fairly similar to, though slightly lower than, the results obtained for daily rainfall accumulations (Table 3). Yet, results for POD are lower

TABLE 4. Evaluation of daily TMPA ResV and RT products in different contexts on the ground. "Mixed" is pixels that include large water bodies, "Co. topogr." is complex topographic terrain, and "Low" is relatively low terrain. For each of these contexts four gauges have been selected (indicated in Figs. 3 and 4). Category A is averaged metrics per category, category B is elevation groups (m MSL), category C is groups based on average monthly rainfall (Fig. 2), and category D is intensity groups (all in mm day^{-1}; bold numbers refer to the thresholds used for the calculation of HSS, POD, and POFA). The results in parentheses for the highest intensity group refer to the same outcome as the second highest intensity group, as the upper threshold for the calculations is the same.

			COR		NME		NMAE		NRMSE		HSS		POD		POFA		Days	
			ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT
A		Mixed	0.43	0.45	−9.80	−43.94	106.97	92.44	215.94	199.21	0.34	0.34	0.60	0.60	0.30	0.30	2396	2396
		Co. topogr.	0.34	0.39	−33.94	−44.80	103.23	100.46	213.94	215.57	0.29	0.29	0.65	0.65	0.35	0.36	4530	4614
		Low	0.43	0.50	−8.55	−28.73	108.51	98.31	255.88	230.30	0.44	0.43	0.66	0.64	0.37	0.37	4350	4350
B	Mixed	>100 mm month^{-1}	0.41	0.43	−10.98	−45.28	102.97	89.11	190.65	175.92	0.30	0.31	0.62	0.62	0.22	0.22	1728	1728
		<100 mm month^{-1}	0.22	0.19	6.42	−25.68	161.52	137.96	414.99	381.34	0.19	0.18	0.44	0.44	0.69	0.70	668	668
	Co. topogr.	>100 mm month^{-1}	0.34	0.33	−34.24	−45.57	102.63	100.29	208.23	211.36	0.27	0.26	0.65	0.65	0.34	0.35	3499	3583
		<100 mm month^{-1}	0.31	0.33	−32.51	−41.09	106.15	101.31	236.67	230.30	0.35	0.36	0.67	0.67	0.40	0.39	1031	1031
	Low	>100 mm month^{-1}	0.41	0.42	−7.30	−30.79	108.30	97.22	231.34	207.62	0.41	0.39	0.66	0.63	0.33	0.33	3247	3247
		<100 mm month^{-1}	0.40	0.53	−23.86	−3.46	111.15	111.66	431.79	404.35	0.41	0.41	0.67	0.68	0.60	0.61	1103	1103
C	Mixed	**0** mm day^{-1}	—	—	—	—	—	—	—	—	0.34	0.34	0.60	0.60	0.30	0.30	1197	1197
		0–**5** mm day^{-1}	0.16	0.14	114.49	26.26	217.42	140.77	376.13	228.30	0.35	0.47	0.48	0.49	0.52	0.36	733	733
		5–**25** mm day^{-1}	0.14	0.16	−37.05	−59.96	75.21	72.29	91.96	84.88	0.21	0.28	0.18	0.18	0.71	0.25	399	399
		>**25** mm day^{-1}	−0.13	−0.11	−62.96	−77.32	69.11	77.32	78.56	83.59	(0.21	0.28)	(0.18	0.18)	(0.71	0.25)	67	67
	Co. topogr.	**0** mm day^{-1}	—	—	—	—	—	—	—	—	0.29	0.29	0.65	0.65	0.35	0.36	2233	2314
		0–**5** mm day^{-1}	0.16	0.17	84.29	50.56	183.52	154.93	324.09	264.46	0.30	0.28	0.39	0.34	0.52	0.51	1336	1339
		5–**25** mm day^{-1}	0.10	0.11	−57.68	−64.99	73.67	74.87	87.58	88.11	0.07	0.03	0.05	0.02	0.76	0.80	789	789
		>**25** mm day^{-1}	0.15	0.19	−78.51	−83.04	79.11	83.04	86.28	89.43	(0.07	0.03)	(0.05	0.02)	(0.76	0.80)	172	172
	Low	**0** mm day^{-1}	—	—	—	—	—	—	—	—	0.44	0.43	0.66	0.64	0.37	0.37	2803	2803
		0–**5** mm day^{-1}	0.19	0.21	85.62	36.73	183.37	140.54	338.38	232.36	0.39	0.45	0.45	0.47	0.49	0.41	854	854
		5–**25** mm day^{-1}	0.17	0.18	−34.65	−48.39	77.62	74.41	99.27	89.59	0.12	0.17	0.13	0.13	0.85	0.73	617	617
		>**25** mm day^{-1}	0.18	0.21	−67.77	−74.38	73.07	77.20	80.07	82.36	(0.12	0.17)	(0.13	0.13)	(0.85	0.73)	76	76

TABLE 5. Evaluation of gauge and TMPA ResV and RT data uncertainty in pixels containing multiple gauges. Analyses are done using daily rainfall estimates: "Mean intergauge correlation" is the average correlation between the gauge data within a pixel; "Mean gauge–TMPA correlation" is the average correlation between each gauge and the same TMPA value of that pixel; "Correlation mean gauge data–TMPA" is the correlation between the averaged gauge data within a pixel and the respective TMPA values; "Mean std dev gauge data" is the average standard deviation (mm day^{-1}) between the gauge measurements in a pixel; and "TMPA ResSE" is the square root of the mean squared residual (mm day^{-1}) in the linear model for TMPA with gauge data as independent variable. Number of days in overlapping time series used for the analysis are presented in parentheses.

	Mean intergauge correlation	Mean gauge–TMPA correlation	Correlation mean gauge data–TMPA	Mean std dev gauge data	TMPA ResSE
ResV					
One four-gauge pixel (699 days)	0.293	0.335	0.472	4.03	4.74
Five two-gauge pixels (2900 days)	0.399	0.399	0.473	2.44	5.22
RT					
One four-gauge pixel (724 days)	0.296	0.336	0.470	3.92	3.86
Five two-gauge pixels (2927 days)	0.395	0.385	0.458	2.42	3.28

and POFA higher (Table 6, category A) than those for the daily values (Table 3). When grouping for rainfall intensities, we find zero probability of detection for rainfall intensities higher than 25 mm/3 h, although TMPA recorded 74 events above this threshold (Table 6, category B). Also POFA degrades quickly for any rainfall detection per 3 h (Table 6, category B). Differences between the performance of both TMPA products at 3-hourly resolution is small overall, with the RT version generally performing better for COR, NMAE, and NRMSE (Table 6).

Validation results for monthly rainfall accumulations show an expected significant improvement relative to 3-hourly and daily TMPA rainfall estimates (Table 7, category A). In an explorative focus on monthly values over four gauges from each ground context (Figs. 3, 4), TMPA encounters more difficulties over complex topographic terrain (Table 7, category B), confirming earlier findings at daily resolution (Table 4).

c. *Daily extremes related to hydrometeorological hazards*

In addition to the above intensity-grouped analyses, TMPA's performance for high-intensity events is evaluated by examining the 100 highest recorded daily rainfall events from the gauge and TMPA datasets for

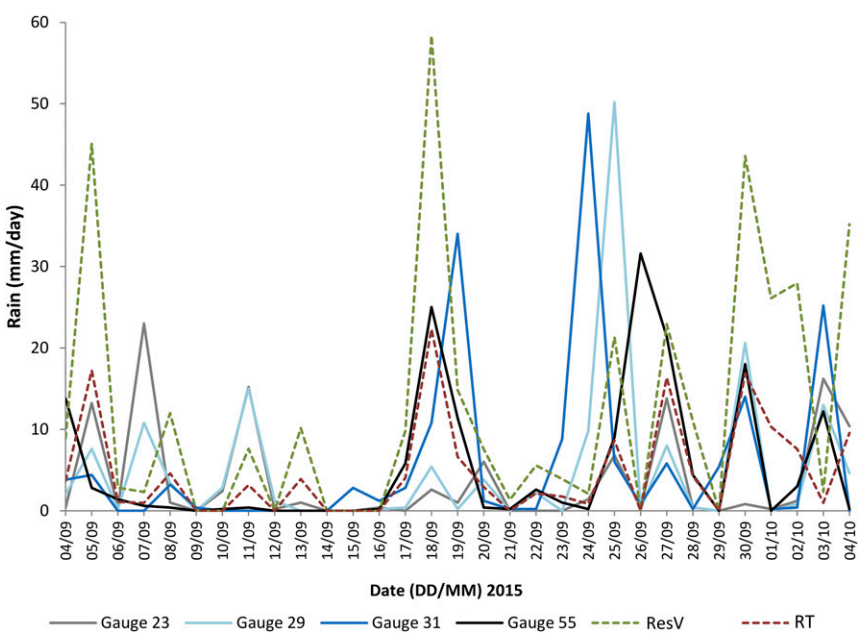

FIG. 6. Daily rainfall time series from 4 Sep to 4 Oct 2015 for four gauges located in one TMPA pixel that is located over the Rwenzori Mountains, together with the TMPA estimates of this pixel (Fig. 3). This presents the variability in observed rainfall in a 25 km × 25 km area.

TABLE 6. Evaluation of TMPA ResV and RT products at 3-hourly resolution. Analyses are based on data from the 13 gauges we maintain (indicated in Fig. 4) during two rainy seasons in 2016 when available (March–May, September–November). The number of observations used for analyses in each category is given under "Time stamps." Category A is averaged results for the selected gauges and time period with their standard deviation (Std dev), and category B is intensity groups (all in mm day^{-1}; bold numbers refer to the thresholds used for the calculation of HSS, POD, and POFA). The results in parentheses for the highest intensity group refer to the same outcome as the second highest intensity group, as the upper threshold for the calculations are the same.

		COR		NME		NMAE		NRMSE		HSS		POD		POFA		Time stamps	
		ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT
A	Average	0.33	0.34	−16.59	−34.49	127.32	115.57	501.54	473.23	0.36	0.36	0.41	0.41	0.47	0.47	16 952	16 952
	Std dev	0.07	0.07	23.75	17.60	13.79	8.95	59.01	55.30	0.04	0.04	0.04	0.04	0.08	0.08		
B	**0** mm/3 h	—	—	—	—	—	—	—	—	0.37	0.37	0.41	0.41	0.46	0.46	14 006	14 006
	>0–**5** mm/3 h	0.21	0.24	35.63	3.69	158.26	129.92	312.56	228.61	0.27	0.25	0.27	0.22	0.68	0.64	2333	2333
	5–**25** mm/3 h	0.10	0.12	−66.43	−73.18	78.28	79.65	93.86	94.34	0.00	0.02	0.00	0.01	1.00	0.80	539	539
	>**25** mm/3 h	0.15	0.06	−87.81	−89.29	87.81	89.37	93.27	95.06	(0.00	0.02)	(0.00	0.01)	(1.00	0.80)	74	74

the time period from March 2000 to January 2018, referred to as "extremes." The highest daily rainfall recorded by TMPA ResV (93 mm day^{-1}) and RT (57 mm day^{-1}) is markedly lower than that of the gauges (142 mm day^{-1}; Figs. 7a,b). Figures 7a and 7b plot TMPA ResV and RT against gauge values for the 100 most extreme gauge and TMPA ResV records, respectively. Results indicate an asymmetric behavior of the recorded extremes, that is, all gauge extremes are underestimated by TMPA (Fig. 7a), whereas most (but not all) TMPA extremes overestimate rainfall measured by the gauges (Fig. 7b), with a bigger bias magnitude in the former case (Fig. 7c) than in the latter (Fig. 7d). Figure 7b clearly illustrates the difference between TMPA ResV and RT as a result of the correction of TMPA ResV through calibration with gauges on a monthly basis. While this correction results in an average lower bias between gauge and SRE values (Table 3, category A), it tends to be problematic for extreme rainfall intensities recorded by TMPA, resulting in increased discrepancy between ResV and gauge data in the study area. Limiting the gauge adjustment factor to a narrower range could result in better SRE for TMPA ResV.

When considering the top 10 gauge extremes for an exploratory analysis on hazard-triggering rainfall events, results indicate that TMPA misses or severely underestimates them (Fig. 7c). By contrast, TMPA extremes overestimate gauge-measured rainfall intensities, especially in their ResV form (Fig. 7d, appendix B). The 10 TMPA and gauge extremes were matched with reported hydrometeorological hazards in the corresponding pixels (Figs. 7c,d and Fig. B1). We found 7 out of the top 10 gauge extremes to be related with hydrometeorological hazards of flooding, landslides, or flash floods. Because of the remoteness of large parts of the study area, there may be unreported hazard occurrences. Six (ResV) and two (RT) out of the respective top 10 TMPA extremes are also related to hydrometeorological hazards, even though some were completely missed by the corresponding field data (Fig. 7d, event 6; see Fig. B1 in appendix B).

5. Discussion

Results of the TMPA ResV and RT analysis identify several biases inherent to SRE and gauge field data in the study area that should be taken into account when interpreting the validation results. First, whereas we initially considered gauge data as a "ground truth" for SRE validation, intrapixel rainfall analyses have shown that gauge point observations are not necessarily representative of the average rainfall over the ~25 km × 25 km area of a TMPA pixel at daily resolution (Fig. 6;

TABLE 7. Evaluation of monthly TMPA ResV and RT products. The number of months used for analyses in each category is given under "Months." Category A is regionally calculated metrics with their standard deviation (Std dev) (all gauge data in Fig. 3 are used). Category B is evaluation in different contexts on the ground. "Mixed" represents pixels that include large water bodies, "Co. topogr." is complex topographic terrain, "Low" is relatively low terrain. For each of these contexts four gauges have been selected (indicated in Figs. 3 and 4).

		COR		NME		NMAE		NRMSE		Months	
		ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT
A	Regional	0.70	0.69	−15.64	−41.09	38.92	50.21	52.97	65.57	3008	2829
	Std dev	0.18	0.19	18.37	14.10	10.52	9.68	14.53	12.63		
B	Mixed	0.85	0.84	−8.03	−44.28	26.29	47.77	36.91	59.17	75	75
	Co. topogr.	0.51	0.45	−34.07	−45.20	52.66	60.24	69.35	77.73	144	147
	Low	0.76	0.80	−4.36	−25.12	37.13	38.60	54.25	54.00	441	415

Table 5). We found that most of the errors in TMPA data with respect to gauge data are linked to the uncertainty in the gauge data (Table 5). This confirms the concerns raised by Satgé et al. (2016), who analyzed the impact of the gauge representation error on daily rainfall detection metrics. It certainly is a major control on the quality of TMPA validation, which is inevitably degraded if a single reference gauge is unable to record local convective storms that cause heavy rains elsewhere in the TMPA pixel or, conversely, records very local heavy rain not representative at the pixel scale. In principle, this issue should be less acute for higher-resolution SRE,

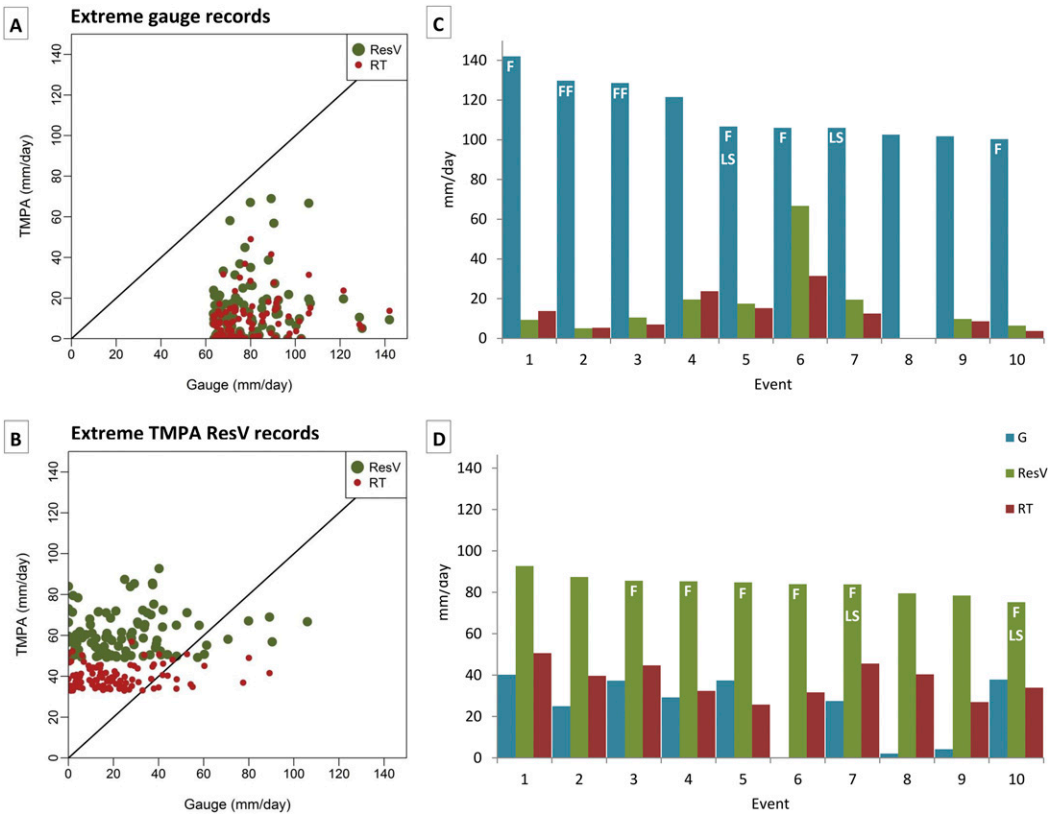

FIG. 7. Extreme daily rainfall recorded by the gauges and by TMPA between 2000 and 2018 based on the availability of rain gauge data in the corresponding TMPA pixel. (a) Scatterplot of the top 100 daily rainfall events as measured by the gauges, indicating that each of these events is underestimated by TMPA ResV and the RT data. (b) Scatterplot of the top 100 daily rainfall events recorded by TMPA ResV. The bias correction of TMPA ResV through calibration with gauges on a monthly basis is clearly visible as a shift between RT and ResV. The top 10 most extreme rainfall events (1 = most extreme) measured by (c) the gauges and (d) TMPA ResV were matched with reported hydrometeorological hazards, if available: Letters in (c) and (d) are F = flood, FF = flash flood, LS = landslide. The corresponding rainfall measured by the gauge, TMPA ResV, and RT in the same pixel is also shown.

such as IMERG. The improvement obtained by averaging multiple gauge data within a pixel is in the line of previous findings showing that spatiotemporal averaging substantially reduces biases and improves performance (Mantas et al. 2015; Tan et al. 2017; Tang et al. 2018). Second, gauge data themselves are likely underestimating rainfall due to wind undercatch, which is known as a severe source for systematic bias (on the order of ~20%) in gauge-based measurements (Sevruk et al. 2009; Mekonnen et al. 2015).

There are very few studies that use a similar study domain and validation context (spatiotemporal resolution, tropical environment, validation approach), limiting a comparison of our results with existing literature. Yuan et al. (2017) validated daily TMPA ResV in the tropical monsoon area of the Chindwin River basin, Myanmar (22°–27°N) by applying normalized validation metrics over the time period from April 2014 to December 2015. There, TMPA ResV regional performance is lower than in the western branch of the East African Rift, showing a maximum COR of 0.356, NMEs ranging between −41.2% and +5%, POD between 0.092 and 0.299, and POFA between 0.404 and 0.626. A 1998–2006 study of TMPA (mainly ResV) in the La Plata basin, South America (Su et al. 2008) found for the tropical upper Paraguay area a POD and POFA of 0.36 and ~0.70, respectively, for rainfall $> 20\,\mathrm{mm\,day^{-1}}$, which is better compared with values of 0.14 and 0.75 in our study area (Table 3, category D). Potential explanatory differences between the two regions might include a larger number of gauges and generally lower elevations and relief in the upper Paraguay area.

The agreement between gauge and TMPA was reduced at TMPA's native temporal resolution, that is, 3-hourly, to a significant extent only for NMAE, NRMSE, POD, and POFA, whereas the average bias, Pearson correlation coefficient, and overall detection skill (HSS) were found similar to those at the daily scale (Table 3, category A, and Table 6). A study by Scheel et al. (2011) found a significantly lower correlation coefficient (0.018) in subtropical Bolivia for TMPA at 3-hourly resolution. By contrast, spatiotemporal averaging substantially reduces biases and improves performance (Mantas et al. 2015; Tan et al. 2017), as confirmed by the better monthly TMPA validation results in Table 7 and other studies in the same area (Adeyewa and Nakamura 2003; Munzimi et al. 2015).

The separate assessment of wet and dry periods (defined based on a 100-mm threshold for monthly average rainfall; Fig. 2) revealed TMPA's performance to be lowest for several metrics (NRMSE, POD, POFA) during the dry months (Tables 3, 4). One possible cause for the poorer performance during the dry season is subcloud evapotranspiration (Mashingia et al. 2014; Hobouchian et al. 2017). Greater evaporation means that, for the same amount of rain produced in the cloud aloft (and hence the same ice scattering signature used in the passive microwave algorithm to estimate the surface rain rate), the rain that actually reaches the surface is lower, leading to a higher probability of false alarms during the dry season. Serrat-Capdevila et al. (2016) also found smaller errors in TMPA within areas that follow the seasonally oscillating ITCZ, attributing this observation to the dependence of TMPA quality on the associated convective rainfall regime. The relatively poor performance of TMPA during the dry months has limited implication on hazard prediction, as hydrometeorological hazards will generally occur less frequently in this period (Fig. 2).

With 20 out of 46 gauges being located in TMPA pixels containing large inland water bodies, the evaluation results are likely to be affected by detection problems of SRE over water–land mixed pixels (e.g., Huffman et al. 2007; Derin and Yilmaz 2014), and thus potentially do not represent the true TMPA performance. A comparison of the validation of mixed pixels only (Table 4) against the overall TMPA performance (Table 3) shows a strong degradation of the results during dry months (Table 4, category B) but, surprisingly, an otherwise increase in TMPA performance compared to the regional validation results (Table 3), with COR reaching 0.45 for TMPA RT in mixed pixels (Table 4, category A). By contrast, a more rugged topography seems to impact the results toward an overall lower TMPA performance (Tables 3, 4). In terms of local elevation, TMPA is better correlated with field data and provides more accurate rainfall estimation for lower elevations (<1700 m MSL; Table 3, category B). This is possibly due to the occurrence of orographically controlled rainfall at higher elevations. Many studies have underlined the difficulty for SRE to estimate such rainfall compared to convectively driven rainfall (Dinku et al. 2010; Mantas et al. 2015; Serrat-Capdevila et al. 2016). This is also confirmed here by the poorer performance of TMPA over the complex topographical setting in the Rwenzori Mountains with respect to mixed pixels and pixels in lower-altitude continental environments (Table 4).

TMPA's performance is drastically degraded for higher rainfall intensities (Table 3, category D; Table 4, category C; Table 6, category B). Decreased detection skills and a transition from over- to underestimation by TMPA when gauge-based rainfall increases is consistent with previous studies (e.g., Dinku et al. 2008; Vila et al. 2009; Scheel et al. 2011; Gao and Liu 2013; Satgé et al. 2016). As a consequence, high-intensity

storms and extreme events are especially poorly reproduced in SREs (Fig. 7). Even though TMPA uses a combination of several satellite estimates, if a subdaily event is of short duration and high intensity, the satellite microwave observations may entirely miss these peak intensities given the more than 3-h revisit time in this area (Huffman et al. 2007). While infrared data are used to fill in gaps between microwave overpasses, these data may also be biased by the relationship between cloud-top temperatures and rainfall intensities in this complex climatologic and topographic setting (Huffman et al. 2007; Kidd and Huffman 2011). Improvements are expected for extreme rainfall detection using IMERG, as shown already in other parts of the world (Prakash et al. 2016a; Hobouchian et al. 2017; Xu et al. 2017).

A performance comparison between TMPA ResV and RT indicates that the latter underestimates rainfall on average more severely over the entire range for all different spatiotemporal scales. While this agrees with the overall findings of previous studies and is explained by the lack of gauge-based adjustment of TMPA RT (e.g., Habib et al. 2009; Shen et al. 2010; Prakash et al. 2016b; Satgé et al. 2016), values for correlation, absolute error, sensitivity to outliers, and categorical validation metrics show here overall better results for TMPA RT compared to TMPA ResV (Tables 3, 4, 6, 7). In contrast to what has been argued in the previous studies, this better performance of RT data is most likely related to the fact that, due to the sparsity of gauge data in central Africa, the GPCC-based monthly correction applied to TMPA ResV does not adequately represent the rainfall variability over the study area. Regarding the use of TMPA for flood simulations, some studies have shown increased uncertainties associated with the adjusted TMPA ResV product compared to the RT product (Su et al. 2008; Bitew and Gebremichael 2011). This might be related to our results on extreme rainfall events, with TMPA RT extremes lying closer to the corresponding gauge values than the ResV extremes (Figs. 7b,d). In addition, the intrapixel analysis has shown that the TMPA ResV product adds uncertainty to the recognized gauge data uncertainty, whereas the uncertainty in TMPA RT is mainly related to the gauge data uncertainty (Table 5). Hence, given its short latency and better performance TMPA RT is probably more relevant in hazard applications over the western branch of the East Africa Rift.

6. Conclusions

While widely available, there remain challenges for accurate rain detection and quantification of SREs. This paper outlines SRE uncertainties in a sparsely gauged, low-latitude region with complex topography in the western branch of the East African Rift. TMPA RT and ResV were evaluated at multiple spatiotemporal scales from 1998 to 2018 with an unprecedented dataset of 46 gauges. Results indicated that the sparse and heterogeneous temporal gauge coverage and high rainfall variability in the study region poses challenges for TMPA validation. The validation approach allowed detection of trends and sources of bias in TMPA and will be applied to the validation of IMERG as soon as the reprocessed product, spanning from 1998 to the present, is released in late 2018. The latter's higher spatiotemporal resolution will allow a more effective use of gauge data for validating rainfall with high variability, which is a cause of large uncertainties in TMPA. Results indicate that TMPA performs relatively better in areas without complex topography, and systematically underestimate precipitation for rainfall $> 5\,\mathrm{mm\,day^{-1}}$. TMPA performance decreases in predictive power during the dry months. Validation results for 3-hourly and daily TMPA are found to be very similar, whereas the performance significantly increases for monthly rainfall accumulations. Trade-offs for the short latency of TMPA RT were found to be small, showing overall higher bias with gauge data, but better rainfall detection skills and lower absolute errors compared to TMPA ResV, probably as a result of the latter's gauge-based calibration. TMPA's error characteristics highlighted in this paper will improve the efficient use of TMPA in hydrometeorological hazard applications. Especially in the study region, TMPA is indispensable to provide the regional rainfall information required in hazard assessment, owing to the sparse gauge network. Despite the key challenges identified for satellite rainfall detection by TMPA in the western branch of the East African Rift, and as long as IMERG products are not available over a long period of time, TMPA remains one of the best sources of regional rainfall information available in the study area. However, and although we recognized that weaknesses of the gauge data might be partly responsible for the somewhat disappointing quality of our validation results, TMPA should be used with caution for hazard assessment.

Acknowledgments. We are grateful to all institutions, research projects, and people who shared rain gauge data, including Appalachian State University (USA), Centre de Recherches en Hydrobiologie (DR Congo), EAGLES (http://www.eagles-kivu.be/project.htm) and CHOLTIC projects (BELSPO, Belgium), Meteo Rwanda

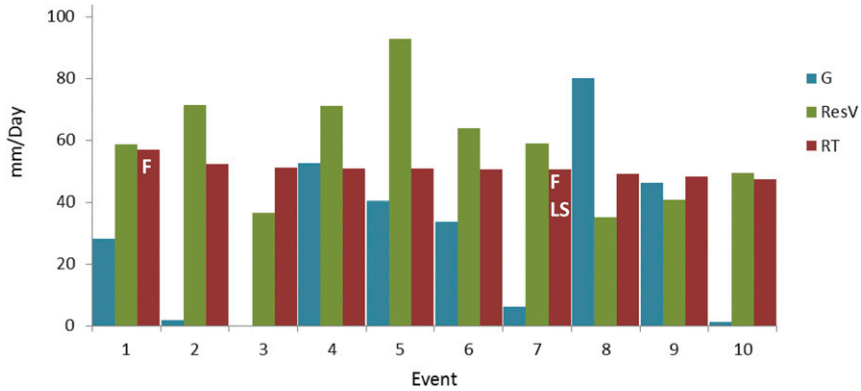

FIG. B1. Top 10 most extreme rainfall events (1 = most extreme) measured between 2000 and 2018 by TMPA RT were matched with reported hydrometeorological hazards, if available: F = flood and LS = landslide. The corresponding rainfall measured by the gauge in the same pixel (G) as well as the TMPA ResV are also shown.

(Rwanda), National Oceanic and Atmospheric Administration (USA), Observatoire Volcanologique de Goma (DR Congo), Pères Blancs Bukavu (DR Congo), University of Burundi (Burundi), Université Polytechnique de Gitega (Burundi), and United States Geological Survey (USA). Special thanks go to our partners at Centre de Recherche en Sciences Naturelles de Lwiro (DR Congo), Université Officielle de Bukavu (DR Congo), Kahuzi-Biega National Park (DR Congo; Mr. Radar Nishuli), and Mountains of the Moon University (Uganda), who are dedicated to maintaining and regularly collecting data from the gauge network in sometimes difficult field conditions. We thank Thom Bogaard for the fruitful discussions on the validation results. We also thank the Civil Protection in DR Congo and Olivier Machiels, who shared their insights on landslide hazard in the study area and made it possible to execute fieldwork. We thank George Huffman for sharing his expertise on the TMPA algorithms. The authors acknowledge NASA Goddard Earth Sciences Data and Information Services Center for providing full access to the precipitation datasets exploited in this study. Datasets can be accessed at https://disc-beta.gsfc.nasa.gov/. Financial support came from BELSPO for RESIST (SR/00/305), AfReSlide (BR/121/A2/AfReSlide), and GeoRisCA (SD/RI/02A) research projects (http://resist.africamuseum.be/, http://afreslide.africamuseum.be/, http://georisca.africamuseum.be/), and an F.R.S.-FNRS PhD scholarship for E.M. The Belgian American Education Foundation facilitated a one-year research stay for E.M. at the Hydrological Sciences Laboratory at NASA Goddard Space Flight Centre. J.T. is supported by an appointment to the NASA Postdoctoral Program at Goddard Space Flight Center, administered by USRA through a contract with NASA (NNH15CO48B). Finally, we thank the reviewers for their help in improving the content of the paper.

APPENDIX A

List of Acronyms and Abbreviations

COR	Pearson correlation coefficient
DR Congo	Democratic Republic of Congo
GPCC	Global Precipitation Climatology Centre
HSS	Heidke skill score
IMERG	Integrated Multisatellite Retrievals for Global Precipitation Measurement
ITCZ	Intertropical convergence zone
NMAE	Normalized mean absolute error
NME	Normalized mean error
NRMSE	Normalized root-mean-square error
POD	Probability of detection
POFA	Probability of false alarm
PR	Precipitation radar
QQ	Quantile–quantile
ResV	Research version
ResSE	Residual standard error
RT	Near–real time
SRE	Satellite rainfall estimate
TMI	TRMM Microwave Imager
TMPA	TRMM Multisatellite Precipitation Analysis
TRMM	Tropical Rainfall Measuring Mission

APPENDIX B

Top 10 Most Extreme Rainfall Events Measured by TMPA RT

Figure B1 shows the top 10 most extreme rainfall events measured by TMPA RT.

REFERENCES

Abdelkareem, M., 2017: Targeting flash flood potential areas using remotely sensed data and GIS techniques. *Nat. Hazards*, **85**, 19–37, https://doi.org/10.1007/s11069-016-2556-x.

Adeyewa, Z. D., and K. Nakamura, 2003: Validation of TRMM radar rainfall data over major climatic regions in Africa. *J. Appl. Meteor.*, **42**, 331–347, https://doi.org/10.1175/1520-0450(2003)042<0331:VOTRRD>2.0.CO;2.

Alfieri, L., B. Bisselink, F. Dottori, G. Naumann, A. De Roo, P. Salomon, K. Wyser, and L. Feyen, 2017: Global projections of river flood risk in a warmer world. *Earth's Future*, **5**, 171–182, https://doi.org/10.1002/2016EF000485.

Avalon Cullen, C., R. Al-Suhili, and R. Khanbilvardi, 2016: Guidance index for shallow landslide hazard analysis. *Remote Sens.*, **8**, 866, https://doi.org/10.3390/rs8100866.

Barnes, L. R., D. M. Schultz, E. C. Gruntfest, M. H. Hayden, and C. C. Benight, 2009: Corrigendum: False alarm rate or false alarm ratio. *Wea. Forecasting*, **24**, 1452–1454, https://doi.org/10.1175/2009WAF2222300.1.

Behera, S. K., and T. Yamagata, 2001: Subtropical SST dipole events in the southern Indian Ocean. *Geophys. Res. Lett.*, **28**, 327–330, https://doi.org/10.1029/2000GL011451.

Bitew, M. M., and M. Gebremichael, 2011: Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands. *Hydrol. Earth Syst. Sci.*, **15**, 1147–1155, https://doi.org/10.5194/hess-15-1147-2011.

Brunetti, M. T., M. Melillo, S. Peruccacci, L. Ciabatta, and L. Brocca, 2018: How far are we from the use of satellite rainfall products in landslide forecasting? *Remote Sens. Environ.*, **210**, 65–75, https://doi.org/10.1016/j.rse.2018.03.016.

Cattani, E., A. Merino, and V. Levizzani, 2016: Evaluation of monthly satellite-derived precipitation products over East Africa. *J. Hydrometeor.*, **17**, 2555–2573, https://doi.org/10.1175/JHM-D-15-0042.1.

Chen, C.-T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *J. Climate*, **21**, 1605–1621, https://doi.org/10.1175/2007JCLI1494.1.

Delvaux, D., J. Mulumba, M. Ntabwoba, S. Sebagenzi, S. Fiama, F. Kervyn, and H. Havenith, 2017: Seismic hazard assessment of the Kivu rift segment based on a new seismotectonic zonation model (western branch, East African Rift system). *J. Afr. Earth Sci.*, **134**, 831–855, https://doi.org/10.1016/j.jafrearsci.2016.10.004.

Derin, Y., and K. K. Yilmaz, 2014: Evaluation of multiple satellite-based precipitation products over complex topography. *J. Hydrometeor.*, **15**, 1498–1516, https://doi.org/10.1175/JHM-D-13-0191.1.

Devrani, R., V. Singh, S. M. Mudd, and H. D. Sinclair, 2015: Prediction of flash flood hazard impact from Himalayan river profiles. *Geophys. Res. Lett.*, **42**, 5888–5894, https://doi.org/10.1002/2015GL063784.

Dezfuli, A. K., 2017: Climate of western and central equatorial Africa. Climate science, Oxford Research Encyclopedias, https://doi.org/10.1093/acrefore/9780190228620.013.511.

——, C. M. Ichoku, G. J. Huffman, K. I. Mohr, J. S. Selker, N. van de Giesen, R. Hochreutener, and F. O. Annor, 2017: Validation of IMERG precipitation in Africa. *J. Hydrometeor.*, **18**, 2817–2825, https://doi.org/10.1175/JHM-D-17-0139.1.

Dinku, T., S. Chidzambwa, P. Ceccato, S. J. Connor, and C. F. Ropelewski, 2008: Validation of high-resolution satellite rainfall products over complex terrain. *Int. J. Remote Sens.*, **29**, 4097–4110, https://doi.org/10.1080/01431160701772526.

——, S. J. Connor, and P. Ceccato, 2010: Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America. *Satellite Rainfall Applications for Surface Hydrology*, M. Gebremichael and H. Faisal, Eds., Springer, 193–204, https://doi.org/10.1007/978-90-481-2915-7_11.

Docquier, D., W. Thiery, S. Lhermitte, and N. P. M. van Lipzig, 2016: Multi-year wind dynamics around Lake Tanganyika. *Climate Dyn.*, **47**, 3191–3202, https://doi.org/10.1007/s00382-016-3020-z.

Ebert, E. E., 2007: Methods for verifying satellite precipitation estimates. *Measuring Precipitation from Space*, V. Levizzani, P. Bauer, and F. J. Turk, Eds., Springer, 345–356, https://doi.org/10.1007/978-1-4020-5835-6_27.

El Kenawy, A. M., J. I. Lopez-Moreno, M. F. McCabe, and S. M. Vicente-Serrano, 2015: Evaluation of the TMPA-3B42 precipitation product using a high-density rain gauge network over complex terrain in northeastern Iberia. *Global Planet. Change*, **133**, 188–200, https://doi.org/10.1016/j.gloplacha.2015.08.013.

Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations — A new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, https://doi.org/10.1038/sdata.2015.66.

Gao, Y. C., and M. F. Liu, 2013: Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau. *Hydrol. Earth Syst. Sci.*, **17**, 837–849, https://doi.org/10.5194/hess-17-837-2013.

Gariano, S. L., and F. Guzzetti, 2016: Landslides in a changing climate. *Earth-Sci. Rev.*, **162**, 227–252, https://doi.org/10.1016/j.earscirev.2016.08.011.

Gasse, F., F. Chalié, A. Vincens, M. A. Williams, and D. Williamson, 2008: Climatic patterns in equatorial and southern Africa from 30,000 to 10,000 years ago reconstructed from terrestrial and near-shore proxy data. *Quat. Sci. Rev.*, **27**, 2316–2340, https://doi.org/10.1016/j.quascirev.2008.08.027.

Gebregiorgis, A. S., P. E. Kirstetter, Y. E. Hong, N. J. Carr, J. J. Gourley, W. Petersen, and Y. Zheng, 2017: Understanding overland multisensor satellite precipitation error in TMPA-RT products. *J. Hydrometeor.*, **18**, 285–306, https://doi.org/10.1175/JHM-D-15-0207.1.

Grimaldi, S., A. Petroselli, L. Baldini, and E. Gorgucci, 2018: Description and preliminary results of a 100 square meter rain gauge. *J. Hydrol.*, **556**, 827–834, https://doi.org/10.1016/j.jhydrol.2015.09.076.

Habib, E., A. Henschke, and R. F. Adler, 2009: Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmos. Res.*, **94**, 373–388, https://doi.org/10.1016/j.atmosres.2009.06.015.

Heidke, P., 1926: Berechnung des Erfolges und der Giite der Windstarkevorhersagen im Sturmwarnungsdienst. *Geogr. Ann.*, **8**, 310–349.

Hobouchian, M. P., P. Salio, Y. García, D. Vila, and R. Garreaud, 2017: Assessment of satellite precipitation estimates over the slopes of the subtropical Andes. *Atmos. Res.*, **190**, 43–54, https://doi.org/10.1016/j.atmosres.2017.02.006.

Hou, A. Y., and Coauthors, 2014: The global precipitation measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

Huffman, G. J., and D. T. Bolvin, 2014: TRMM and other data precipitation data set documentation. NASA TRMM Doc., 42 pp., ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf.

——, and Coauthors, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor

precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

——, R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM Multi-Satellite Precipitation Analysis (TMPA). *Satellite Applications for Surface Hydrology*, F. Hossain and M. Gebremichael, Eds., Springer, 3–22.

——, D. T. Bolvin, and E. J. Nelkin, 2015: Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code 612 Tech. Doc., 48 pp., http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf.

——, A. Pendergrass, and NCAR Research Staff, Eds., 2017: The Climate Data Guide: TRMM: Tropical Rainfall Measuring Mission. UCAR, accessed 6 January 2017, https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., https://doi.org/10.1017/CBO9781107415324.

Islam, N., S. Das, and H. Uyeda, 2010: Calibration of TRMM derived rainfall over Nepal during 1998-2007. *Open Atmos. Sci. J.*, **4**, 12–23, https://doi.org/10.2174/1874282301004010012.

Islam, T., M. A. Rico-Ramirez, D. Han, P. K. Srivastava, and A. M. Ishak, 2012: Performance evaluation of the TRMM precipitation estimation using ground-based radars from the GPM validation network. *J. Atmos. Solar-Terr. Phys.*, **77**, 194–208, https://doi.org/10.1016/j.jastp.2012.01.001.

Jacobs, L., O. Dewitte, J. Poesen, D. Delvaux, W. Thiery, and M. Kervyn, 2016a: The Rwenzori Mountains, a landslide-prone region? *Landslides*, **13**, 519–536, https://doi.org/10.1007/s10346-015-0582-5.

——, and Coauthors, 2016b: Reconstruction of a flash flood event through a multi-hazard approach: Focus on the Rwenzori Mountains, Uganda. *Nat. Hazards*, **84**, 851–876, https://doi.org/10.1007/s11069-016-2458-y.

——, and Coauthors, 2017: Landslide diversity in the Rwenzori Mountains (Uganda). *WLF 2017: Advancing Culture of Living with Landslides*, M. Mikos et al., Eds., Springer, 79–86, https://doi.org/10.1007/978-3-319-53498-5_10.

Kidd, C., 2001: Satellite rainfall climatology: A review. *Int. J. Climatol.*, **21**, 1041–1066, https://doi.org/10.1002/joc.635.

——, and G. J. Huffman, 2011: Global precipitation measurement. *Meteor. Appl.*, **18**, 334–353, https://doi.org/10.1002/met.284.

Kim, K., J. Park, J. Baik, and M. Choi, 2017: Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.*, **187**, 95–105, https://doi.org/10.1016/j.atmosres.2016.12.007.

Kirschbaum, D. B., and T. Stanley, 2018: Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness. *Earth's Future*, **6**, 505–523, https://doi.org/10.1002/2017EF000715.

——, ——, and J. Simmons, 2015: A dynamic landslide hazard assessment system for Central America and Hispaniola. *Nat. Hazards Earth Syst. Sci.*, **15**, 2257–2272, https://doi.org/10.5194/nhess-15-2257-2015.

Kjekstad, O., and L. Highland, 2009: Economic and social impacts of landslides. *Landslides – Disaster Risk Reduction*, K. Sassa and P. Canuti, Eds., Springer, 573–587, https://doi.org/10.1007/978-3-540-69970-5_30.

Kumar, A., A. K. L. Asthana, R. Singh, R. Jayangondaperumal, A. K. Gupta, and S. S. Bhakuni, 2017: Geomorphology assessment of landslide hazards induced by extreme rainfall event in Jammu and Kashmir Himalaya, northwest India. *Geomorphology*, **284**, 72–87, https://doi.org/10.1016/j.geomorph.2017.01.003.

Kummerow, C. D., and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.*, **39**, 1965–1982, https://doi.org/10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2.

——, D. L. Randel, M. Kulie, N. Y. Wang, R. Ferraro, S. J. Munchak, and V. Petkovic, 2015: The evolution of the Goddard profiling algorithm to a fully parametric scheme. *J. Atmos. Oceanic Technol.*, **32**, 2265–2280, https://doi.org/10.1175/JTECH-D-15-0039.1.

Li, L., and Coauthors, 2009: Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Nat. Hazards*, **50**, 109–123, https://doi.org/10.1007/s11069-008-9324-5.

Maki Mateso, J., and O. Dewitte, 2014: Towards an inventory of landslide processes and the elements at risk on the Rift flanks West of Lake Kivu (DRC). *Geo-Eco-Trop*, **38**, 137–154.

Mantas, V. M., Z. Liu, C. Caro, and A. J. S. C. Pereira, 2015: Validation of TRMM multi-satellite precipitation analysis (TMPA) products in the Peruvian Andes. *Atmos. Res.*, **163**, 132–145, https://doi.org/10.1016/j.atmosres.2014.11.012.

Martelloni, G., S. Segoni, R. Fanti, and F. Catani, 2012: Rainfall thresholds for the forecasting of landslide occurrence at regional scale. *Landslides*, **9**, 485–495, https://doi.org/10.1007/s10346-011-0308-2.

Mashingia, F., F. Mtalo, and M. Bruen, 2014: Validation of remotely sensed rainfall over major climatic regions in Northeast Tanzania. *Phys. Chem. Earth*, **67–69**, 55–63, https://doi.org/10.1016/j.pce.2013.09.013.

Mekonnen, G. B., S. Matula, F. Doležal, and J. Fišák, 2015: Adjustment to rainfall measurement undercatch with a tipping-bucket rain gauge using ground-level manual gauges. *Meteor. Atmos. Phys.*, **127**, 241–256, https://doi.org/10.1007/s00703-014-0355-z.

Michellier, C., P. Pigeon, and F. Kervyn, 2016: Contextualizing vulnerability assessment: A support to geo-risk management in central Africa. *Nat. Hazards*, **82**, 27–42, https://doi.org/10.1007/s11069-016-2295-z.

Monsieurs, E., and Coauthors, 2017: Constraints on landslide-climate research imposed by the reality of fieldwork in Central Africa. *Third North American Symp. on Landslides*, Roanoke, VA, Association of Environmental and Engineering Geologists, 158–168.

——, and Coauthors, 2018: Landslide inventory for hazard assessment in a data-poor context: A regional-scale approach in a tropical African environment. *Landslides*, https://doi.org/10.1007/s10346-018-1008-y, in press.

Munzimi, Y. A., M. C. Hansen, B. Adusei, and G. B. Senay, 2015: Characterizing Congo basin rainfall and climate using Tropical Rainfall Measuring Mission (TRMM) satellite data and limited rain gauge ground observations. *J. Appl. Meteor. Climatol.*, **54**, 541–555, https://doi.org/10.1175/JAMC-D-14-0052.1.

Nelkin, E., 2017: Equator-crossing times (local). Microwave satellite overpass time history. Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center, accessed 20 June 2017, https://precip.gsfc.nasa.gov/times_allsat.jpg.

Nicholson, S. E., 1996: A review of climate dynamics and climate variability in eastern Africa. *The Limnology, Climatology, and Paleoclimatology of East African Lakes*, I. Johnson and E. O. Odada, Eds., Gordon and Breach Publishers, 25–56.

Nobile, A., A. Dille, E. Monsieurs, J. Basimike, T. Mugaruka Bibentyo, N. d'Oreye, F. Kervyn, and O. Dewitte, 2018:

Multi-temporal DInSAR to characterise landslide ground deformations in a tropical urban environment: Focus on Bukavu (DR Congo). *Remote Sens.*, **10**, 626, https://doi.org/10.3390/rs10040626.

Peel, M. C., B. L. Finlayson, and T. A. McMahon, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, **11**, 1633–1644, https://doi.org/10.5194/hess-11-1633-2007.

Poméon, T., D. Jackisch, and B. Diekkrüger, 2017: Evaluating the performance of remotely sensed and reanalysed precipitation data over West Africa using HBV light. *J. Hydrol.*, **547**, 222–235, https://doi.org/10.1016/j.jhydrol.2017.01.055.

Prakash, S., A. K. Mitra, D. S. Pai, and A. Aghakouchak, 2016a: From TRMM to GPM: How well can heavy rainfall be detected from space? *Adv. Water Resour.*, **88**, 1–7, https://doi.org/10.1016/j.advwatres.2015.11.008.

——, ——, E. N. Rajagopal, and D. S. Pai, 2016b: Assessment of TRMM-based TMPA-3B42 and GSMaP precipitation products over India for the peak southwest monsoon season. *Int. J. Climatol.*, **36**, 1614–1631, https://doi.org/10.1002/joc.4446.

Rossi, M., S. Luciani, D. Valigi, D. Kirschbaum, M. T. Brunetti, S. Peruccacci, and F. Guzzetti, 2017: Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data. *Geomorphology*, **285**, 16–27, https://doi.org/10.1016/j.geomorph.2017.02.001.

Salio, P., M. Paula, Y. García, and D. Vila, 2015: Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmos. Res.*, **163**, 146–161, https://doi.org/10.1016/j.atmosres.2014.11.017.

Satgé, F., and Coauthors, 2016: Assessment of satellite rainfall products over the Andean plateau. *Atmos. Res.*, **167**, 1–14, https://doi.org/10.1016/j.atmosres.2015.07.012.

Scheel, M. L. M., M. Rohrer, C. Huggel, D. Santos Villar, E. Silvestre, and G. J. Huffman, 2011: Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.*, **15**, 2649–2663, https://doi.org/10.5194/hess-15-2649-2011.

Segoni, S., L. Piciullo, and S. L. Gariano, 2018: A review of the recent literature on rainfall thresholds for landslide occurrence. *Landslides*, **15**, 1483–1501, https://doi.org/10.1007/s10346-018-0966-4.

Serrat-Capdevila, A., M. Merino, J. B. Valdes, and M. Durcik, 2016: Evaluation of the performance of three satellite precipitation products over Africa. *Remote Sens.*, **8**, 836, https://doi.org/10.3390/rs8100836.

Sevruk, B., M. Ondrás, and B. Chvíla, 2009: The WMO precipitation measurement intercomparisons. *Atmos. Res.*, **92**, 376–380, https://doi.org/10.1016/j.atmosres.2009.01.016.

Shaaban, A. A., and P. E. Roundy, 2017: OLR perspective on the Indian Ocean dipole with application to East African precipitation. *Quart. J. Roy. Meteor. Soc.*, **143**, 1828–1843, https://doi.org/10.1002/qj.3045.

Shen, Y., A. Xiong, Y. Wang, and P. Xie, 2010: Performance of high-resolution satellite precipitation products over China. *J. Geophys. Res.*, **115**, D02114, https://doi.org/10.1029/2010JD014481.

Sidle, R. C., and T. A. Bogaard, 2016: Dynamic earth system and ecological controls of rainfall-initiated landslides. *Earth-Sci. Rev.*, **159**, 275–291, https://doi.org/10.1016/j.earscirev.2016.05.013.

Smets, B., D. Delvaux, K. Ann, S. Poppe, M. Kervyn, and F. Kervyn, 2016: The role of inherited crustal structures and magmatism in the development of rift segments: Insights from the Kivu basin, western branch of the East African Rift. *Tectonophysics*, **683**, 62–76, https://doi.org/10.1016/j.tecto.2016.06.022.

Souverijns, N., W. Thiery, M. Demuzere, and N. P. M. van Lizpig, 2016: Drivers of future changes in East African precipitation. *Environ. Res. Lett.*, **11**, 114011, https://doi.org/10.1088/1748-9326/11/11/114011.

Stanley, T., and D. B. Kirschbaum, 2017: A heuristic approach to global landslide susceptibility mapping. *Nat. Hazards*, **87**, 145–164, https://doi.org/10.1007/s11069-017-2757-y.

Su, F., Y. Hong, and D. P. Lettenmaier, 2008: Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *J. Hydrometeor.*, **9**, 622–640, https://doi.org/10.1175/2007JHM944.1.

Sun, X., and A. P. Barros, 2010: An evaluation of the statistics of rainfall extremes in rain gauge observations, and satellite-based and reanalysis products using universal multifractals. *J. Hydrometeor.*, **11**, 388–404, https://doi.org/10.1175/2009JHM1142.1.

Tan, J., W. A. Petersen, P. E. Kirstetter, and Y. Tian, 2017: Performance of IMERG as a function of spatiotemporal scale. *J. Hydrometeor.*, **18**, 307–319, https://doi.org/10.1175/JHM-D-16-0174.1.

Tang, G., A. Behrangi, D. Long, C. Li, and Y. Hong, 2018: Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *J. Hydrol.*, **559**, 294–306, https://doi.org/10.1016/j.jhydrol.2018.02.057.

Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, V. Levizzani, and A. De Roo, 2012: Validation of satellite-based precipitation products over sparsely gauged African river basins. *J. Hydrometeor.*, **13**, 1760–1783, https://doi.org/10.1175/JHM-D-12-032.1.

Thiery, W., E. L. Davin, H. J. Panitz, M. Demuzere, S. Lhermitte, and N. Van Lipzig, 2015: The impact of the African Great Lakes on the regional climate. *J. Climate*, **28**, 4061–4085, https://doi.org/10.1175/JCLI-D-14-00565.1.

——, ——, S. I. Seneviratne, K. Bedka, S. Lhermitte, and N. P. M. van Lizpig, 2016: Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.*, **7**, 12786, https://doi.org/10.1038/ncomms12786.

——, L. Gudmundsson, K. Bedka, F. H. M. Semazzi, P. Willems, N. P. M. van Lipzig, and S. I. Seneviratne, 2017: Early warnings of hazardous thunderstorms over Lake Victoria. *Environ. Res. Lett.*, **12**, 074012, https://doi.org/10.1088/1748-9326/aa7521.

Tian, Y., and C. D. Peters-Lidard, 2007: Systematic anomalies over inland water bodies in satellite-based precipitation estimates. *Geophys. Res. Lett.*, **34**, L14403, https://doi.org/10.1029/2007GL030787.

Vandecasteele, I., J. Moeyersons, and P. Trefois, 2010: An assessment of the spatial and temporal distribution of natural hazards in Central Africa. *African Palaeoenvironments and Geomorphic Landscape Evolution*, J. Runge, Ed., Palaeoecology of Africa, Vol. 30, CRC Press, 279–300.

Vila, D. A., L. G. G. de Goncalves, D. L. Toll, and J. R. Rozante, 2009: Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. *J. Hydrometeor.*, **10**, 533–543, https://doi.org/10.1175/2008JHM1048.1.

Villarini, G., and W. F. Krajewski, 2007: Evaluation of the research-version TMPA three-hourly 0.25° × 0.25° rainfall estimates over Oklahoma. *Geophys. Res. Lett.*, **34**, L05402, https://doi.org/10.1029/2006GL029147.

Washington, R., and Coauthors, 2006: African climate change: Taking the shorter route. *Bull. Amer. Meteor. Soc.*, **87**, 1355–1366, https://doi.org/10.1175/BAMS-87-10-1355.

Wilks, D. S., 2006: Forecast verification. *Statistical Methods in the Atmospheric Sciences*, Academic Press, 260–268.

Xu, R., F. Tian, L. Yang, H. Hu, H. Lu, and A. Hou, 2017: Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.*, **122**, 910–924, https://doi.org/10.1002/2016JD025418.

Yaduvanshi, A., P. K. Srivastava, and A. C. Pandey, 2015: Integrating TRMM and MODIS satellite with socio-economic vulnerability for monitoring drought risk over a tropical region of India. *Phys. Chem. Earth*, **83–84**, 14–27, https://doi.org/10.1016/j.pce.2015.01.006.

Yuan, F., L. Zhang, K. Wah, W. Win, L. Ren, and C. Zhao, 2017: Assessment of GPM and TRMM multi-satellite precipitation products in streamflow simulations in a data-sparse mountainous watershed in Myanmar. *Remote Sens.*, **9**, 302, https://doi.org/10.3390/rs9030302.

Zambrano-Bigiarini, M., A. Nauditt, C. Birkel, K. Verbist, and L. Ribbe, 2017: Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrol. Earth Syst. Sci.*, **21**, 1295–1320, https://doi.org/10.5194/hess-21-1295-2017.