TITLE

Evaluating TMPA rainfall over the sparsely gauged East African Rift

AUTHORS

- MONSIEURS Elise[1] – Department of Earth Sciences, Royal Museum for Central Africa, Tervuren, Belgium - Department of Geography, University of Liège, Liège, Belgium - NASA Goddard Space Flight Center, Hydrological Sciences Laboratory, Greenbelt, Maryland, USA – F.R.S.-FNRS Research Fellow
- KIRSCHBAUM Dalia Bach - NASA Goddard Space Flight Center, Hydrological Sciences Laboratory, Greenbelt, Maryland, USA
- TAN Jackson - Universities Space Research Association, Columbia, Maryland, USA - NASA Goddard Space Flight Center, Climate and Radiation Laboratory, Greenbelt, Maryland, USA
- MAKI MATESO Jean-Claude - Department of Geophysics, Centre de Recherche en Sciences Naturelles, Lwiro, DR Congo
- JACOBS Liesbet - Department of Geography and Tourism , KU Leuven, Leuven, Belgium
- PLISNIER Pierre-Denis – G.L. Eco, Grez-Doiceau, Belgium
- THIERY Wim - Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium – Institute for Atmospheric an Climate Science, Swiss Federal Institute of Technology Zurich (ETH Zurich), Zürich, Switzerland
- UMUTONI Augusta - Ministry of Infrastructure, Kigali, Rwanda
- MUSONI Didace - Rwanda Meteorology Agency, Kigali, Rwanda

[1] Corresponding author: elise.monsieurs@africamuseum.be; +32 2769 5456

- MUGARUKA BIBENTYO Toussaint – Department of Geology, Université Officielle de Bukavu, Bukavu, DR Congo
- GANZA Gloire Bamulezi- Department of Geology, Université Officielle de Bukavu, Bukavu, DR Congo
- ILOMBE MAWE Guy - Department of Geology, Université Officielle de Bukavu, Bukavu, DR Congo
- BAGALWA Luc - Department of Geophysics, Centre de Recherche en Sciences Naturelles, Lwiro, DR Congo
- KANKURIZE Clairia - Faculté des Sciences de l'Environnement, Université Polytechnique de Gitega, Gitega, Burundi
- MICHELLIER Caroline - Department of Earth Sciences, Royal Museum for Central Africa, Tervuren, Belgium
- STANLEY Thomas - Universities Space Research Association, Columbia, Maryland, USA - Goddard Earth Sciences Technology and Research, Maryland, USA - NASA Goddard Space Flight Center, Hydrological Sciences Laboratory, Greenbelt, Maryland, USA
- KERVYN Francois - Department of Earth Sciences, Royal Museum for Central Africa, Tervuren, Belgium
- KERVYN Matthieu - Department of Geography, Earth System Science, Vrije Universiteit Brussel, Brussels, Belgium
- DEMOULIN Alain – F.R.S.-FNRS Senior Research Associate - Department of Geography, University of Liège, Liège, Belgium
- DEWITTE Olivier - Department of Earth Sciences, Royal Museum for Central Africa, Tervuren, Belgium

ABSTRACT

Accurate precipitation data is 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 (SRE) are of great value in such areas, but rigorous validation is required to identify the uncertainties linked to SRE 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. This data was used to validate the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite 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. TMPA Near Real-Time 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 towards regional hazard assessment, compared to data from scarcely distributed gauges with limited representativeness in the context of high rainfall variability.

1. Introduction

66 Hydrometeorological hazards triggered by extreme rainfall, such as floods and rainfall-initiated
67 landslides, pose a serious socioeconomic threat in many parts of the world and more particularly
68 in mountainous areas (Kjekstad and Highland 2009; Jacobs et al. 2016a; Alfieri et al. 2017; Kumar
69 et al. 2017). Moreover, in the context of ongoing climate change, it is estimated that rainfall
70 extremes may intensify, particularly in the tropics (IPCC 2013; Gariano and Guzzetti 2016;
71 Souverijns et al. 2016, Thiery et al. 2016), which would also increase the vulnerability of the
72 population (Washington et al. 2006). Accurate rainfall data are fundamental to better
73 characterizing extremes as well as ultimately mitigating the weather-related hazards in areas such
74 as tropical Africa. Hydrometeorological hazards (herein referring to both hydrological and
75 hydrologically-triggered hazards such as landslides), can be related to high intensity - short
76 duration events and/or long duration rainfall (Gariano and Guzzetti 2016; Sidle and Bogaard
77 2016). High-resolution rainfall data are therefore fundamental for accurate hazard assessment
78 (Brunetti et al. 2018; Kirschbaum and Stanley 2018).

79

80 Ground-based rainfall measurements in tropical Africa are either sparse or nonexistent and often
81 include erroneous data or large gaps (Serrat-Capdevila et al. 2016; Dezfuli et al. 2017b). Main
82 reasons are the lack of African states' means and political will or interest to support such data
83 collection, as well as the relatively high cost of establishing and maintaining infrastructure in these
84 areas (Washington et al. 2006; Monsieurs et al. 2017). This problem is even more prominent in
85 mountainous areas, where rainfall presents large spatial variability due to strong topographic
86 transitions (Dinku et al. 2008) and the susceptibility to hydrometeorological hazards is generally
87 high (Sidle and Bogaard 2016). Spatial and temporal discontinuities in rainfall data from gauges
88 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, e.g., rainfall undercatch and gauge malfunctioning (Sevruk et al. 2009; Grimaldi et al. 2015, 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 (SRE) with high spatial and temporal resolution and large areal coverage provide an opportunity for regional rainfall data acquisition in remote areas. Yet, SRE 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

111 comprising inland water bodies, inducing a complicated microwave signal due to the cold water
112 and the warm land (Tian and Peters-Lidard 2007).

113

114 This paper focuses on the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite
115 Precipitation Analysis (TMPA) product for six reasons. TMPA (1) is currently amongst the most
116 widely recognized multi-satellite rainfall products (Gebregiorgis et al. 2017; Hobouchian et al.
117 2017); (2) was designed to improve tropical rainfall observations by combining microwave and
118 infrared rainfall estimates, at a spatial resolution of $0.25° \times 0.25°$ and 3-hourly temporal resolution
119 (Huffman et al. 2007); (3) provides one of the longest consistent records (1998 – present) of freely
120 available, spatially homogeneous SRE products over the tropics; (4) has been validated with
121 satisfactory results in many parts of the world (Dinku et al. 2008; Habib et al. 2009; Islam et al.
122 2012; El Kenawy et al. 2015; Munzimi et al. 2015); (5) was proven successful in several
123 hydrometeorological hazard applications (e.g., Li et al. 2009; Kirschbaum et al. 2015; Yaduvanshi
124 et al. 2015; Cullen et al. 2016; Abdelkareem 2017; Kumar et al. 2017; Thiery et al. 2017;
125 Kirschbaum and Stanley 2018); and (6) serves as important input data for high-resolution satellite-
126 based rainfall estimates such as CHIRPS (0.05°, Funk et al. 2015). While the TRMM satellite is
127 no longer operating, the multi-satellite TMPA product will continue to be produced until the
128 Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (IMERG; 2014 -
129 present, 0.1°, half-hourly) retrospective reprocessing is completed (expected by late-2018). This
130 reprocessing will provide a higher spatial and temporal resolution from 2000 to present, and will
131 extend the latitudinal coverage to 65°N-S (Huffman et al. 2015).

132

133 Rigorous validation is necessary to characterize uncertainties in the SRE records in order to more
134 effectively use these data for hydrometeorological hazard modelling and support local authorities
135 in risk management (Mashingia et al. 2014). Even though many current SRE data are freely
136 available, little validation has been done on a regional scale in tropical Africa, particularly at daily
137 and sub-daily time scales and in mountainous regions (Cattani et al. 2016). In this paper, we
138 therefore aim to evaluate the TMPA product over a data-scarce region in tropical Africa with
139 complex topography. We focus on the western branch of the East African Rift, a region known for
140 being highly susceptible to hydrometeorological hazards (Maki Mateso and Dewitte 2014; Jacobs
141 et al. 2016a, 2017; Monsieurs et al. 2017, 2018; Stanley and Kirschbaum 2017). The paper is
142 organized as follows: section two describes the validation region and the data used for analysis,
143 section three outlines the validation approach, section four provides results of this approach, and
144 the final section discusses the findings. A list with acronyms is given in Appendix 1.

145

146 **2. Setting and data description**

147 *a) Study area*

148 The validation region stretches over the western branch of the East African Rift from Lake
149 Tanganyika up to the Rwenzori Mountains (Fig. 1). The regional climate is driven by three
150 principal surface airstreams and two major surface convergence zones (Fig. 1). A comprehensive
151 description of the complex system of climate drivers in Equatorial Africa is given by Dezfuli
152 (2017a). The region has a bimodal rainfall regime (Fig. 2), with the first rainy season starting in
153 September (Monsieurs et al. 2018). While December to February is considered a 'dry' season
154 (Nicholson 1996), results from the average monthly TMPA data in Fig. 2 show that there are still

155 over 100 mm of average rainfall, whereas the period from June to August can clearly be
156 distinguished as a dry season

157

158 Rainfall anomalies are related to the strong interannual variability of the ITCZ (Souverijns et al.
159 2016), El Ninõ-Southern Oscillation and Indian Ocean Dipole events (Behera and Yamagate 2001;
160 Shaaban and Roundy 2017). This is reflected by the large standard deviations of monthly rainfall
161 in Fig. 2. In addition, the local climate is strongly modulated by the complex rift topography
162 (Jacobs et al. 2016a; Smets et al. 2016), and the presence of large lakes (Thiery et al. 2015;
163 Docquier et al. 2016). Consequently, there is large rainfall variability across the study area. The
164 complex topography, presence of lakes, and variable rainfall regimes within the study area
165 represent many of the typical features that pose difficulties for estimating rainfall from satellite
166 observations.

167

168 Several types of natural hazards threaten the densely populated areas in the study region, including
169 hydrological (e.g., flood) and geophysical hazards (e.g., earthquake, volcanic eruption, landslide),
170 and their interactions (Jacobs et al. 2016b, 2017; Michellier et al. 2016; Delvaux et al. 2017; Thiery
171 et al. 2017; Nobile et al. 2018). Similar to rain gauge data collection, collecting information on
172 hydrometeorological hazards is tedious in the study area (Monsieurs et al. 2017, 2018) and beyond
173 the scope of the present study. Fig. 2 provides a summary of the most extensive landslide hazard
174 inventory currently available in the region, which is an updated version of the inventory compiled
175 by Monsieurs et al. (2018) comprising 199 landslide events with known location and date over a
176 time span of 50 years (1968 – 2018). There is a clearly observable signal related to the wetting of

177 the soil throughout the two rainy seasons and subsequent landslide occurrences, which reach a
178 peak in May.

179

180 *b) Data description*

181 1) RAIN GAUGE DATA

182 Gauge data with a sampling frequency of one day or better were collected for 24 gauges from a
183 variety of sources, including international research projects, universities, and local research,
184 religious, and governmental institutions. In addition, a gauge network was installed and maintained
185 as part of the RESIST and AfReSlide projects since 2014 (http://resist.africamuseum.be/;
186 http://afreslide.africamuseum.be/) which comprises 10 gauges in the Rwenzori Mountains
187 (Uganda) and 12 on the Rift flanks in DR Congo. These gauges have a temporal resolution ranging
188 between 30 min and one hour. The 46 gauges in total (herein referred to as 'gauge network') cover
189 the latitudinal range of the study area; however an optimal distribution of gauges in the longitude
190 is mainly hampered by inaccessibility due to low levels of security west of Lake Kivu (Fig. 3).
191 Minimum and maximum rain gauge elevations in this network are 664 m and 2435 m, with a mean
192 of 1600 m and a standard deviation of 428 m. The temporal extent of the gauge data over the
193 TMPA record used in this study (1/1/1998 – 1/1/2018) is presented in Fig. 4. This gauge network
194 currently represents the most extensive data compilation from gauges with a temporal resolution
195 of 24 h or better over the study area. Moreover, validation of SRE in Equatorial Africa has, to the
196 authors' knowledge, never been performed with a comparable rain gauge density at the present
197 spatiotemporal resolution. However, continuous time series are almost non-existent because of
198 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 Supplementary Material.

Six gauges coincide with gauges in the Global Precipitation Climatology Centre (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.

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 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-S every 3 hours (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-hour window, i.e., not the actual 3-hourly average rain rate (Villarini and Krajewski 2007) over an area of ~775 km² (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 1/1/1998 to present, over 50°N-S, with a 0.25° x 0.25° - 3 h 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/hour (Huffman et al. 2007). Due to this post-processing, 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 post-processing 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-S with experimental data currently extending from 50 to 60° N-S. Further details can be found in Kummerow et al. (2000, 2015) and Huffman et al. (2007, 2017).

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

247

248 **3. Methodology**

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

256

257 *a. Validation statistics*

258 In this work three validation approaches are used. All statistical tests are described in Table 1,
259 based on the overview of standard and diagnostic validation methods by Wilks (2006) and Ebert
260 (2007). The entire validation was performed using the R open-source software, release 3.3.2
261 (http://www.r-project.org/). The first approach comprises visual comparison of the rainfall value
262 distributions, using quantile-quantile (QQ) plots. The second validation approach measures

263 TMPA's accuracy for continuous variables including rainfall amount and intensity, while
264 normalizing by the gauge-measured rainfall rate of the respective period. This approach relies on
265 a combination of accuracy metrics to evaluate scatter, or random error, by use of the Pearson
266 correlation coefficient (COR); error direction, by Normalized Mean Error (NME; also known as
267 mean bias); error magnitude, by Normalized Mean Absolute Error (NMAE); and importance of
268 extremes, by Normalized Root Mean Square Error (NRMSE). Lastly, TMPA's accuracy in rainfall
269 detection is evaluated based on a contingency table (Table 2). A comprehensive statistic
270 summarizing the contingency table is the Heidke Skill Score (HSS) (Heidke 1926; Wilks 2006).
271 HSS measures how well estimates perform compared to random chance, with negative values
272 indicating a worse performance than random chance, positive values indicating a better-than-
273 random performance and a value of one corresponding to perfect skill. For the purpose of
274 evaluating SRE for hydrometeorological hazard applications, additional highly relevant metrics
275 are the Probability of Detection (POD) and the Probability of False Alarm (POFA) (Martelloni et
276 al. 2012; Rossi et al. 2017; Thiery et al. 2017). The use of the term POFA is preferred above *false*
277 *alarm ratio* (Wilks 2006) because of its common confusion with *false alarm rate* which is a
278 different metric (Barnes et al. 2009). POD refers to TMPA's ability to correctly identify rain
279 occurrence, i.e., separate wet and dry days or, if rain thresholds are imposed to evaluate TMPA's
280 efficiency in identifying heavy rains, separating days drier or wetter than the threshold. POFA
281 works the same way with respect to thresholds.

282

283 *b. Spatiotemporal validation scales*

284 Validation has been conducted for three temporal rainfall accumulation periods: 3-hourly (TMPA
285 native resolution), daily, and monthly. We focus mostly on the daily time scale to utilize the entire
286 gauge dataset, since 35% of the gauges have no sub-daily information. Furthermore, daily
287 resolution is highly relevant for regional hazard model calibration due to the dearth of information
288 on the exact time of the occurrence of hydrometeorological hazards (e.g., Kirschbaum et al. 2015),
289 especially in the context of Central Africa (Jacobs et al. 2016a; Monsieurs et al. 2018).
290 Mountainous environments, however, are characterized by high rainfall variability and induce
291 accelerations of streamflow volume concentration that might cause hydrometeorological hazards
292 such as flash floods (Devrani et al. 2015). Therefore, we validated TMPA on a 3-hourly resolution
293 for the rainy seasons in 2016, i.e. March – May and September – November, using 13 gauges
294 from the network we maintain with sub-daily temporal resolution and which have data over the
295 selected period. Lastly, TMPA performance is tested on a monthly scale, because accumulated
296 antecedent rainfall conditions are proven to play a crucial role for hazards such as landslides
297 (Segoni et al. 2018). However, monthly TMPA accumulations are strongly influenced by GPCC.
298 Only continuous metrics are computed at the monthly scale as detection of rainfall for monthly
299 accumulations is irrelevant to SRE skill evaluation. Months comprising gaps have been excluded
300 from this analysis.

301

302 We evaluated the different temporal rainfall accumulations by season, ground conditions, and
303 rainfall characteristics. These factors are known to affect the errors in SREs (Kidd and Huffman
304 2011). Two elevation classes (500 – 1700; 1700 – 2500 m a.s.l.) are chosen such that the amount
305 of gauge data in each class is approximately equal. A threshold of 100-mm average monthly
306 rainfall was chosen to define wet versus dry periods and test the impact of seasonality on TMPA's

307 performance (Fig. 2). To evaluate TMPA's sensitivity to rainfall intensity, data were grouped into
308 ten intensity classes as measured by the gauges, and adapted from Kim et al. (2017).

309

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

321

322 *c. Point-to-pixel approach*

323 The sparse and temporally and spatially heterogeneous gauge coverage does not allow for an
324 accurate rainfall interpolation, especially given the strong topographic gradients within the
325 domain. Therefore, gauge data are not extrapolated and aggregated to TMPA's resolution, as
326 recommended by Chen and Knutson (2008). Instead we apply the point-to-pixel validation
327 approach (Islam et al. 2010; Thiemig et al. 2012; El Kenawy et al. 2015). A severe limitation to

328 this approach is the discrepancy between local-scale gauge data and spatially averaged TMPA data
329 in the context of local convective storms and orographic rainfall.

330 Local rainfall variability brings uncertainty in the gauge data used as a reference to validate TMPA.
331 This source of uncertainty is especially relevant for extreme rainfall analyses (Chen and Knutson
332 2008; Sun and Barros 2010). To evaluate the gauge data uncertainty, we study the correlation and
333 standard deviation of the daily rainfall accumulations from the simultaneously but independently
334 measuring gauges in a pixel. Five TMPA pixels contain two gauges and one pixel contains four
335 gauges with overlapping time series. There are a total of 3,599 (ResV), 3651 (RT) overlapping
336 days between intra-pixel gauges over these six pixels, with an average overlap per pixel of 604
337 days. The average distance between the gauges in one pixel is 18.72 (±7.28) km. The measures on
338 local rainfall variability within a pixel are compared with TMPA uncertainty in the respective
339 pixel, calculated by the mean Residual Standard Error (ResSE) of TMPA. ResSE (mm/day) is
340 calculated as the square root of the mean squared residual in a simple linear model for TMPA with
341 gauge data as independent variable. We have to take into account that ResSE is affected by the
342 uncertainty of both TMPA and gauge data.

343

344 *d. Analyses of extremes*

345 A final analysis considers TMPA's performance in capturing extreme rainfall events. The hundred
346 highest daily rainfall events for TMPA and gauge data over the same spatial and temporal domain
347 are evaluated, of which the ten highest records are related to reported hydrometeorological hazards.
348 Information on hydrometeorological hazards is collected using the following principal sources:
349 Monsieurs et al. (2018) (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 March 2000 to January 2018 in order to be able to compare for each event the three datasets, i.e. gauge, TMPA ResV and RT.

4. Results

Validation results relevant for using TMPA in hydrometeorological hazard assessment are firstly presented for the daily temporal resolution, allowing a broad picture of TMPA's performance over multiple spatiotemporal scales. This is followed by the 3-hourly and monthly validation results for the selected gauges and time periods. Lastly, 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 analyse 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 3A). 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 data sets at a daily scale is moderately low (COR, Table 3A). 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

372 detection skills are relatively good (Table 3A). However, these metrics have a relatively high
373 standard deviation, indicating a significant variation in performance among the gauges.
374
375 When grouping the validation by elevation, we are able to reveal an improvement of the correlation
376 between TMPA RT and the gauges, which outperformed TMPA ResV for both elevation
377 categories. Average daily TMPA rainfall underestimation is lower for low-altitude sites (Table
378 3B). The other statistics do not significantly vary for the two elevation categories (Table 3B),
379 pointing to no obvious topographic control on the correlation and detection skills on a regional
380 scale.
381
382 Rainfall underestimation by TMPA, error magnitude and importance of outliers tend to increase
383 during the drier months of the year (Table 3C). Even though HSS is similar for wet and dry months,
384 the POD is higher and POFA lower for the wet months (Table 3C). The higher correlation for the
385 relatively dry months is related to the many zero-rain days.
386
387 When examining the performance of TMPA as a function of daily rainfall intensities, we found
388 that all rainfall intensities above 5 mm/day are underestimated by TMPA with increasing
389 magnitudes for higher rainfall intensities (Table 3D). We note that at the lowest intensities,
390 disproportionally high normalized errors are due to the normalization against low gauge averages.
391 In terms of rain detection, the HSS decreases for higher rainfall intensities but remains positive at
392 all rain intensities, indicating a declining performance in rain detection but still better than random
393 chance (Table 3D). The POD has relatively high values at 0 and 1 mm/day, 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 in the RT version just because no values appear above this threshold due to strong underestimation (Table 3D).

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 4A), with TMPA ResV consistently showing lower values. TMPA's performance for pixels characterized with a complex topography is also significantly lower (Table 4A) 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 4A).

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 the Pearson correlation coefficient, HSS, POD, and POFA (Table 4B). Regarding the rainfall intensity control, HSS and POD are found to degrade most severely with increasing intensities in complex topographic conditions (Table 4C).

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 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 (NW: UG6, UG9; SE: UG3, UG4, Fig. 3), knowing that the SE flank is the rain shadow side (Jacobs et al. 2016). Validation results for COR, NME, HSS and POFA are significantly better for the rainy side of the mountain (NW). Rainfall is more severely underestimated on the rain shadow (SE) side, with NME = -39.23 (ResV), -49.47 (RT), compared to the NW flank where NME = -14.91 (ResV), -26.73 (RT).

3) INTRA-PIXEL 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 correlation and standard deviation of the daily samples. With about 700 days of overlapping gauge data, analysing this single pixel is statistically significant from the temporal point of view. A representative time series was chosen over a one-month period, from 4 September to 4 October 2015, in order to show the daily rainfall variability measured by the four gauges in one pixel (Fig. 6). We observe that while the average correlation between the four 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 inter-gauge variability 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 representativity 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 data, which is for the pixel containing four gauges ~4 mm/day (ResV: 4.03; RT: 3.92). The error on TMPA RT data with respect to ground "truth" 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 therefore conclude that most of the residual error is linked to the uncertainty on gauge data and TMPA appears as a good performing indicator of relative rainfall at pixel scale in the study area. If we correct for TMPA underestimation, it may also be considered as working well in absolute terms.

Using TMPA ResV data adds uncertainty to the recognized gauge data uncertainty, with a higher 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 that the gauge-calibrated TMPA ResV product has a lower accuracy 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 the study area.

However, with one single pixel, we cannot illustrate a possible spatial variability in gauge uncertainty. Therefore we present also results for the five pixels containing two gauges, but with a prior call for caution in discussing them as two samples are far from ideal to get realistic standard deviation estimates and tend to underestimate them on average. Results show values of ~2.4 (ResV: 2.44; RT: 2.42), which are smaller compared to the standard error on residuals for TMPA in the respective pixels (ResV: 5.22; RT: 3.28) (Table 5). These results confirm that a large part

459 of the uncertainty in TMPA is related to the gauge uncertainty, and that TMPA ResV introduces
460 more uncertainty in the data compared to TMPA RT.

461 These results highlight the high local rainfall variability over the study area and suggest that
462 multiple rain gauges are needed to represent the rainfall in a 25 km x 25 km area (TMPA pixel
463 size). Hence, the limited quality of SRE validation at the TMPA resolution might be partly due to
464 the inadequacy of using individual rain gauges as reference data. This also highlights the relevance
465 of TMPA area-averaged rainfall estimates for regional hydrometeorological hazard analyses in
466 this area, as point-observations represent only a small area in the proximity of the gauge.

467

468 *b. 3-hourly and monthly resolution*

469 We also tested TMPA's performance at its 3-hourly native resolution during the rainy seasons of
470 2016. Validation results in Table 6 are fairly similar to, though slightly lower than, the results
471 obtained for daily rainfall accumulations (Table 3). Yet, results for POD are lower and POFA
472 higher (Table 6A) than those for the daily values (Table 3). When grouping for rainfall intensities,
473 we find zero probability of detection for rainfall intensities higher than 25 mm/3 hour although
474 TMPA recorded 74 events above this threshold (Table 6B). Also POFA degrades quickly for any
475 rainfall detection per 3 hours (Table 6B). Differences between the performance of both TMPA
476 products at 3-hourly resolution is small overall, with the RT version generally performing better
477 for COR, NMAE, NRMSE (Table 6).

478

479 Validation results for monthly rainfall accumulations show an expected significant improvement
480 relative to 3-hourly and daily TMPA rainfall estimates (Table 7A). In an explorative focus on

481 monthly values over four gauges from each ground context (Fig. 3, Fig. 4), TMPA encounters
482 more difficulties over complex topographic terrain (Table 7B), confirming earlier findings at daily
483 resolution (Table 4).

484

485 *c. Daily extremes related to hydrometeorological hazards*

486 In addition to the above intensity-grouped analyses, TMPA's performance for high intensity
487 events is evaluated by examining the 100 highest recorded daily rainfall events from the gauge and
488 TMPA datasets for the time period March 2000 to January 2018, referred to as 'extremes'. The
489 highest daily rainfall recorded by TMPA ResV (93 mm/day) and RT (57 mm/day), is markedly
490 lower than that of the gauges (142 mm/day) (Fig. 7A, B). Fig. 7A and 7B plot TMPA ResV and
491 RT against gauge values for the 100 most extreme gauge and TMPA ResV records respectively.
492 Results indicate an asymmetric behaviour of the recorded extremes, i.e., all gauge extremes are
493 underestimated by TMPA (Fig. 7A), whereas most (but not all) TMPA extremes overestimate
494 rainfall measured by the gauges (Fig. 7B), with a bigger bias magnitude in the former case (Fig.
495 7C) than in the latter (Fig. 7D). Fig. 7B clearly illustrates the difference between TMPA ResV and
496 RT as a result of the correction of TMPA ResV through calibration with gauges on a monthly
497 basis. While this correction results in an average lower bias between gauge and SRE values (Table
498 3A), it tends to be problematic for extreme rainfall intensities recorded by TMPA, resulting in
499 increased discrepancy between ResV and gauge data in the study area. Limiting the gauge
500 adjustment factor to a narrower range could result in better SRE for TMPA ResV.

501

When considering the top ten 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 2). The ten TMPA and gauge extremes were matched with reported hydrometeorological hazards in the corresponding pixels (Fig. 7C, D, see top ten TMPA RT extremes in Appendix 2). We found seven out of the top ten gauge extremes to be related with hydrometeorological hazards, either 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 ten TMPA extremes are also related to hydrometeorological hazards, even though some were completely missed by the corresponding field data (Fig. 7D, event 6; see RT extremes in Appendix 2).

4. 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, intra-pixel rainfall analyses have shown that gauge point observations are not necessarily representative of the average rainfall over the ~25×25 km area of a TMPA pixel at daily resolution (Fig. 6; Table 5). We found that most of the error on TMPA data with respect to gauge data is linked to the uncertainty in the gauge data (Table 5). This confirms the concerns raised by Satgé et al. (2016), who analysed 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

524 degraded if a single reference gauge is unable to record local convective storms that cause heavy
525 rains elsewhere in the TMPA pixel or, conversely, records very local heavy rain not representative
526 at the pixel scale. In principle, this issue should be less acute for higher-resolution SRE, such as
527 IMERG. The improvement obtained by averaging multiple gauge data within a pixel is in the line
528 of previous findings showing that spatiotemporal averaging substantially reduces biases and
529 improves performance (Mantas et al. 2015; Tan et al. 2017; Tang et al. 2018). Second, gauge data
530 themselves are likely underestimating rainfall due to wind undercatch, which is known as a severe
531 source for systematic bias (in the order of ~20%) in gauge-based measurements (Sevruk et al.
532 2009; Mekonnen et al. 2015).

533

534 There are very few studies that use a similar study domain and validation context (spatiotemporal
535 resolution, tropical environment, validation approach), limiting a comparison of our results with
536 existing literature. Yuan et al. (2017) validated daily TMPA ResV in the tropical monsoon area of
537 the Chindwin River basin, Myanmar (22-27°N) by applying normalized validation metrics over
538 the time period April 2014 - December 2015. There, TMPA ResV regional performance is lower
539 than in the western branch of the East African Rift, showing a maximum COR of 0.356, NMEs
540 ranging between -41.2% and +5%, POD between 0.092 and 0.299, and POFA between 0.404 and
541 0.626. A 1998-2006 study of TMPA (mainly ResV) in the La Plata basin, South America (Su et
542 al. 2008) found for the tropical Upper Paraguay area a POD and POFA of 0.36 and ~0.70,
543 respectively, for rainfall > 20 mm/day, which is better compared with values of 0.14 and 0.75 in
544 our study area (Table 3D). Potential explanatory differences between the two regions might
545 include a larger number of gauges and generally lower elevations and relief in the Upper Paraguay
546 area.

547

548 The agreement between gauge and TMPA was reduced at TMPA's native temporal resolution, i.e.,
549 3-hourly, to a significant extent only for NMAE, NRMSE, POD, and POFA, whereas the average
550 bias, Pearson correlation coefficient, and overall detection skill (HSS) were found similar to those
551 at the daily scale (Table 3A, Table 6). A study by Scheel et al. (2011) found a significantly lower
552 correlation coefficient (0.018) in sub-tropical Bolivia for TMPA at 3-hourly resolution. By
553 contrast, spatiotemporal averaging substantially reduces biases and improves performance
554 (Mantas et al. 2015; Tan et al. 2017), as confirmed by the better monthly TMPA validation results
555 in Table 7 and other studies in the same area (Adeyewa and Nakamura 2003; Munzimi et al. 2015).

556

557 The separate assessment of wet and dry periods (defined based on a 100 mm threshold for monthly
558 average rainfall, Fig. 2) revealed TMPA's performance to be lowest for several metrics (NRMSE,
559 POD, POFA) during the dry months (Table 3, Table 4). One possible cause for the poorer
560 performance during the dry season is subcloud evapotranspiration (Mashingia et al. 2014;
561 Hobouchian et al. 2017). Greater evaporation means that, for the same amount of rain produced in
562 the cloud aloft (and hence the same ice scattering signature used in the passive microwave
563 algorithm to estimate the surface rain rate), the rain that actually reaches the surface is lower,
564 leading to a higher probability of false alarms during the dry season. Serrat-Capdevila et al. (2016)
565 also found smaller errors in TMPA within areas that follow the seasonally oscillating ITCZ,
566 attributing this observation to the dependence of TMPA quality on the associated convective
567 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 4B) 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 4A). By contrast, a more rugged topography seems to impact the results towards an overall lower TMPA performance (Table 3, Table 4). In terms of local elevation, TMPA is better correlated with field data and provides more accurate rainfall estimation for lower elevations (< 1700 m a.s.l.) (Table 3B). 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 3D, Table 4C, Table 6B). Decreased detection skills and a transition from over- to underestimation by TMPA

590 when gauge-based rainfall increases is consistent with previous studies (e.g., Dinku et al. 2008;
591 Vila et al. 2009; Scheel et al. 2010; Gao and Liu 2013; Satgé et al. 2016). As a consequence, high-
592 intensity storms and extreme events are especially poorly reproduced in SREs (Fig. 7). Even
593 though TMPA uses a combination of several satellite estimates, if a sub-daily event is of short
594 duration and high intensity, the satellite microwave observations may entirely miss these peak
595 intensities given the 3+ hour revisit time in this area (Huffman at al. 2007). While infrared data
596 are used to fill in gaps between microwave overpasses, this data may also be biased by the
597 relationship between cloud top temperatures and rainfall intensities in this complex climatologic
598 and topographic setting (Huffman et al. 2007; Kidd and Huffman 2011). Improvements are
599 expected for extreme rainfall detection using IMERG, as shown already in other parts of the world
600 (Prakash et al. 2016a; Hobouchian et al. 2017; Xu et al. 2017).

601

602 A performance comparison between TMPA ResV and RT indicates that the latter underestimates
603 rainfall on average more severely over the entire range for all different spatiotemporal scales.
604 While this agrees with the overall findings of previous studies and is explained by the lack of
605 gauge-based adjustment of TMPA RT (e.g., Habib et al. 2009; Shen et al. 2010; Prakash et al.
606 2016b; Satgé et al. 2016), values for correlation, absolute error, sensitivity to outliers, and
607 categorical validation metrics show here overall better results for TMPA RT compared to TMPA
608 ResV (Table 3, Table 4, Table 6, Table 7). In contrast to what has been argued in the previous
609 studies, this better performance of RT data is most likely related to the fact that, due to the sparsity
610 of gauge data in Central Africa, the GPCC-based monthly correction applied to TMPA ResV does
611 not adequately represent the rainfall variability over the study area. Regarding the use of TMPA
612 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 (Fig. 7B, D). In addition, the intra-pixel 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.

5. Conclusion

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 2000 to present, will be available 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 mm/day. TMPA

635 performance decreases in predictive power during the dry months. Validation results for 3-hourly
636 and daily TMPA are found very similar, whereas the performance significantly increases for
637 monthly rainfall accumulations. Trade-offs for the short latency of TMPA RT were found to be
638 small, showing overall higher bias with gauge data, but better rainfall detection skills and lower
639 absolute errors compared to TMPA ResV, probably as a result of the latter's gauge-based
640 calibration. TMPA's error characteristics highlighted in this paper will improve the efficient use
641 of TMPA in hydrometeorological hazard applications. Especially in the study region, TMPA is
642 indispensable to provide the regional rainfall information required in hazard assessment, owing to
643 the sparse gauge network. Despite the key challenges identified for satellite rainfall detection by
644 TMPA in the western branch of the East African Rift, and as long as IMERG products are not
645 available over a long period of time, TMPA remains one of the best sources of regional rainfall
646 information available in the study area. However, and although we recognized that weaknesses of
647 the gauge data might be partly responsible for the somewhat disappointing quality of our validation
648 results, TMPA should be used with caution for hazard assessment.

649

650 *Acknowledgments*

651 We are grateful to all institutions, research projects and people who shared rain gauge data,
652 including: the Appalachian State University (USA), Centre de Recherches en Hydrobiologie (DR
653 Congo), EAGLES (http://www.eagles-kivu.be/project.htm) and CHOLTIC projects (BELSPO,
654 Belgium), Meteo Rwanda (Rwanda), National Oceanic and Atmospheric Administration (USA),
655 Observatoire Volcanologique de Goma (DR Congo), Pères Blancs Bukavu (DR Congo),
656 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 maintain and regularly collect 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 EM. The Belgian American Education Foundation facilitated a one-year research stay for EM at the Hydrological Sciences Laboratory at NASA Goddard Space Flight Centre. JT 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

A1. List of acronyms and abbreviations used in this paper.

679 Table A1 here.

680 A2. Top ten most extreme rainfall events measured by TMPA Near Real-Time.

681 Figure A2 here.

682

683 **REFERENCES**

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

686 Adeyewa, Z. D., and K. Nakamura, 2003: Validation of TRMM radar rainfall data over major
687 climatic regions in Africa. J. Appl. Meteorol., **42**, 331–347, doi:10.1175/1520-
688 0450(2003)042<0331:VOTRRD>2.0.CO;2.

689 Alfieri, L., B. Bisselink, F. Dottori, G. Naumann, A. De Roo, P. Salomon, K. Wyser, and L.
690 Feyen, 2017: Global projections of river flood risk in a warmer world. Earth's Futur., **5**,
691 171–182, doi:10.1002/eft2.183.

692 Barnes, L. R., Schultz, D. M., Gruntfest, E. C., Hayden, M. H., and C. C. Benight, 2009:
693 Corrigendum: False alarm rate or false alarm ratio. Weather Forecast., **24**, 1452-1454,
694 doi:org/10.1175/2009WAF2222300.1.

695 Behera, S. K., and T. Yamagata, 2001: Subtropical SST dipole events in the southern Indian
696 Ocean. Geophy. Res. Lett., **28**, 327-330, doi:10.1029/2000GL011451.

697 Bitew, M.M., M. Gebremichael, 2011: Assessment of satellite rainfall products for streamflow
698 simulation in medium watersheds of the Ethiopian highlands. Hydrol. Earth Syst. Sci. **15**,
699 1147–1155, doi:10.5194/hess-15- 1147-2011.

700 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, doi: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. Hydrometeorol., **17**, 2555–2573, doi: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, doi:10.1175/2007JCLI1494.1.

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

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. African Earth Sci., **134**, 831-855, doi: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, doi: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. Geoph. Res. Lett., **42**, 5888-5894, doi:10.1080/07900627.2018.1435409.

Dezfuli, A. K., 2017a: Climate of western and central equatorial Africa. Oxford Res. Enc., doi:10.1093/acrefore/9780190228620.013.511.

Dezfuli, A. K., C. M., Ichoku, G. J., Huffman, K. I., Mohr, J. S., Selker, N., van de Giesen, R., Hochreutener, and F. O. Annor, 2017b: Validation of IMERG Precipitation in Africa. J. Hydrometeorol., **18**, 2817-2825, doi:10.1175/JHM-D-17-0139.1.

724 Dinku, and Coauthors, 2008: Validation of high-resolution satellite rainfall products over
725 complex terrain. Int. J. Remote Sens., **29**, 4097–4110, doi:10.1080/01431160701772526.
726 Dinku, T., S. J. Connor, and P. Ceccato, 2010: Comparison of CMORPH and TRMM-3B42 over
727 mountainous regions of Africa and South America. Satellite Rainfall Applications for
728 Surface Hydrology, M. Gebremichael and H. Faisal, Eds., Springer Netherlands, 193–204.
729 Docquier, D., Thiery, W., Lhermitte, S., and N.P.M. van Lipzig, 2016: Multi-year wind
730 dynamics around Lake Tanganyika. Clim. Dyn., **47**, 3191-3202, doi:10.1007/s00382-016-
731 3020-z.
732 Ebert, E. E., 2007: Methods for verifying satellite precipitation estimates. Measuring
733 precipitation from space, Springer Netherlands, 345–356.
734 El Kenawy, A. M., J. I. Lopez-Moreno, M. F. McCabe, and S. M. Vicente-Serrano, 2015:
735 Evaluation of the TMPA-3B42 precipitation product using a high-density rain gauge
736 network over complex terrain in northeastern Iberia. Glob. Planet. Change, **133**, 188–200,
737 doi:10.1016/j.gloplacha.2015.08.013.
738 Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations — a new
739 environmental record for monitoring extremes. Sci. data, **2**, 150066,
740 doi:10.1038/sdata.2015.66.
741 Gao, Y. C., and M. F. Liu, 2013: Evaluation of high-resolution satellite precipitation products
742 using rain gauge observations over the Tibetan Plateau. Hydrol. Earth Syst. Sci., **17**, 837–
743 849, doi:10.5194/hess-17-837-2013.
744 Gariano, S. L., and F. Guzzetti, 2016: Landslides in a changing climate. Earth Sci. Rev., **162**,
745 227–252, doi:10.1016/j.earscirev.2016.08.011.
746 Gasse, F., Chalié, F., Vincens, A., Williams, M. A., and D. Williamson, 2008: Climatic patterns

747 in equatorial and southern Africa from 30,000 to 10,000 years ago reconstructed from
748 terrestrial and near-shore proxy data. Quaternary Sci. Rev., **27**, 2316-2340,
749 doi:10.1016/j.quascirev.2008.08.027.
750 Gebregiorgis, A. S., P. E. Kirstetter, Y. E. Hong, N. J. Carr, J. J. Gourley, W. Petersen, and Y.
751 Zheng, 2017: Understanding overland multisensor satellite precipitation error in TMPA-RT
752 products. J. Hydrometeorol., **18**, 285–306, doi:10.1175/JHM-D-15-0207.1.
753 Grimaldi, S., A. Petroselli, L. Baldini, and E. Gorgucci, 2015: Description and preliminary
754 results of a 100 square meter rain gauge. J. Hydrol., **556**, 827-834,
755 doi:10.1016/j.jhydrol.2015.09.076.
756 Habib, E., A. Henschke, and R. F. Adler, 2009: Evaluation of TMPA satellite-based research and
757 real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana,
758 USA. Atmos. Res., **94**, 373–388, doi:10.1016/j.atmosres.2009.06.015.
759 Heidke, P., 1926: Berechnung des Erfolges und der Giite der Windstarkevorhersagen im
760 Sturmwarnungsdienst. Geogr. Ann., **8**, 310–349, doi:10.2307/519729.
761 Hobouchian, M. P., P. Salio, Y. García, D. Vila, and R. Garreaud, 2017: Assessment of satellite
762 precipitation estimates over the slopes of the subtropical Andes. Atmos. Res., **190**, 43–54,
763 doi:10.1016/j.atmosres.2017.02.006.
764 Hou, A. Y., and Coauthors, 2014: The global precipitation measurement mission. Bull. Am.
765 Meteorol. Soc., **95**, 701–722, doi:10.1175/BAMS-D-13-00164.1.
766 Huffman, G. J., and D. T. Bolvin, 2014: TRMM and other data precipitation data set
767 documentation. NASA/GSFC, 42.
768 ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf (Accessed June 20, 2017).
769 ——, and Coauthors, 2007: The TRMM multisatellite precipitation analysis (TMPA): quasi-

770 global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol.,
771 **8**, 38–55, doi:10.1175/JHM560.1.
772 Huffman, G. J., R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM multi-satellite
773 precipitation analysis (TMPA). Satellite Applications for Surface Hydrology, F. Hossain
774 and M. Gebremichael, Eds., Springer, Dordrecht, The Netherlands, 3–22.
775 Huffman, G. J., J. David, T. Bolvin, and E. Nelkin, 2015: Integrated Multi-satellitE Retrievals
776 for GPM (IMERG). Technical documentation. NASA/GSFC Code 612:47.
777 ——, A. Pendergrass, and N. C. for A. R. Staff, 2017: The climate data guide: TRMM: Tropical
778 Rainfall Measuring Mission. https://climatedataguide.ucar.edu/climate-data/trmm-tropical-
779 rainfall-measuring-mission (Accessed January 6, 2017).
780 IPCC, 2013: The physical science basis. Contribution of working group I to the fifth assessment
781 report of the intergovernmental panel on climate change. doi:10.1017/CBO9781107415324.
782 Islam, N., S. Das, and H. Uyeda, 2010: Calibration of TRMM derived rainfall over Nepal during
783 1998-2007. Open Atmos. Sci., **4**, 12–23.
784 Islam, T., M. A. Rico-Ramirez, D. Han, P. K. Srivastava, and A. M. Ishak, 2012: Performance
785 evaluation of the TRMM precipitation estimation using ground-based radars from the GPM
786 validation network. J. Atmos. Solar-Terrestrial Phys., **77**, 194–208,
787 doi:10.1016/j.jastp.2012.01.001.
788 Jacobs, L., O. Dewitte, J. Poesen, D. Delvaux, W. Thiery, and M. Kervyn, 2016a: The Rwenzori
789 Mountains, a landslide-prone region? Landslides, **13**, 519–536, doi:10.1007/s10346-015-
790 0582-5.
791 Jacobs, L., Maes, J., Mertens, K., Sekajugo, J., Thiery, W., Lipzig, N. Van, Poesen, J., Kervyn,
792 M., and O. Dewitte, 2016b: Reconstruction of a flash flood event through a multi-hazard

793 approach: focus on the Rwenzori Mountains, Uganda. Nat. Hazards, 1–31.
794 doi:10.1007/s11069-016-2458-y.
795 ——, and Coauthors, 2017: Landslide diversity in the Rwenzori Mountains (Uganda).
796 Advancing Culture of Living with Landslides. WLF 2017, M. Mikos, N. Casagli, Y. Yin,
797 and K. Sassa, Eds., Springer, Cham, 79–86, doi:10.1007/978-3-319-53498-5_10.
798 Kidd, C., 2001: Satellite rainfall climatology: a review. Int. J. Climatol., **21**, 1041–1066,
799 doi:10.1002/joc.635.
800 Kidd, C., and G. J. Huffman, 2011: Global precipitation measurement. Meteorol. Appl., **18**, 334–
801 353, doi:10.1002/met.284.
802 Kim, K., J. Park, J. Baik, and M. Choi, 2017: Evaluation of topographical and seasonal feature
803 using GPM IMERG and TRMM 3B42 over Far-East Asia. Atmos. Res., **187**, 95–105,
804 doi:10.1016/j.atmosres.2016.12.007.
805 Kirschbaum, D. B., T. Stanley, and J. Simmons, 2015: A dynamic landslide hazard assessment
806 system for Central America and Hispaniola. Nat. Hazards Earth Syst. Sci., **15**, 2257–2272,
807 doi:10.5194/nhess-15-2257-2015.
808 Kirschbaum, D. B., T. Stanley, 2018: Satellite-Based Assessment of Rainfall-Triggered
809 Landslide Hazard for Situational Awareness. Earth's Future, **6**,
810 doi:https://doi.org/10.1002/2017EF000715.
811 Kjekstad, O., and L. Highland, 2009: Economic and social impacts of landslides. Landslides –
812 Disaster Risk Reduction, S. Kyoji and P. Canuti, Eds., Springer Berlin Heidelberg, 573–
813 587.
814 Kumar, A., A. K. L. Asthana, R. Singh, R. Jayangondaperumal, A. K. Gupta, and S. S. Bhakuni,
815 2017: Geomorphology assessment of landslide hazards induced by extreme rainfall event in

816 Jammu and Kashmir Himalaya, northwest India. Geomorphology, **284**, 72–87,
817 doi:10.1016/j.geomorph.2017.01.003.
818 Kummerow, C., and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission
819 (TRMM) after two years in orbit. J. Appl. Meteorol., **39**, 1965–1982, doi:10.1175/1520-
820 0450(2001)040<1965:TSOTTR>2.0.CO;2.
821 Kummerow, C. D., D. L. Randel, M. Kulie, N. Y. Wang, R. Ferraro, S. Joseph Munchak, and V.
822 Petkovic, 2015: The evolution of the Goddard profiling algorithm to a fully parametric
823 scheme. J. Atmos. Ocean. Technol., **32**, 2265–2280, doi:10.1175/JTECH-D-15-0039.1.
824 Li, L., and Coauthors, 2009: Evaluation of the real-time TRMM-based multi-satellite
825 precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake
826 Victoria, Africa. Nat. Hazards, **50**, 109–123, doi:10.1007/s11069-008-9324-5.
827 Maki Mateso, J., and O. Dewitte, 2014: Towards an inventory of landslide processes and the
828 elements at risk on the Rift flanks West of Lake Kivu (DRC). Geo. Eco. Trop., **38**, 137–
829 154.
830 Mantas, V. M., Z. Liu, C. Caro, and A. J. S. C. Pereira, 2015: Validation of TRMM multi-
831 satellite precipitation analysis (TMPA) products in the Peruvian Andes. Atmos. Res., **163**,
832 132–145, doi:10.1016/j.atmosres.2014.11.012.
833 Martelloni, G., S. Segoni, R. Fanti, and F. Catani, 2012: Rainfall thresholds for the forecasting of
834 landslide occurrence at regional scale. Landslides, **9**, 485–495, doi:10.1007/s10346-011-
835 0308-2.
836 Mashingia, F., F. Mtalo, and M. Bruen, 2014: Validation of remotely sensed rainfall over major
837 climatic regions in Northeast Tanzania. Phys. Chem. Earth, **67**, 55–63,
838 doi:10.1016/j.pce.2013.09.013.

839 Mekonnen, G. B., Matula, S., Doležal, F., & Fišák, J. (2015). Adjustment to rainfall
840 measurement undercatch with a tipping-bucket rain gauge using ground-level manual
841 gauges. Meteorology and Atmospheric Physics, 127(3), 241-256.
842 Michellier, C., P. Pigeon, and F. Kervyn, 2016: Contextualizing vulnerability assessment : a
843 support to geo-risk management in central Africa. Nat, **82**, 27–42, doi:10.1007/s11069-016-
844 2295-z.
845 Monsieurs, E., and Coauthors, 2017: Constraints on landslide-climate research imposed by the
846 reality of fieldwork in Central Africa. 3rd North American Symposium on Landslides,
847 Roanoke, VA, USA, 158–168.
848 Monsieurs, E., and Coauthors, 2018: Landslide inventory for hazard assessment in a data-poor
849 context: a regional-scale approach in a tropical African environment. Landslides, in press.
850 Munzimi, Y. A., M. C. Hansen, B. Adusei, and G. B. Senay, 2015: Characterizing Congo basin
851 rainfall and climate using Tropical Rainfall Measuring Mission (TRMM) satellite data and
852 limited rain gauge ground observations. J. Appl. Meteorol. Climatol., **54**, 541–555,
853 doi:10.1175/JAMC-D-14-0052.1.
854 Nelkin, E., 2017: Microwave satellite overpass time history. Equator-Crossing Times (Local).
855 Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center.
856 https://precip.gsfc.nasa.gov/times_allsat.jpg (Accessed June 20, 2017).
857 Nicholson, S. E., 1996: A review of climate dynamics and climate variability in eastern Africa.
858 The limnology, climatology, and paleoclimatology of East African lakes, I. Johnson and
859 E.O. Odada, Eds., Gordon and Breach Publishers, 25–56.
860 Nobile, A., A. Dille, E. Monsieurs, J. Basimike, T. Mugaruka Bibentyo, N. d'Oreye, F. Kervyn,
861 O. Dewitte, 2018: Multi-Temporal DInSAR to Characterise Landslide Ground

862 Deformations in a Tropical Urban Environment: Focus on Bukavu (DR Congo). Remote
863 Sens., **10**: 626, doi:10.3390/rs10040626.
864 Peel, M. C., B. L. Finlayson, and T. A. Mcmahon, 2007: Updated world map of the Köppen-
865 Geiger climate classification. Hydrol. earth Syst. Sci. Discuss., **4**, 439–473.
866 Poméon, T., D. Jackisch, and B. Diekkrüger, 2017: Evaluating the performance of remotely
867 sensed and reanalysed precipitation data over West Africa using HBV light. J. Hydrol., **547**,
868 222–235, doi:10.1016/j.jhydrol.2017.01.055.
869 Prakash, S., A. K. Mitra, D. S. Pai, and A. Aghakouchak, 2016a: From TRMM to GPM : How
870 well can heavy rainfall be detected from space ? Adv. Water Resour., **88**, 1–7,
871 doi:10.1016/j.advwatres.2015.11.008.
872 ——, ——, E. N. Rajagopal, and D. S. Pai, 2016b: Assessment of TRMM-based TMPA-3B42
873 and GSMaP precipitation products over India for the peak southwest monsoon season. Int.
874 J. Climatol., **36**, 1614–1631, doi:10.1002/joc.4446.
875 Rossi, M., S. Luciani, D. Valigi, D. Kirschbaum, M. T. Brunetti, S. Peruccacci, and F. Guzzetti,
876 2017: Statistical approaches for the definition of landslide rainfall thresholds and their
877 uncertainty using rain gauge and satellite data. Geomorphology, **285**, 16–27,
878 doi:10.1016/j.geomorph.2017.02.001.
879 Salio, P., M. Paula, Y. García, and D. Vila, 2015: Evaluation of high-resolution satellite
880 precipitation estimates over southern South America using a dense rain gauge network.
881 Atmos. Res., **163**, 146–161, doi:10.1016/j.atmosres.2014.11.017.
882 Satgé, F., and Coauthors, 2016: Assessment of satellite rainfall products over the Andean
883 plateau. Atmos. Res., **167**, 1–14, doi:10.1016/j.atmosres.2015.07.012.
884 Scheel, M. L. M., M. Rohrer, C. Huggel, D. Santos Villar, E. Silvestre, and G. J. Huffman, 2010:

885 Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the
886 Central Andes region and its dependency on spatial and temporal resolution. Hydrol. Earth
887 Syst. Sci. Discuss., **7**, 8545–8586, doi:10.5194/hessd-7-8545-2010.
888 Segoni, S., L., Piciullo, and S. L. Gariano, 2018: A review of the recent literature on rainfall
889 thresholds for landslide occurrence. Landslides, 1-19, doi:10.1007/s10346-018-0966-4.
890 Serrat-Capdevila, A., M. Merino, J. B. Valdes, and M. Durcik, 2016: Evaluation of the
891 performance of three satellite precipitation products over Africa. Remote Sens., **8**, 1–23,
892 doi:10.3390/rs8100836.
893 Sevruk, B., Ondrás, M., and B. Chvíla, 2009: The WMO precipitation measurement
894 intercomparisons. Atmos, Res., **92**, 376-380, doi: 10.1016/j.atmosres.2009.01.016.
895 Shaaban, Ahmed A., and P. E. Roundy, 2017: OLR perspective on the Indian Ocean Dipole with
896 application to East African precipitation. Q. J. Roy. Meteor. Soc., **143**, 1828-1843, doi:
897 10.1002/qj.3045.
898 Shen, Y., Xiong, A.,Wang, Y., and P. Xie, 2010: Performance of high-resolution satellite
899 precipitation products over China. J. Geophys. Res., **115**, D02114,
900 doi:10.1029/2009JD012097.
901 Sidle, R. C., and T. A. Bogaard, 2016: Dynamic earth system and ecological controls of rainfall-
902 initiated landslides. Earth Sci. Rev., **159**, 275–291, doi:10.1016/j.earscirev.2016.05.013.
903 Smets, B., D. Delvaux, K. Ann, S. Poppe, M. Kervyn, and F. Kervyn, 2016: The role of inherited
904 crustal structures and magmatism in the development of rift segments: Insights from the
905 Kivu basin, western branch of the East African Rift. Tectonophysics, **683**, 62–76,
906 doi:10.1016/j.tecto.2016.06.022.
907 Souverijns, N., Thiery, W., Demuzere, M., and N.P.M. van Lizpig, 2016: Drivers of future

908 changes in East African precipitation. Environ. Res. Lett., **11**, 114011, doi:10.1088/1748-
909 9326/11/11/114011.
910 Stanley, T., and D. B. Kirschbaum, 2017: A heuristic approach to global landslide susceptibility
911 mapping. Nat. Hazards, doi:10.1007/s11069-017-2757-y.
912 Su, F., Y. Hong, and D. P. Lettenmaier, 2008: Evaluation of TRMM Multisatellite Precipitation
913 Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. J.
914 Hydrometeorol., 622–640, doi:10.1175/2007JHM944.1.
915 Sun, X., and A. P. Barros,2010: An evaluation of the statistics of rainfall extremes in rain gauge
916 observations, and satellite-based and reanalysis products using universal multifractals. J.
917 Hydrometeorol., **11**, 388-404, doi: 10.1175/2009JHM1142.1.
918 Tan, J., Petersen, W. A., Kirstetter, P. E., & Tian, Y. (2017). Performance of IMERG as a
919 Function of Spatiotemporal Scale. Journal of Hydrometeorology, 18(2), 307-319.
920 Tang, G., A. Behrangi, D. Long, C. Li, and Y. Hong, 2018: Accounting for spatiotemporal errors
921 of gauges: A critical step to evaluate gridded precipitation products. J. Hydrol., **559**, 294-
922 306, doi:10.1016/j.jhydrol.2018.02.057.
923 Thiemig, V., and Coauthors, 2012: Validation of satellite-based precipitation products over
924 sparsely gauged african river basins. J. Hydrometeorol., **13**, 1760–1783, doi:10.1175/JHM-
925 D-12-032.1.
926 Thiery, W., E. L. Davin, H. J. Panitz, M. Demuzere, S. Lhermitte, and N. Van Lipzig, 2015: The
927 impact of the African Great Lakes on the regional climate. J. cli, **28**, 4061–4085,
928 doi:10.1175/JCLI-D-14-00565.1.

929 Thiery, W., Davin, E.L., Seneviratne, S.I., Bedka, K., Lhermitte, S., and N.P.M. van Lizpig,
930 2016: Hazardous thunderstorm intensification over Lake Victoria. Nat. Commun., **7**, 12786,
931 doi: 10.1038/ncomms12786.
932 Thiery, W., L., Gudmundsson, K., Bedka, F.H.M., Semazzi, P., Willems, N.P.M., van Lipzig,
933 and S.I., Seneviratne, 2017: Early warnings of hazardous thunderstorms over Lake Victoria,
934 Env. Res. Lett., **12**, 074012, doi: 10.1088/1748-9326/aa7521.
935 Tian, Y., and C. D. Peters-Lidard, 2007: Systematic anomalies over inland water bodies in
936 satellite-based precipitation estimates. Geophys. Res. Lett., **34**, L14403,
937 doi:10.1029/2007GL030787.
938 Vandecasteele, I., Moeyersons, J. and P. Trefois, 2010: An assessment of the spatial and
939 temporal distribution of natural hazards in Central Africa."African Palaeoenvironments and
940 Geomorphic Landscape Evolution. Series Palaeoecology of Africa, **30**, 279-300.
941 Vila, D. A., L. G. G. de Goncalves, D. L. Toll, and J. R. Rozante, 2009: Statistical evaluation of
942 combined daily gauge observations and rainfall satellite estimates over continental South
943 America. J. Hydrometeorol., **10**, 533–543, doi:10.1175/2008JHM1048.1.
944 Villarini, G., and W. F. Krajewski, 2007: Evaluation of the research-version TMPA three-hourly
945 0.25°×0.25° rainfall estimates over Oklahoma. Geophys Res Lett., **34**, L05402,
946 doi:10.1029/2006GL029147.
947 Washington, R., and Coauthors, 2006: African climate change: Taking the shorter route. Bull.
948 Am. Meteorol. Soc., **87**, 1355–1366, doi:10.1175/BAMS-87-10-1355.
949 Wilks, D. S., 2006: Forecast verification. Statistical methods in the atmospheric sciences, 260–
950 268.
951 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 Ran. J. Geophys. Res. Atmos., **122**, 910–924, doi: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**, 14–27, doi: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, doi: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. Sc., **21**, 1295-1320, doi: 10.5194/hess-21-1295-2017.

966 **TABLES**

967 TABLE 1. Validation metrics for the evaluation of satellite rainfall estimates (SRE). G refers to gauge observations, S to satellite
968 observations. $\overline{\boldsymbol{G}}$, $\overline{\boldsymbol{S}}$ are the means and δ_G, δ_S are the standard deviations of G and S respectively. N is the total amount of observations.
969 The terms in HSS, POD, and POFA are explained in Table 2. Normalized metrics are expressed in %, all metrics are unitless.

Statistical measure	Equation	Range	Optimum Score	Description
Pearson correlation coefficient	$COR = \frac{\sum_i^N (G_i - \bar{G})(S_i - \bar{S})}{\sqrt{\sum_i^N (G_i} - \bar{G})^2 \sqrt{\sum_i^N (S_i - \bar{S})^2}$	[-1, 1]	1	Reflects the degree of linear correlation between SRE and gauge observations.
Normalized Mean Error (%)	$NME = \frac{\sum_i^N (S_i - G_i)}{\sum_i^N G_i}$ x 100	[-100, +∞]	0	Average difference between SRE and gauge observations Overestimation is represented as positive bias.
Normalized Mean Absolute Error (%)	$NMAE = \frac{\sum_i^N \lvert S_i - G_i \rvert}{\sum_i^N G_i}$ x 100	[0, +∞]	0	Mean magnitude of the errors without considering their direction.
Normalized Root Mean Square Error (%)	$NRMSE = \frac{\sqrt{\frac{1}{N}\sum_i^N (S_i - G_i)^2}}{\frac{1}{N}\sum_i^N Gi}$ x 100	[0, +∞]	0	Mean magnitude of the errors, giving greater weight to the larger errors.
Heidke Skill Score	$HSS = \frac{2(ad - bc)}{(a+b)(c+d) + (a+b)(b+d)}$	[-∞, 1]	1	Fractional improvement of SRE over random forecast.
Probability of Detection	$POD = \frac{Hit}{Hit + Miss}$	[0, 1]	1	Likelihood that the event would be forecasted, given that it occurred.
Probability of False Alarm	$POFA = \frac{False\ Alarm}{Hit + False\ Alarm}$	[0, 1]	0	Proportion of forecasted events that fail to materialize.

970

971 TABLE 2. Contingency table for categorical validation metrics. SRE = satellite rainfall estimate. Letters a-d are the same as in Table 1

972 for HSS.

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

973

974 TABLE 3. Summary of validation results for daily TMPA Research Version (ResV) and Near
975 Real-Time (RT) products (both Version 7). Validation statistics include: Pearson correlation
976 coefficient (COR), Normalized Mean Error (NME), Normalized Mean Absolute Error (NMAE),
977 Normalized Root Mean Square Error (NRMSE), Heidke Skill Score (HSS), Probability of
978 Detection (POD), and Probability of False Alarm (POFA). The number of days used for analyses
979 in each category is given under 'Days'. A: Regionally calculated metrics. B: Elevation groups,
980 given in m a.s.l.. C: Groups based on average monthly rainfall (Fig. 2). D: Intensity groups, all in
981 mm/day (bold numbers referring to the thresholds used for the calculation of HSS, POD, and
982 POFA). Standard deviations (Stdev) for regionally calculated metrics are presented between
983 brackets.

		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	92941	87357
	(Stdev)	(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	49226	47264
	>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	43715	40093
C.	>100 mm/month	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	69592	65450
	<100 mm/month	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	23349	21907
D.	**0** mm/day	NA	NA	NA	NA	NA	NA	NA	NA	0.40	0.39	0.65	0.64	0.32	0.32	51160	47901
	<**1** mm/day	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	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	13980	13224
	5-**10** mm/day	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	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	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	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	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	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**	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

TABLE 4. Evaluation of daily TMPA ResV and RT products in different contexts on the ground. Mixed: pixels that include large water bodies, Co. Topogr.: complex topographic terrain, Low: relatively low terrain. For each of these contexts four gauges have been selected (indicated in Fig. 3, Fig. 4). A: Averaged metrics per category. B: Elevation groups, given in m a.s.l.. C: Groups based on average monthly rainfall (Fig. 2). D: Intensity groups, all in mm/day (bold numbers referring to the thresholds used for the calculation of HSS, POD, and POFA).

			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
		C. 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	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	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	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	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	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	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	NA	NA	NA	NA	NA	NA	NA	NA	0.34	0.34	0.60	0.60	0.30	0.30	1197	1197
		0-**5** mm/day	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	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	-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	NA	NA	NA	NA	NA	NA	NA	NA	0.29	0.29	0.65	0.65	0.35	0.36	2233	2314
		0-**5** mm/day	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	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	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	NA	NA	NA	NA	NA	NA	NA	NA	0.44	0.43	0.66	0.64	0.37	0.37	2803	2803
		0-**5** mm/day	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	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	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 inter-gauge 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 standard deviation gauge data' is the average standard deviation (mm/day) between the gauge measurements in a pixel; and 'TMPA Residual Standard Error' is the square root of the mean squared residual (mm/day) in the linear model for TMPA with gauge data as independent variable.

	Mean inter-gauge correlation	Mean gauge-TMPA correlation	Correlation mean gauge data - TMPA	Mean standard deviation gauge data	TMPA Residual Standard Error
ResV					
One 4-gauge pixel 699 days	0.293	0.335	0.472	4.03	4.74
Five 2-gauge pixels 2900 days	0.399	0.399	0.473	2.44	5.22
RT					
One 4-gauge pixel 724 days	0.296	0.336	0.470	3.92	3.86
Five 2-gauge pixels 2927 days	0.395	0.385	0.458	2.42	3.28

1010 TABLE 6. Evaluation of TMPA ResV and RT products at 3-h resolution. Analyses are based on data from the 13 gauges we maintain (indicated
1011 in Fig. 4) during two rainy seasons in 2016 when available (March-May, September-November). The number of observations used for analyses
1012 in each category is given under 'Time Stamps'. A: Averaged results for the selected gauges and time period. B: Intensity groups, all in mm/day
1013 (bold numbers referring to the thresholds used for the calculation of HSS, POD, and POFA).

1014

		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	16952	16952
	(Stdev)	(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 hr	NA	NA	NA	NA	NA	NA	NA	NA	0.37	0.37	0.41	0.41	0.46	0.46	14006	14006
	>0-**5** mm/3 hr	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 hr	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 hr	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

1015

1016 TABLE 7. Evaluation of monthly TMPA ResV and RT products. The number of months used for analyses in each category is given under 'Months'.
1017 A: Regionally calculated metrics (all gauge data in Fig. 3 are used). B: Evaluation in different contexts on the ground. Mixed: pixels that include large
1018 water bodies, Co. Topogr.: complex topographic terrain, Low: relatively low terrain. For each of these contexts four gauges have been selected
1019 (indicated in Fig. 3, Fig. 4).

1020

		COR		NME		NMAE		NRMSE		Months	
		ResV	RT	ResV	RT	ResV	RT	ResV	RT	ResV	RT
A.	Regional (Stdev)	0.70 (0.18)	0.69 (0.19)	-15.64 (18.37)	-41.09 (14.10)	38.92 (10.52)	50.21 (9.68)	52.97 (14.53)	65.57 (12.63)	3008	2829
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

1021

1022 Table A1. List of acronyms and abbreviations used in this paper.

COR	Pearson Correlation Coefficient
DR Congo	Democratic Republic of Congo
GPCC	Global Precipitation Climatology Centre
HSS	Heidke Skill Score
IMERG	Integrated Multi-satellitE 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 Multi-satellite Precipitation Analysis
TRMM	Tropical Rainfall Measuring Mission

1023

1024

1025

1026

1027

1028

1029

1030

1031

FIGURES

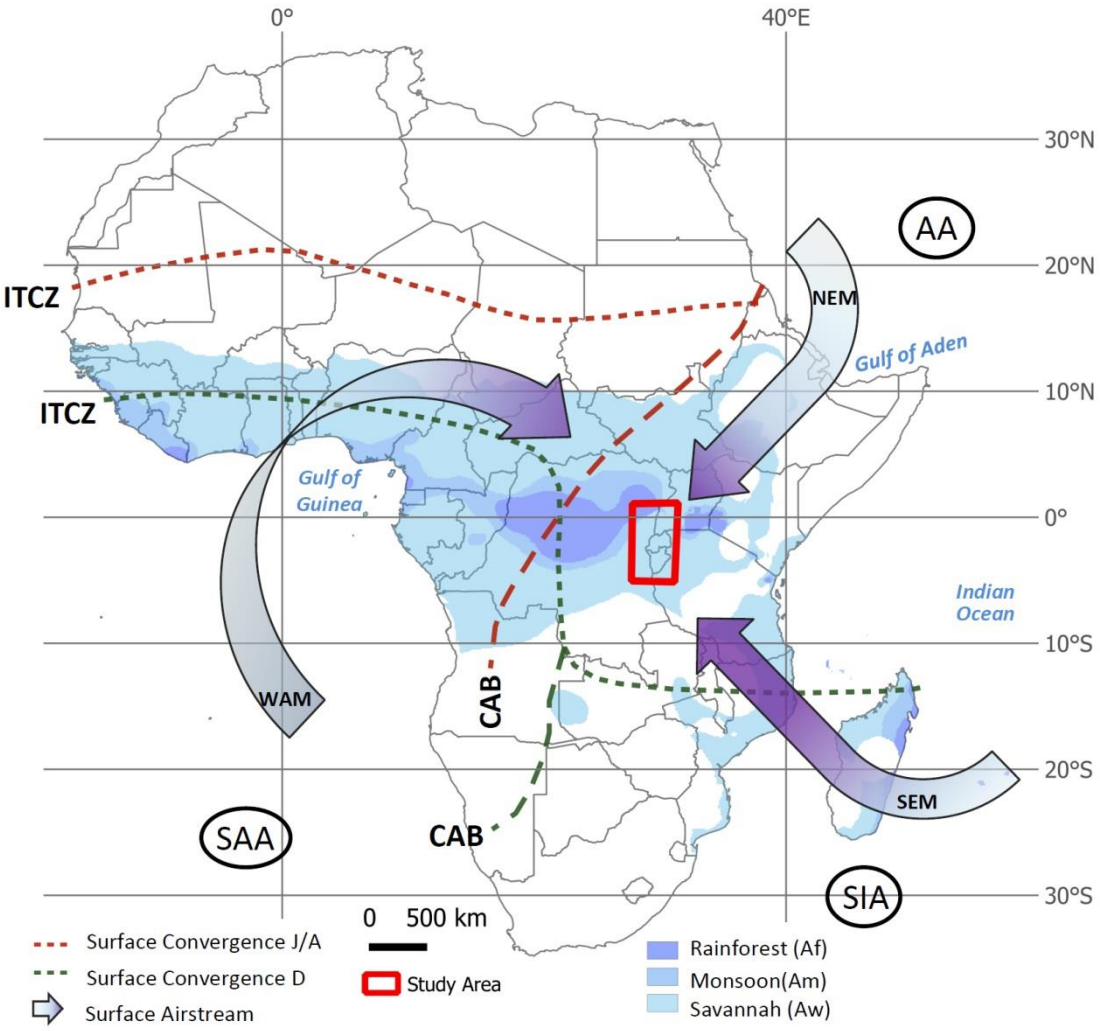
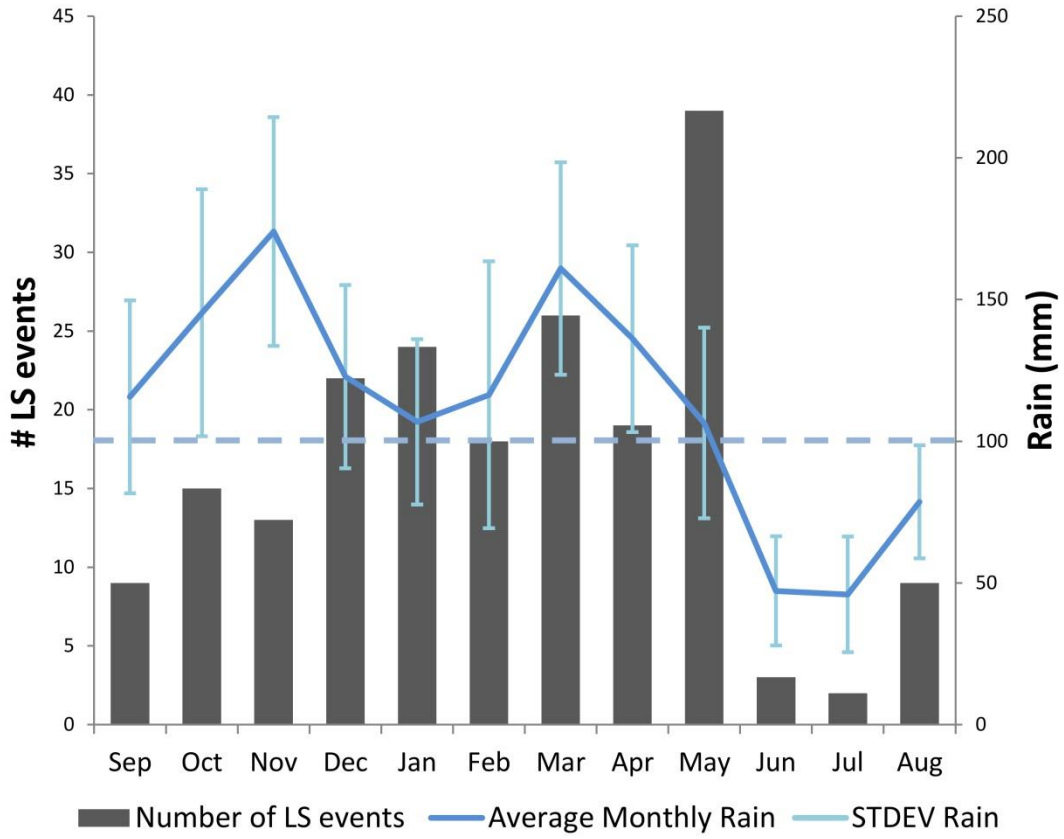


FIG. 1. Surface pressure patterns, airstreams and convergence zones that affect the climate in the study area (modified after Nicholson 1996; Gasse et al. 2008). SAA = South Atlantic Anticyclone, AA = Arabian Anticyclone, SIA = South Indian Anticyclone, WAM = West African Monsoon, NEM = northerly East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

1041



1042

1043 FIG. 2. Rainfall seasonality and distribution of landslide events (LS) in the study area. The
1044 dashed line highlights the 100 mm monthly average rainfall threshold that was used to group
1045 the analyses in wet and dry months. The presented mean monthly rainfall is based on 20 years
1046 (1998-2018) TMPA 3B42 daily data. Modified after Monsieurs et al. (2018).

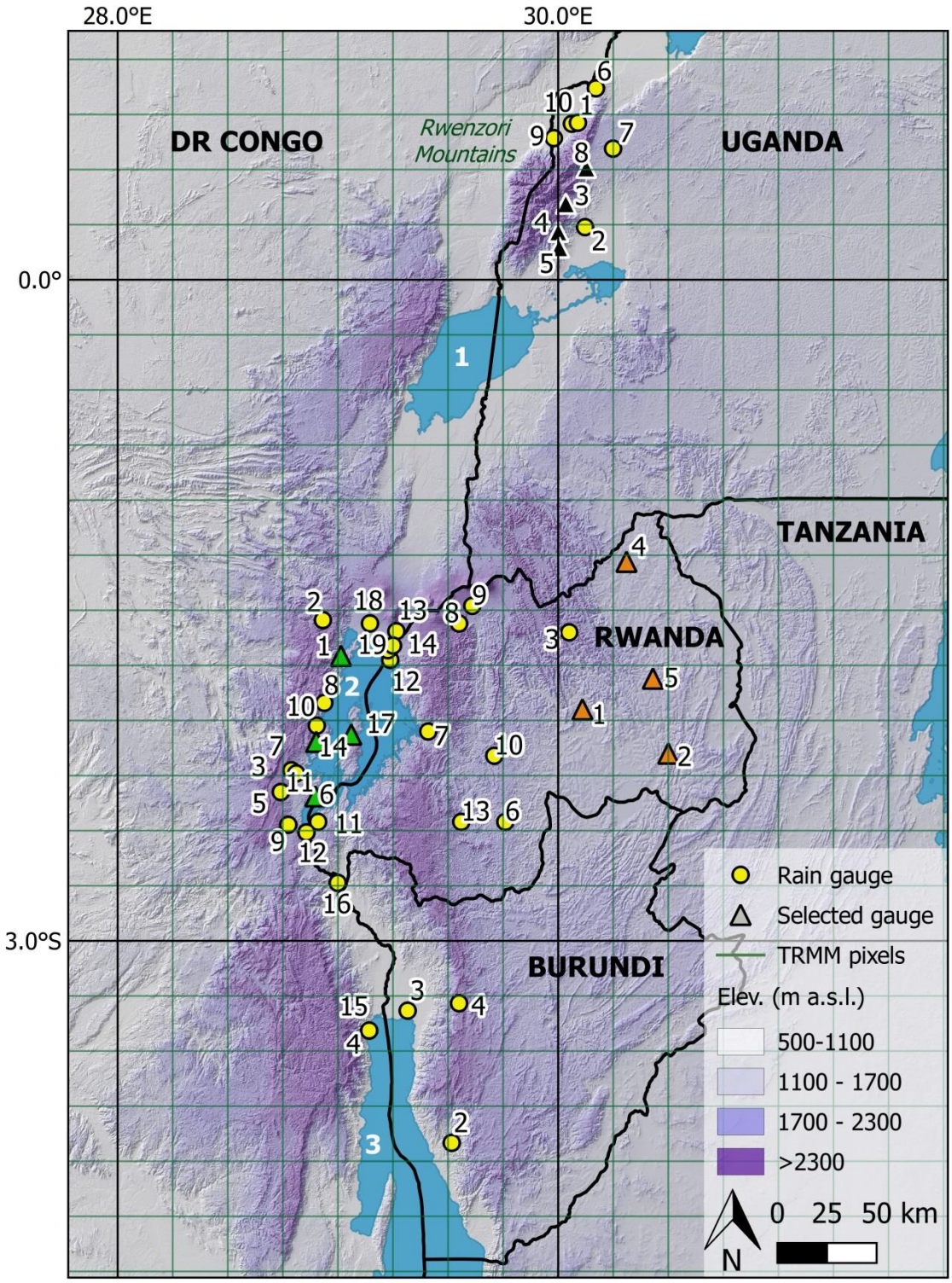


FIG. 3. Spatial distribution of rain gauges in the study area with a temporal resolution of 24 h or better. Gauges represented as triangles refer to the selected gauges for the validation in distinct contexts on the ground: Black = Topographic complex terrain, Green = Presence of large water

1065 bodies, Orange = Relative low-altitude continental environment. Numbers and colours refer to
1066 gauge codes in Fig. 4. Numbers in the lakes: 1 = Lake Edward, 2 = Lake Kivu, 3 = Lake
1067 Tanganyika. Background hillshade SRTM (90m).

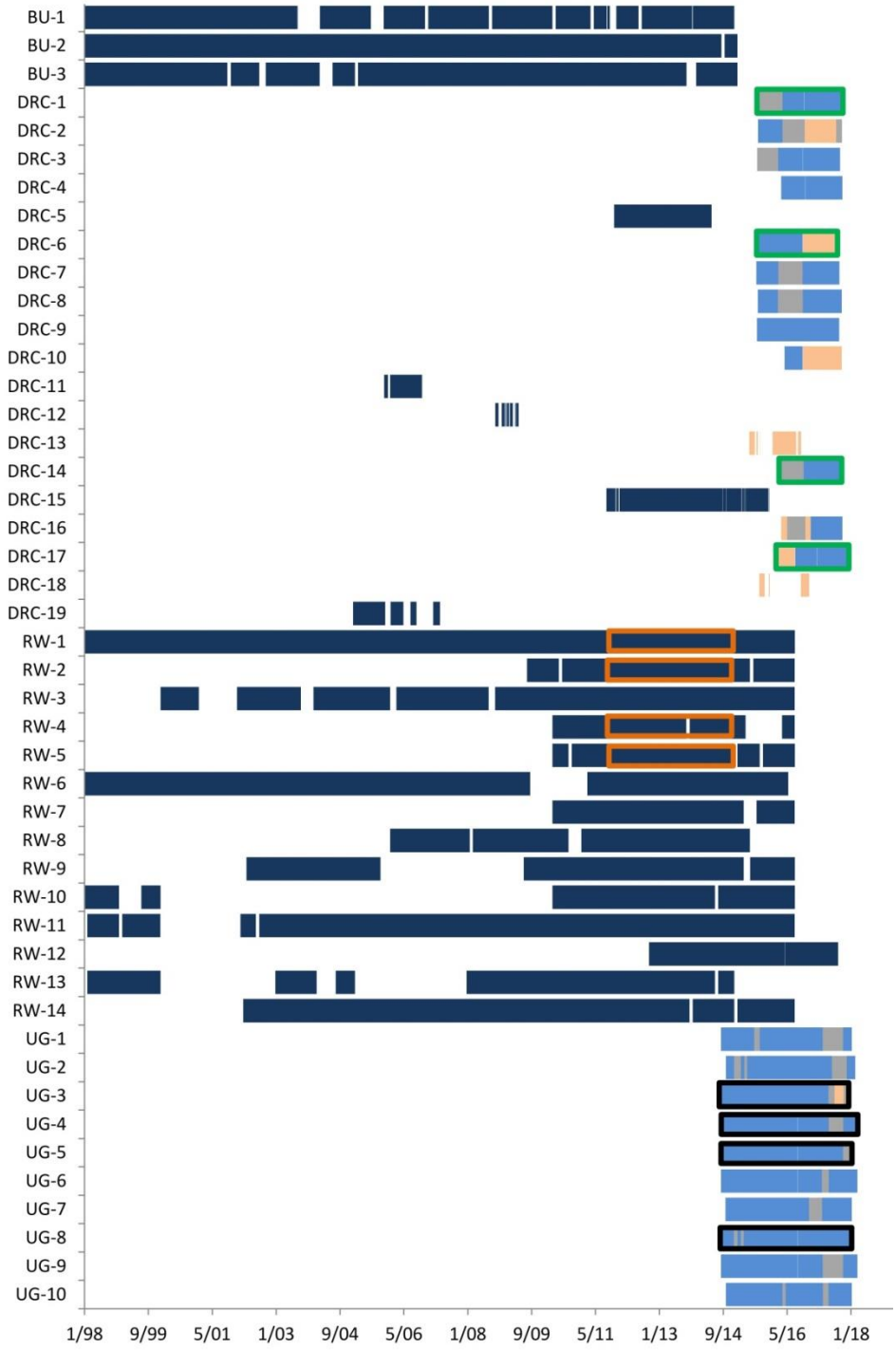
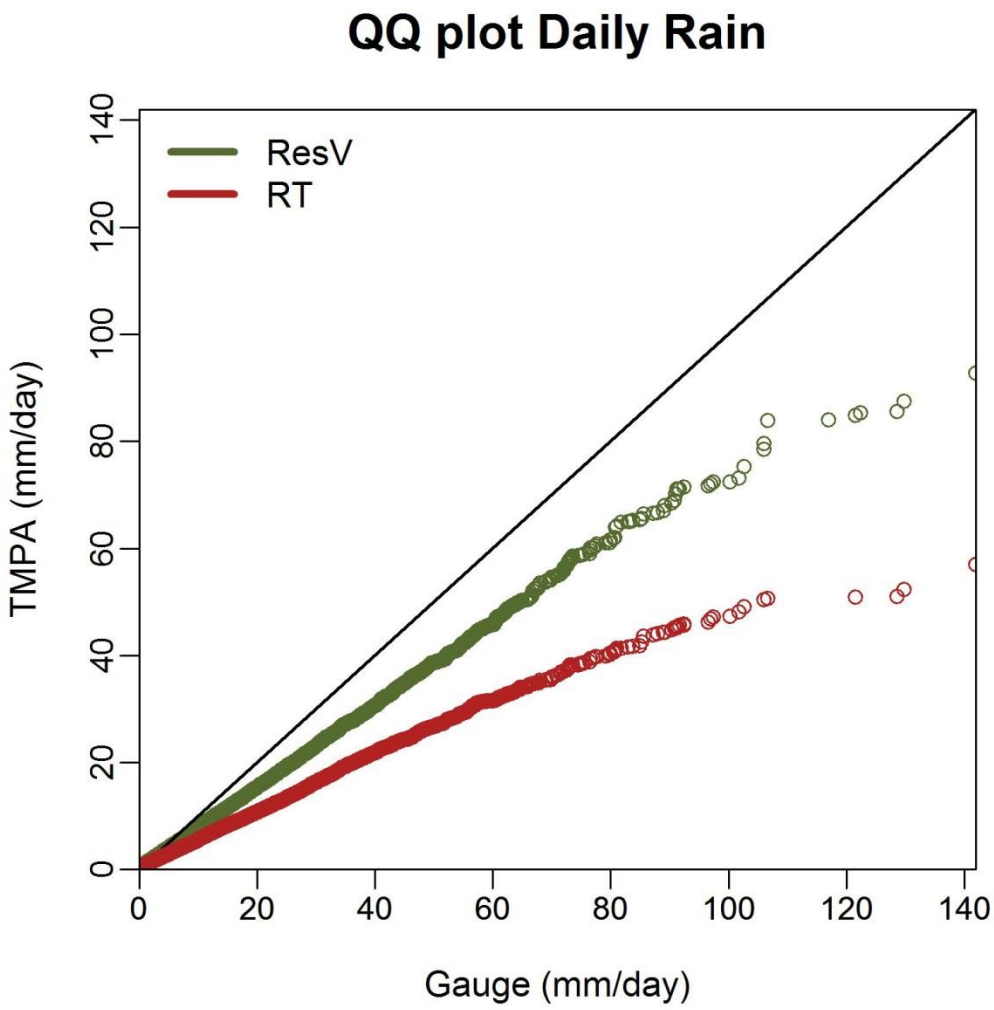


FIG. 4. Temporal characterization of rain gauge data from 1/1/1998 to 1/1/2018. Gauge codes refer to the numbers on the map in Fig. 3 with an additional land code: BU = Burundi, DRC = DR Congo, RW = Rwanda, UG = Uganda. Gauges installed and maintained by the authors are light

blue, gauge data collected from other sources are dark blue. Periods that contain suspicious data are shaded light orange, data gaps are shaded grey. The coloured 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.



1094

1095 FIG. 5. Quantile-quantile plot for daily rainfall measured by the gauges and the equivalent TMPA
1096 data based on the availability of rain gauge data in that pixel for the entire study area. The full
1097 black line presents the 1:1 line. TMPA distributions significantly differ from that of gauge data, of
1098 which the Near Real-Time TMPA product (RT) deviates more than the TMPA Research Version
1099 product (ResV).

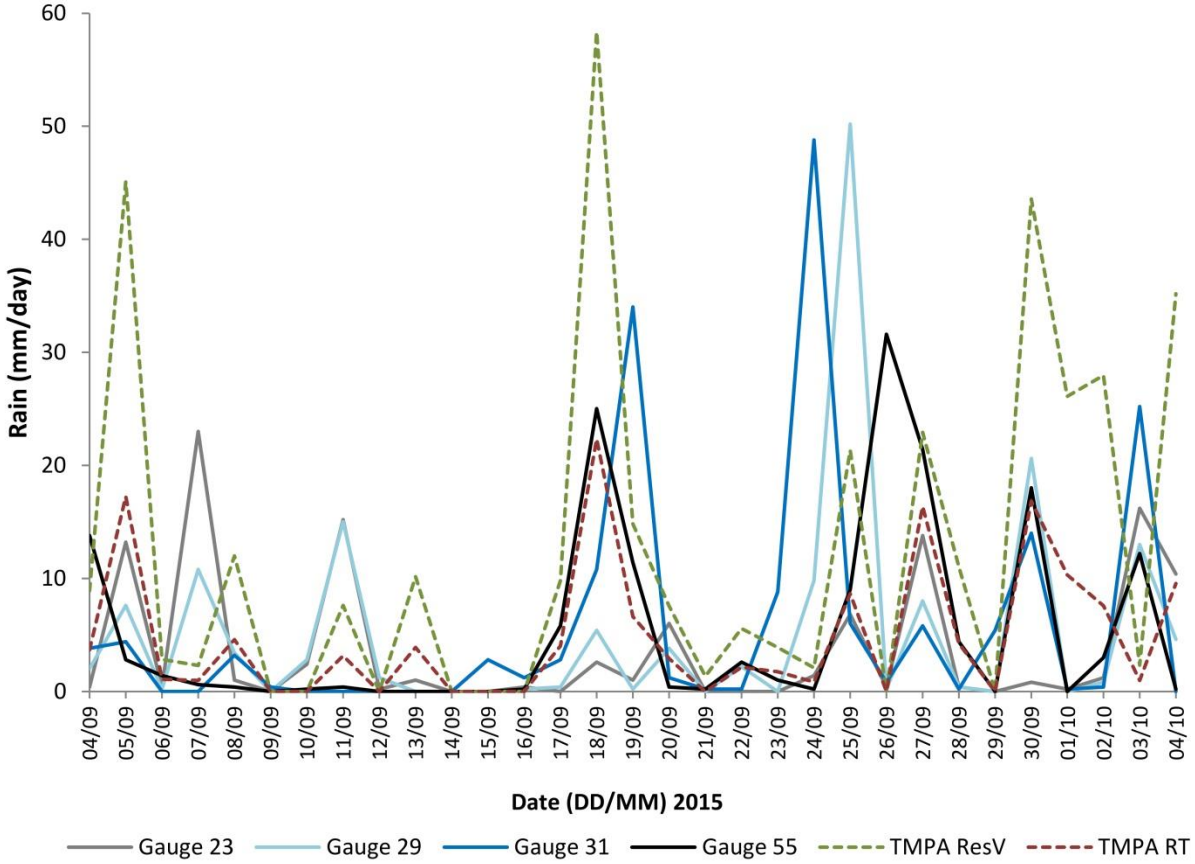


FIG. 6. Daily rainfall time series from 4 September to 4 October 2015 for four gauges located in one TMPA pixel that is located over the Rwenzori Mountains and, together with the TMPA estimates of this pixel (Fig. 3). This presents the variability in observed rainfall in a 25 km x 25 km area.

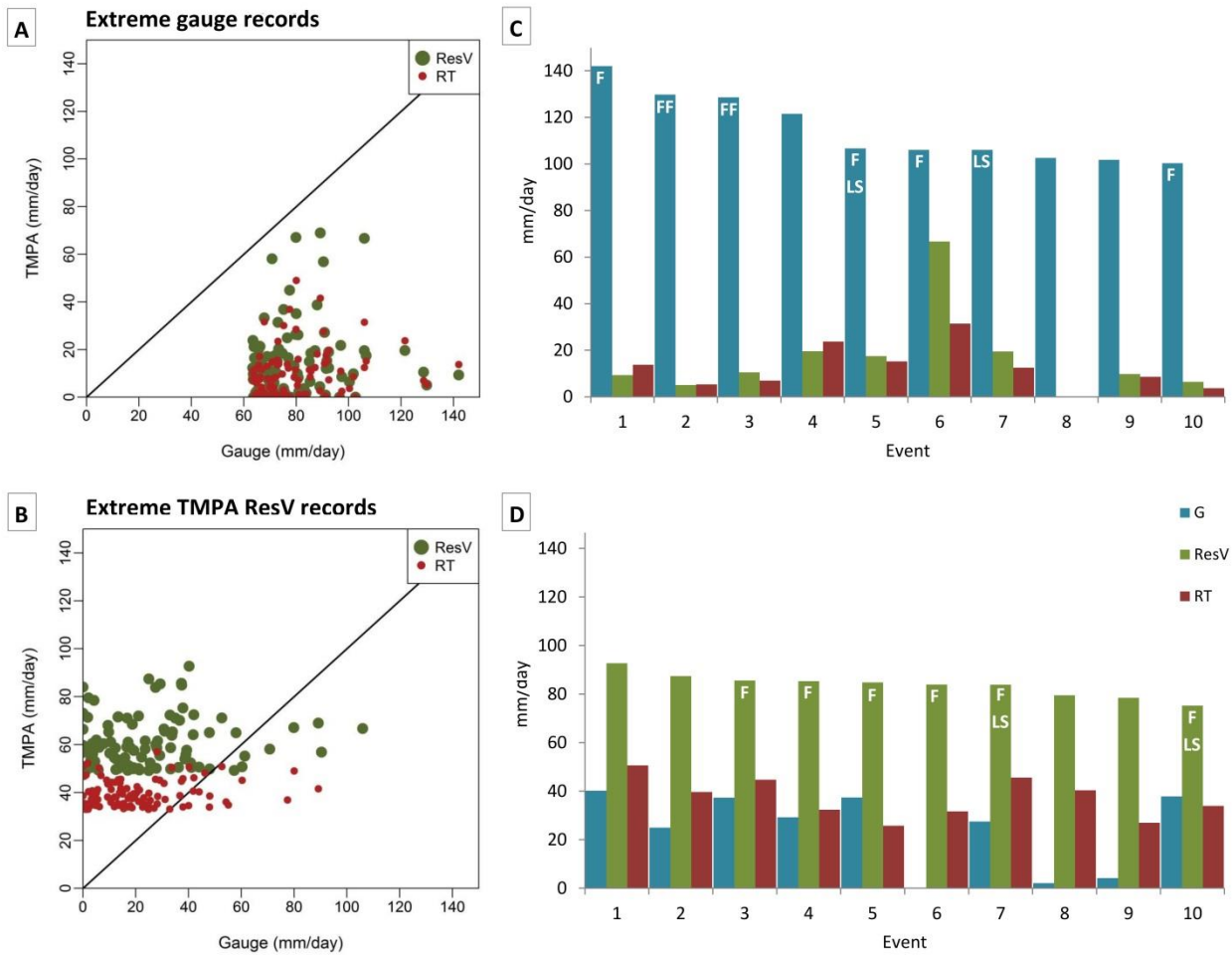


FIG. 7. Extreme daily rainfall recorded by the gauges and by TMPA between 2000 and 2018 based on the availability of rain gauge data in the corresponding TMPA pixel. A: Scatterplot of the top 100 daily rainfall events as measured by the gauges, indicating that each of these events is underestimated by 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 ten most extreme rainfall events (1 = most extreme) measured by (C) the gauges and (D) TMPA ResV were matched with

reported hydrometeorological hazards, if available: F: Flood, FF: Flash Flood, LS: Landslide. The corresponding rainfall measured by the gauge, TMPA ResV and RT in the same pixel are also shown.

1154 FIG. A2. Top ten most extreme rainfall events (1 = most extreme) measured between 2000 and
1155 2018 by TMPA Near Real-Time (RT) were matched with reported hydrometeorological hazards,
1156 if available: F: Flood, LS: Landslide. The corresponding rainfall measured by the gauge in the
1157 same pixel (G) as well as the TMPA Research Version (ResV) are also shown.

1158