

Comparative evaluation of different satellite rainfall estimation products and bias correction in the Upper Blue Nile (UBN) basin

Wuletawu Abera^a, Luca Brocca^b, Riccardo Rigon^a

^a Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy

^b Research Institute for Geo-Hydrological Protection, National Research Council, Perugia, Italy

Abstract

In a region where ground-based gauge data are scarce, satellite rainfall estimates (SREs) are a viable option for proper space-time rainfall characterization. However, their accuracy and performances vary from region to region, and must be assessed. In this study, five high resolution satellite products (3B42V7, CMORPH, TAMSAT, SM2R-CCI, and CFSR) are compared and analysed using the available rain gauge data in one of the most topographically and climatologically complex regions, the Upper Blue Nile basin. The basin rainfall is investigated systematically, and it is found that, at some locations, the difference in mean annual rainfall estimates between these SREs could be as much as about 2700 mm. Considering three goodness-of-fit indexes, correlation, bias and root mean square error (RMSE) between the SREs and ground-based gauge rainfall, CMORPH, TAMSAT and SM2R-CCI outperform the other two. Furthermore, a confusion matrix is used to investigate the detection ability of satellite rainfall products for different rainfall intensities. TAMSAT has the highest (91%) detection skill for dry days, followed by CFSR (77%). On the contrary, SM2R-CCI has the highest accuracy index for medium rainfall ranges (10-20 mm). The empirical cumulative distribution (*ecdf*) mapping technique is used to correct the SREs intensities distribution. This method provides a means to improve the rainfall estimation of all SREs, and the highest improvement is obtained for CMORPH (bias reduction from -72% to -1%).

Keywords: Remote sensing of rainfall, TRMM 3B42V7, CFSR, CMORPH, TAMSAT, SM2R-CCI, Upper Blue Nile

1. Introduction

Rainfall is the primary component of the water balance and is, therefore, the key input for hydrological modelling and water resources analysis. Before any analysis for water resource modelling and/or management can be carried out, it is essential to answer how much water is flowing to the basin (Shaw et al., 2010). Traditionally, rainfall estimation for hydrological modelling and water resource analysis is available from ground-based rain gauges. However, there are many regions in the world where ground-based gauge data are scarce, if available at all, and estimation of spatial rainfall field from ground-based gauge data is questionable. Recently, satellite rainfall estimates (SREs) provide viable alternative options (Ward et al., 2011; Tian et al., 2007; del Jesus et al., 2015; Gebremichael and Hossain, 2010). The key advantage of SREs is that they provide spatially and temporally continuous data that can be used to capture the variability of rainfall. However, the retrieval of rain rate from satellite data is not straightforward and, hence, over the past decades, numerous satellite rainfall estimation algorithms, applied to different satellite sensors, have been developed. For instance, we can mention the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA, (Huffman et al., 2007; Kummerow et al., 1998)), the Climate Prediction Center (CPC) morphing technique (CMORPH, Joyce et al. (2004)), the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN, Sorooshian et al. (2000)) and

recently products based on the novel SM2RAIN algorithm (Brocca et al., 2014). All these products vary in terms of resolution (spatial and temporal) and accuracy, hence, their performances and drawbacks for operational application in water resources modelling must be assessed.

The Blue Nile basin, particularly the Upper Blue Nile basin (hereinafter UBN), is one of the most physically, socially and politically complex basins in Africa, but poorly covered by rain gauges. The Basin is the part of trans-boundary river basin, where different countries have different policies and legal regimes, sometimes with contrasting interest. This represents a challenge on the basin water resource development. Due to the lack of in situ data, most hydrological studies in the region are limited to small basins where there are relatively better hydrometeorological data (Rientjes et al., 2011; Uhlenbrook et al., 2010; Tekleab et al., 2011; Wale et al., 2009; Kebede et al., 2006; Bewket and Sterk, 2005; Steenhuis et al., 2009; Conway, 1997; Mishra et al., 2004; Mishra and Hata, 2006). These studies usually pay little attention to a proper spatio-temporal rainfall characterisation (Kim et al., 2008).

Several validation studies of SREs have been conducted in the Ethiopian UBN basin (Dinku et al., 2007, 2008; Haile et al., 2013; Gebremichael et al., 2014; Worqlul et al., 2014; Romilly and Gebremichael, 2011; Hirpa et al., 2010; Habib et al., 2012). For instance, the two comparative studies by Dinku et al. (2007, 2008) on high temporal (less than and equal to 10 days) and spatial (less than or equal to 1⁰) resolution products show that

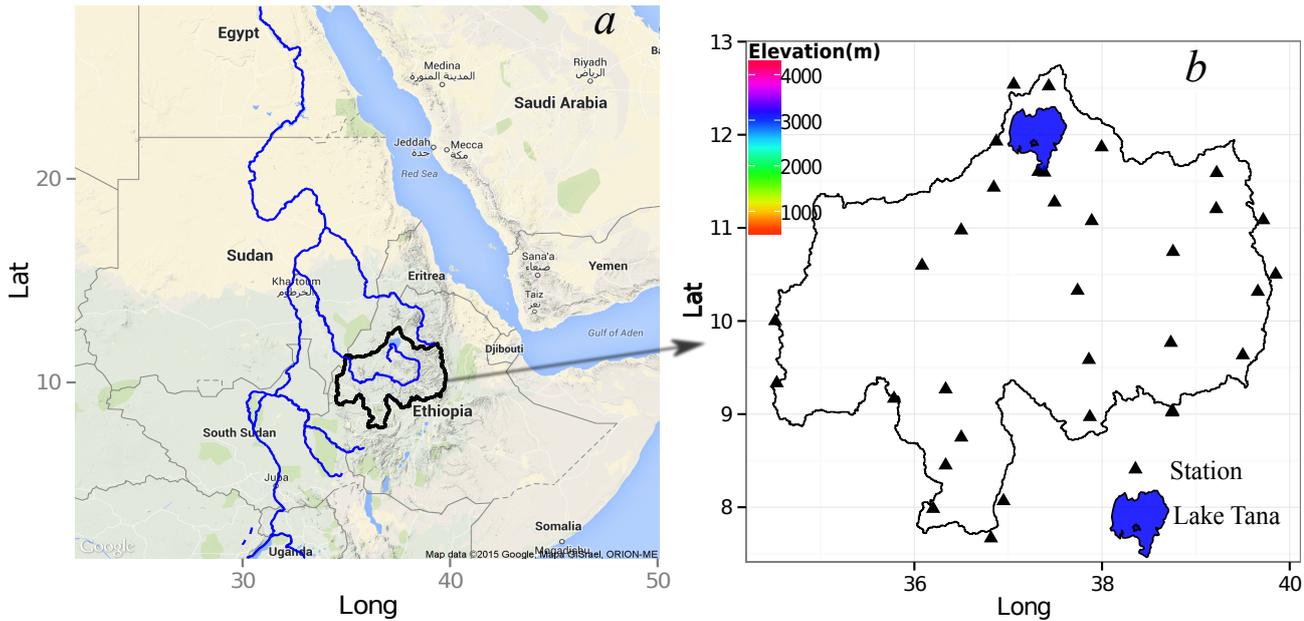


Figure 1: The geographic location of Upper Blue Nile basin in the Nile basin (a) and digital elevation model of the basin (b). The points in figure b are the meteorological stations used for this study.

CMORPH, TAMSAT (Grimes et al., 1999) and TRMM 3B42 (the gauge-corrected version of TMPA products, Huffman et al. (2007)) are three SREs with such a good accuracy to be potentially useful for hydrological applications in the region. Dinku et al. (2008) reported that CMORPH works better in Ethiopia than other regions of Africa, while Haile et al. (2013), studying the accuracy of CMORPH over a subbasin of UBN basin for three months, found poor accuracy with respect to other regions. More recently, Gebremichael et al. (2014), by designing experimental rain gauges for two summer seasons in two experimental locations (one in the lowlands and one in the highlands) of the UBN basin, examined the accuracy of three high-resolution satellite rainfall products (CMORPH, TRMM 3B42RT - the real-time version of TMPA - and TRMM 3B42). Regarding the relationships between SREs goodness-of-fit values and topography (particularly elevation) of the experimental sites, SREs overestimate the mean rainfall rate in the lowlands and, vice versa, underestimate at the highland site. On 3-hourly time scales, the SREs miss more rainfall at the highland sites than at the lowland sites. Worqlul et al. (2014) extended these studies by comparing TRMM-3B42 with two other products (Multi-Sensor Precipitation Estimate Geostationary (MPEG) and the Climate Forecast System Reanalysis (CFSR)) in the Lake Tana subbasin in 2010 and found that MPEG and CFSR are far better than TRMM-3B42. CFSR is one of the highest resolution (time and space), multiyear, global gridded reanalysis rainfall dataset, and is used as input to the rainfall-runoff modelling in Gumera basin (subbasin of the UBN) and provides similar performances to the ground-gauge data model inputs (Fuka et al., 2014). This study compliments the above mentioned researches (Dinku et al., 2007, 2008; Haile et al., 2013;

Gebremichael et al., 2014; Worqlul et al., 2014), which have been conducted either for some subbasin or only for short time series in some designated experimental sites.

The objectives of this paper are twofold: 1. to comparatively evaluate five daily SREs, i.e. four commonly used products (TRMM 3B42, CMORPH, TAMSAT, and CFSR) and the new SM2R-CCI (see Data sets section) product; and 2. to assess the potential of improving the SREs by using a bias correction method. In the view of evaluating the SREs for hydrological modelling, the study aims to unify all the above results by analyzing relatively high resolution (daily) SREs against available daily gauged rainfall data for a 10 year period (2003-2012). It is known that some products such as CMORPH and IMERG are available at higher resolution (at sub daily time steps i.e. 30 min and 3 hour). However, in this study, we have focused at daily time steps due to two reasons. Firstly, at large basin scale as UBN, the hydrological simulation at daily time step can be considered high resolution. Secondly, there are no in-situ measured data to evaluate the SREs estimation skills at sub daily time steps.

The paper is organized as follows: section 2 provides a basic introduction to the study basin, followed by data descriptions (section 3) of both SREs (subsection 3.1) and ground-based data (subsection 3.2). Section 4 elaborates the methods used for comparison and for bias correction of SRE. The comparison of SREs in relation to the ground-based data and the bias correction results, and the conclusions of the study are presented in section 5 and 6 respectively.

2. Study area

The study basin is the Upper Blue Nile (UBN) basin. The UBN basin contributes to 60% Of the total contribution (85%) of the Ethiopian highlands to the Nile river flows (Abu-Zeid and Biswas, 1996; Conway, 2000). The total area of UBN enclosed at the Ethio-Sudan border is about 175,315 km² (figure 1). The UBN originates at Lake Tana, and flows to Sudan, where it meets the White Nile River at Khartoum. The large scale hydrological behavior of the basin is described in a series of studies (Conway, 1997, 2000, 2005; Conway and Hulme, 1993). Its hydrological behavior is characterized by high spatio-temporal variability. Since UBN basin has the lion's share of the total Nile flow, it is the economic mainstay of downstream countries (Sudan and Egypt). Moreover, the Ethiopian highlands are highly populated and have high water demands for irrigation and domestic uses.

The topographic distribution of the basin is shown in figure 1 b. The topography of UBN is very complex, with elevation ranging from 500 m in the lowlands at the Sudan border to 4160 m in the upper parts of the basin. Due to the topographic variations, the climate of the basin varies from cool (in the highlands) to hot (in the lowlands), with large variations in a limited elevation range. The wet season, with low temperatures, is from June to September, while the hot season is from March to May. Three controlling mechanisms of the rainfall characteristics of the UBN, and Ethiopia in general (Seleshi and Zanke, 2004), are the Intertropical Convergence Zone (ITCZ) that mainly derive the wet monsoon rainfall in the wet season (June to September), the Saharan anticyclone that generates the dry and cool northeasterly winds in the dry season (from October-February), and the Arabian highlands that produce thermal lows in the mild season (February - May). The mean annual rainfall and potential evapotranspiration of the UBN basin are estimated to be in the ranges of 1200-1600 mm and 1000-1800 mm respectively (Conway, 1997, 2000), with high spatio-temporal variability. The annual temperature mean is 18.5° with small seasonal variability.

3. Data sets

3.1. Satellite Rainfall Estimate (SRE) products

In this section, we describe five high resolution SRE products, TRMM 3B42, CMORPH, TAMSAT, CFSR, and SM2R-CCI, that will be used for rainfall estimates in the following sections. All the products are available at spatial resolution of 0.25° (except TAMSAT which is 0.035°) and temporal resolution of subdaily to daily. To obtain the rainfall estimation, one of the product (i.e. SM2R-CCI) has a different procedure than the rainfall retrieval algorithm from PMW/IR observations. Nevertheless, here, we are referring SRE (satellite rainfall estimate) for all the five products for simplicity. The detailed description of each algorithm/product can be referred to in the literature cited in table 1.

TRMM is the joint NASA and JAXA mission originally aimed at studying tropical rainfall (Kummerow et al., 2000;

Rozante et al., 2010). TRMM 3B42 version 7 (hereinafter defined as 3B42V7) (Huffman and Bolvin, 2013; Huffman et al., 2007) is among the TMPA products with high spatial resolution of (0.25°) and high temporal resolution (3-hour). The spatial coverage extends from 50 degrees south to 50 degrees north latitude. For this study, 3B42V7 is obtained from NASA's TRMM Online Visualization and Analysis System (Liu et al., 2007, 2012) (<http://disc.sci.gsfc.nasa.gov/precipitation/tovas/>). The detailed information on processing and generation of the 3B42V7 can be found in the literature cited in table 1. We underline here that the 3B42V7 product is corrected with rain gauge observations (i.e., it is not based only on satellite data). Indeed, the 3-hourly rainfall fields are corrected on a monthly basis with the Global Precipitation Climatology Centre (GPCC) monthly rain gauge dataset by using inverse-error-variance weighting methods (Huffman and Bolvin, 2013; Huffman et al., 2007).

The CMORPH product (Joyce et al., 2004) is a rainfall estimation method that mainly relies on PMW observation. It uses precipitation estimates that have been derived from low orbit satellite PMW observations exclusively, and whose features are transported via spatial propagation information that is obtained from geostationary satellite IR data (Joyce et al., 2004). CMORPH is a near global product (60° north and south) and provides data at 3 hourly and daily resolution since 1998; and at temporal resolution of 30-min since December 2002. The 3-hourly precipitation at 0.25° spatial resolution data are used for this study.

CFSR is one of the SREs assessed in this study. It is a re-analysis product that combines the weather forecasts generated by the National Weather Service's NCEP Global Forecast System, and satellite data (Saha et al. (2010)). The analysis is reinitialized every 6-hours, and it is obtained at hours 00:00, 06:00, 12:00, 18:00 UTC. The spatial resolution of CFSR is about 38 km, and it is available since 1979.

Tropical Applications of Meteorology Using Satellite and Ground-Based Observations (TAMSAT) (Grimes et al., 1999; Tarnavsky et al., 2014), is a rainfall product specialized for Africa. It has been providing 10-day rainfall estimates since 1983 and recently (since 2013) daily products are available. The TAMSAT rainfall estimation algorithm is calibrated using the historical rain gauges, TAMSAT African Rainfall Climatology And Time-series. The methodology and algorithms of the TAMSAT rainfall product derive from Meteosat imagery and gauge data using contemporaneous cold cloud duration fields (Tarnavsky et al., 2014; Maidment et al., 2014). The product is available at a resolution of 0.0375° at nadir (~4 km). Some studies (Laurent et al., 1998; Thorne et al., 2001; Chadwick et al., 2010; Maidment et al., 2013; Jobard et al., 2011) find that TAMSAT estimates (over ten days) are comparable to the other SREs, even better in some cases, when compared with gauge data. Dinku et al. (2007) showed that TAMSAT (over ten days) rainfall estimates perform as well as CMORPH and TRMM (version 6) in Ethiopias central highlands. However, the new TAMSAT daily rainfall estimates performances have not yet been assessed.

The SM2R-CCI rainfall product is based on the SM2RAIN

Table 1: Summary of the different satellite products used in this study. Because CMORPH is available since December 2002, the study is based on only 10 years of data. However, the statistical indexes analyzed for all the years with available data is almost the same and the difference is not statistically significant.

SREs	Spatial res. ^o	Temporal res.	Available	Data used	References
3B42V7	0.25	3-hourly	1998-present	2003-2012	Zhao and Weng (2002); Huffman et al. (2007); Prakash et al. (2015); Kumar et al. (2014); Duan and Bastiaanssen (2013); Romilly and Gebremichael (2010)
CMORPH	0.25	3-hourly	1998-present	2003-2012	Ferraro (1997); Joyce et al. (2004); Haile et al. (2015); Romilly and Gebremichael (2010); Gao and Liu (2013)
TAMSAT	0.0375	daily	1983-present	2003-2012	Grimes et al. (1999); Maidment et al. (2014); Tarnavsky et al. (2014)
CFSR	0.3125	6-hourly	1979-present	2003-2012	Saha et al. (2010, 2006); Wang et al. (2011); Xue et al. (2011); Saha et al. (2014)
SM2R-CCI	0.25	daily	1990-present	2003-2012	Brocca et al. (2013, 2014); Ciabatta et al. (2015)

method (Brocca et al., 2013) and is obtained from the CCI soil moisture product (Dorigo et al., 2015) for the period 1990-2013 at 0.25 spatial resolution. The SM2RAIN algorithm is based on the inversion of the water balance equation by considering the soil as a natural rain gauge (Brocca et al. (2014)). The SM2RAIN-derived product provides an integrated estimate of rainfall (daily in this case). In this study, SM2RAIN is applied for the first time to the CCI soil moisture product, allowing to obtain a long-term (24-year) daily rainfall product, by using the same approach given in Ciabatta et al. (2015). Specifically, the SM2RAIN parameter values are calibrated by using the 3B42V7 product as benchmark in the period 1999-2005. In the evaluation period, there is a three years (2003-2005) of overlap with the calibration period (1999-2005) between 3B42V7 and SM2R-CCI. This could (slightly) affect the independency between the two products. However, the use of data sets of different length for SRE evaluation statistics (see section 4.1), which also have longer period of independent data, helps to understand if that this dependency does not affects the results considerably. We note that this calibration is performed here for the UBN study area but, potentially, it can be easily applied on a global scale.

3.2. Rain Gauge rainfall data sets

Rain gauges across the UBN basin are very scarce. In this study, 35 daily rain gauges are used. The in situ data is obtained from the national meteorological agency of Ethiopia. The spatial location and elevation of the stations used for this study are shown in figure 1b. The data are manually checked for quality and the final refined data are used in the following analysis. Looking at the distribution of the stations, the upper part of the basin has relatively higher station density than the lowland and the middle of the basin (figure 1b).

4. Methodology

The daily rainfall data estimates from 3B42V7, CMORPH, TAMSAT, CFSR, and SM2R-CCI at grid level are compared

with daily rainfall data from raingauges. Since the density of the rain gauges is very low and the topography of the study area is complex, we decided not to interpolate the rainfall data over the spatial fields but to compare directly point rainfall at the stations, with the the grid rainfall as in (Porcù et al., 2014; Worqlul et al., 2014).

4.1. Performance metrics

To make the comparison we used the products at daily time steps. Therefore, all the SRE products at sub daily time steps are accumulated into daily totals. Since we are interested in assessing the entire performances of the SREs in the basin, the statistics are calculated for the whole data set (including zeroes). We used the Pearson linear correlation coefficient (r), the Root Mean Square Error (RMSE), and the BIAS goodness-of-fit (GOF) indexes, as defined in Appendix B. An important question concerning SREs is their error sensitivity and variability to the length of data used for evaluation. In order to assess the effects of data series length on the performance statistics (GOF statistics), we used different sets of data series length for estimation of GOFs between SREs and gauge observed data. We have evaluated the SREs error statistics progressively for 1, 4, 7 and 10 year period. The 1 year means the analysis is based on only the first year of the data set (2003), 4 years from 2003 to 2006, 7 years from 2003 to 2009 and 10 years the whole period. Accordingly, we calculate the GOF statistics for the four data sets with different length.

To further investigate the errors and prediction capacity of SREs, we decomposed the errors in classes as in (Salio et al., 2015; Cohen Liechti et al., 2012; Haile et al., 2013). They use two class of data, i.e., rain and no-rain distinction, and the products are evaluated using the accuracy index, given by:

$$Accuracy = \frac{N_a + N_d}{N_a + N_b + N_c + N_d} \quad (1)$$

Where the terms of the equation are described by the confusion matrix given in table 2.

In this study, we further analyze the detection skill of the SREs by subdividing the SREs into several rainfall classes based

Table 2: Confusion matrix based on the four possibilities of SRE detection of the gauge observed rainfall. The four possibilities are true positive (a), false positive (b), false negative (c), and true negative(d).

	SREs:Yes	SREs:NO
Gauge: Yes	a	b
Gauge: NO	c	d

on observed rainfall amounts. Specifically, the SREs are divided into 7 rainfall classes (all in mm/day): ≤ 0.1 , 0.1-2; 2-5; 5-10; 10-20; 20-40; ≥ 40 . The first class (class 0) characterizes dry days. In literature, separation of dry days is obtained by imposing a minimum detectable rainfall threshold. This threshold value is controversial, as it ranges from 0.1 to 1 mm/day (Jakob Themeßl et al., 2011; Moon et al., 1994; Lázaro et al., 2001; Schmidli and Frei, 2005; Kisaka et al., 2015). In this study, we use 0.1 mm as suggested by World Meteorological Organization (Jarraud, 2008).

The confusion (or matching) matrix of different SRE amounts versus rain gauge measured amounts is estimated for each rainfall class. As it can be seen in the results section, multi-class confusion matrix tells the proportion of SREs in different classes, shown on the y-axis, for a given observed rainfall classes, in the x-axis. In the ideal situation, where the SREs are perfectly consistent with the observed data, the matrix would be an anti-diagonal matrix where all the entries are zero except on the diagonal going from the lower left corner to the upper right corner with ones.

4.2. Bias correction

The comparative evaluation study and selection of the best product may not always give an accurate enough product for hydrological modelling. Therefore, further bias correction of SREs could be important to obtain reasonable results (AghaKouchak et al., 2012). In the second section of the analysis, systematic bias correction of SREs is performed. The empirical cumulative distribution function (*ecdf*) mapping bias correction technique, initially proposed by (Panofsky et al., 1958), and recently applied for hydrological data correction by (Themeßl et al., 2012; Michelangeli et al., 2009; Iizumi et al., 2011; Maurer and Pierce, 2014; Hwang and Graham, 2013), is used. In this method, the *ecdf* of the SREs is first matched to the *ecdf* of the gauges observed rainfall estimates, generating a correction function depending on the data percentile. The correction function is derived for each percentile. If a new extreme value is available in the new SREs, the correction function is linearly extrapolated. Then, the correction function is applied to the *ecdf* of the SREs ($ecdf_i^{S,cal}$) and the *ecdf* of the ground-gauge data at calibration points ($ecdf_i^{G,cal^{-1}}$) to transfer the original SREs ($S_{t,i}^{val}$) to the corrected SREs ($S_{t,i}^{Corr,val}$) at other validation points, as given by (Michelangeli et al., 2009; Themeßl et al., 2012)

$$S_{t,i}^{Corr,val} = ecdf_i^{G,cal^{-1}}(ecdf_i^{S,cal}(S_{t,i}^{val})) \quad (2)$$

where $S_{t,i}^{Corr,val}$ is the new corrected SRE for a given stations i and t time steps. The $ecdf^{-1}$ is the inverse of the *ecdf*. S and G are SRE and gauge rainfall estimate, respectively. The subscript t, i are the time steps and station number, respectively.

As shown in different studies (Cai et al., 2015; Hossain and Huffman, 2008; Ebert et al., 2007), biases and errors in SREs exhibit space and time variability. To reduce the effect of spatial variability in bias correction procedure, the 35 stations are systematic divided into two groups: 17 stations for calibration and 18 stations for validation as shown in (figure A.8). The correction function (equation 2) is constructed based on all stations combined *ecdf* matching between the observed rainfall and SREs at calibration stations. The function generated is used to develop new corrected SREs that can be evaluated in the appropriate meteorological stations. While this approach may not provide the highest performance at a particular station, it is useful to obtain a good performance on average, at the basin. Similarly, to reduce the effect of temporal SREs errors, the bias correction procedures is applied for each season independently (as in Bennett et al., 2014). For the sake of simplicity, as the basin receives 75% of the rain in summer (Mellander et al., 2013), we apply the procedure only in the summer season (June, July, and August).

For validation of time series bias correction procedure, the data set (10 years of data for all stations) is split in to two; the first 5 years (2003-2007) to develop the transfer function, and the next 5 years (2008-2012) for subsequent evaluation of the bias correction procedures using the same approach of spatial cross validation. The procedures of transfer function generation and validation are applied for all stations combined.

5. Results and Discussions

The study results are presented in two subsections. The first part comparatively evaluates the five daily SREs against the gauge observed rainfall data. The second subsection focuses on the results of SREs bias correction using *ecdf* matching method.

5.1. Comparative evaluation of SREs

The five SREs are analyzed based on GOF statistics (r , RMSE and BIAS) and on their detection capability (confusion matrix and accuracy index). In figure 2, we summarized the GOF statistics between the SREs and gauge observations by splitting the 10 years of data into subsets of 1, 4, 7 and 10 years.

Figure 2 illustrates that GOF statistics generally show stability with increasing length of data series used for the evaluation of SREs. For instance, for 3B42V7, the median value of r slightly decreases (from 0.5 to 0.47) with increasing the length of data period from 1 to 10 years. Similar patterns are observed for all the other SREs, except SM2R-CCI which shows an overall improvement when the whole period is considered. The latter result is expect due to the higher accuracy and temporal resolution of the CCI soil moisture product (Dorigo et al., 2015) on which the SM2R-CCI product is built on. The correlation coefficient comparison between the SREs shows that CMORPH,

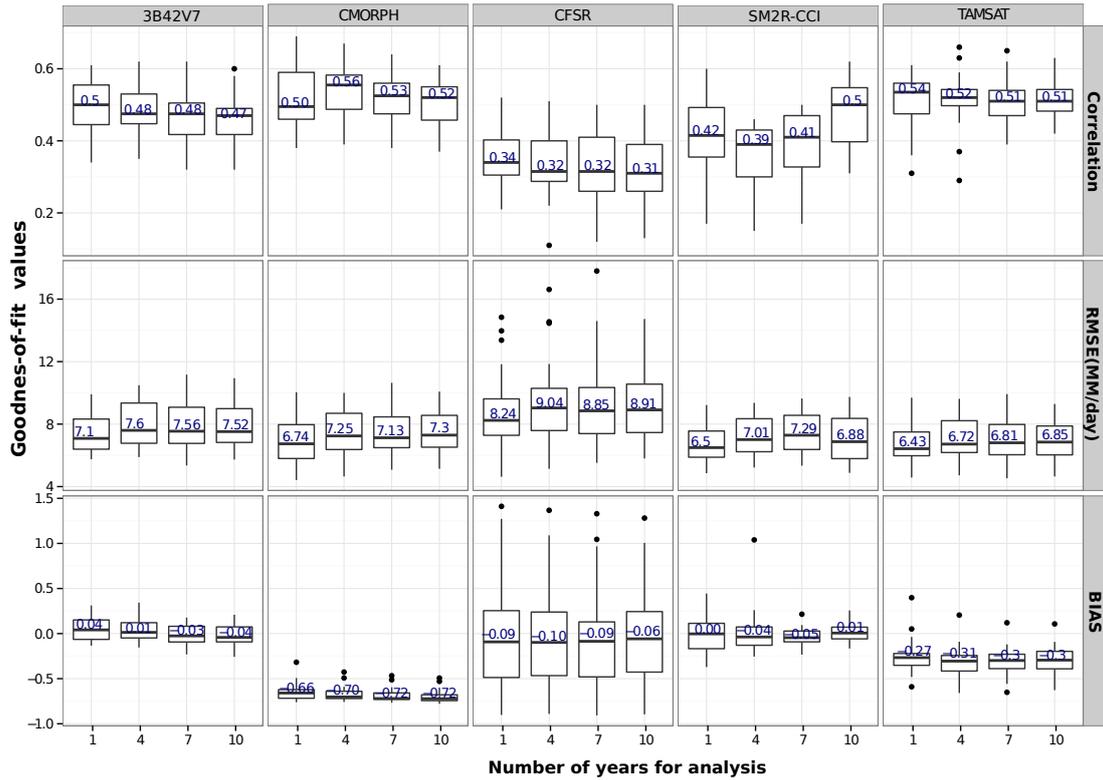


Figure 2: Comparison of SREs with the ground-based rainfall estimation using (above) correlation (-), (middle) Root mean square error (mm/day) and (below) BIAS (-). In the boxplot, the horizontal line in the middle shows the median, the bottom and top end of the box shows the 25th and 75th percentile, respectively, the whiskers (vertical line) shows the range of the data, of the GOF values. For CFSR and TAMSAT, some stations (with high annual gauge observed rainfall) shows outliers GOF values (dot in in the figure).

TAMSAT and SM2R-CCI perform better than the other two SREs. The same figure demonstrates that the correlation of CFSR is low and characterized by high variability as shown by the long boxplot. The median r value of 3B42V7 (0.47) is generally higher than the values reported by Worqlul et al. (2014), which however has a maximum value of 0.29 (monthly scale analysis).

The data analysis for 1, 4, 7, and 10 year data sets illustrates that also the effect of length of the data set on RMSE values is rather limited (figure 2, second row). For most SREs, RMSE median values for 1 year validation is smaller (likely due to the selection of 2003 that is a dry year), and it increases to 4 year, and then it keeps stable for 7 and 10 years. The comparison between the SREs for the 10 years evaluation results is consistent with the correlation results, i.e., the RMSE value of TAMSAT (6.85 mm/day), SM2R-CCI (6.88 mm/day), and CMORPH (7.3 mm/day) shows better performances in comparison to the other SREs.

The daily RMSE values of CFSR and 3B42V7 by Worqlul et al. (2014) at a specific subbasin (Lake Tana watershed) is 6.2 and 4.0 mm/day, respectively. These values are smaller than our results (CFSR=8.9 mm/day, 3B42V7 = 7.52 mm/day). The difference (as discussed below) could be due to the stations in the lowland area of the basin (the western part) which shows higher RMSE values than the eastern and north-eastern highlands where Lake Tana basin is located. Hence, it can be interpreted that RMSE of CFSR is smaller than the 3B42V7 in

the highlands (to be specific Lake Tana basin), while the results in this study shows that at the whole UBN basin, 3B42V7 shows better performance (RMSE = 7.52 mm/day) than CFSR (RMSE=8.91 mm/day). The study of Young et al. (2014), in the Oromia region of Ethiopia, find similar results: TAMSAT and CMORPH have lower RMSE values in comparison to 3B42V7 at daily time steps. Based on the r and RMSE statistics, CMORPH outperforms 3B42V7, which is consistent with the results of Bitew et al. (2012) for the small subbasin of UBN basin, the Koga watershed, but for a single year.

In general, the effect of data length on BIAS is very small, and it is valid for all the SREs (figure 2, third row). For instance, the BIAS for 3B42V7 decreases from 4% for 1 year evaluation to -4% in 10 years, the same level of BIAS but opposite sign. A similar slight decline in BIAS is shown for CMORPH (from -66% to -72%) when the number of years in the analysis increases. The comparison of the five products using BIAS is not consistent with the products comparison using r and RMSE (figure 2, the third row). For instance, SM2R-CCI (0.001) has the lowest BIAS, followed by 3B42V7 (-0.042) and CFSR (-0.06). The low BIAS of SM2R-CCI has to be attributed to the use of 3B42V7 for the calibration of the parameter values of the SM2RAIN algorithm. Note that while CMORPH is better in estimating ground-gauge rainfall using the two previous statistics (i.e., r and RMSE), it is underestimating by 72%, thus being the most biased product of the five SREs. This could be because CMORPH is only based on satellite products, and not

corrected using ground data as 3B42V7. TAMSAT, on average, is underestimating rainfall by 30%.

The spatial distribution of the the three GOF values (r , RMSE, BIAS) are presented in figure 3. Overall the distribution of the statistics can depict a spatial pattern, i.e., the correlations in the eastern and northeastern part of the basin are higher than western and southwestern part. Similar pattern can be inferred from the RMSE and BIAS statistics are both smaller in the eastern part (the highlands), and they are higher in the western lowlands of the basin. The BIAS statistic shows a less pronounced pattern in comparison to r and RMSE. While CMORPH shows high underestimation at all stations, 3B42V7, CFSR, and SM2R-CCI shows the same described pattern. Generally, CFSR has less spatial pattern in the statistics, and has relatively high spatially mixed results (figure 3). As elevation increases from the west towards the east and northeast part of the basin, the pattern can be partly explained by elevation. The general pattern is that SRE performance increases with elevation (data not shown). Gebremichael et al. (2014) did a comparison study between the highland and lowland sites, when using RMSE and BIAS, they found the similar result. I.e., the highland sites show better performances than the lowland sites.

Following the results of GOF analysis, the comparison of SREs using the confusion matrix and the accuracy index is carried out. The confusion matrix in figure 4 summarizes the relative bias of SREs for a given observed rainfall class. To analyses the performances of each rainfall class, the SREs values were grouped based on the observed rainfall classes in the x-axis. For instance, the first column shows the distribution of SREs when the gauge recorded rainfall values ≤ 0.1 : it describes the distribution of SRE records for this specific observed rainfall class.

3B42V7 (figure 4a) has the higher detection capacity for the first class (≤ 0.1 mm/day, 68%), with lower performances for the other rainfall classes ($\leq 26\%$). Except for the first and second rainfall classes (≤ 0.1 , 0.1-2mm), which have high detection capacity (73% and 52%, respectively), the higher rainfall classes (≥ 10 mm) show very low detection by CMORPH (below 5%). It is important to note that apparently CMORPH does not detect observed rainfall ≥ 20 mm at all. In general, compared to 3B42V7, CMORPH has a systematic underestimation for all the rainfall ranges (figure 4b). The confusion matrix analysis results of CFSR (figure 4c) show that the pattern in the detection capacity is similar to CMORPH for the first class and to 3B42V7 for the other classes.

Two important results of the SM2R-CCI confusion matrix (figure 4d) are its relatively higher detection capacities for medium rainfall values (from 5mm to 20mm) and relatively lower detection for lower rainfall intensities (38%), likely due to noise in soil moisture observations (see e.g., Ciabatta et al. (2015)). It is only SM2R-CCI that has detection capacity of 25% and above for rainfall ranges between 5 to 20 mm. In general, SM2R-CCI, being based on soil moisture, tends to aggregate the rainfall volume thus being more accurate in the estimation of accumulated rainfall than 3B42V7 and CMORPH that relies on the estimation of instantaneous rainfall rates. This result is also important as it highlights the large potential that could be gained by the

integration of these two kinds of SREs (e.g., SM2R-CCI and CMORPH or 3B42V7 or TAMSAT).

Among the five SREs, TAMSAT has the highest detection capacity for lowest rainfall intensities (91%). For all classes, TAMSAT has the highest missing rate and the highest recorded is for the 0.1-2 mm observed rainfall class (54%), while the systematic bias for all the classes is relatively low (figure 4e). The SREs detection capacity is further evaluated by the overall accuracy capacity, and the comparison is shown in figure 4f. The result confirms the confusion matrix analysis.

The time series rainfall summary analysis is useful for comparative evaluation, but does not provide insight into the aggregate effects of using different SREs on water resource modelling. Figure 5 shows the comparison of long term (2003-2012, 10 years), mean cumulative rainfall for different SREs and measured data. A sample of three stations systematically selected to represent different ranges of elevation and spatial location is used in the analysis. These are Mehal Meda, Debre Markos, and Assosa which are located at high (3084 meters), medium (2446 meters) and low (1600 meters) elevations, respectively. The spatial location of the three stations is shown in the maps plotted in figure 5. Four comments can be drawn:

1. Based on the three stations, the observed long term annual rainfall shows that the effect of elevation is just partially masked by the rainfall climatological regime difference (Mellander et al., 2013). Mehal Meda which is at the highest elevation has a lower mean annual rainfall than Debre Markos.
2. Although CFSR and CMORPH show consistent patterns across elevations, CMORPH always underestimates, whereas CFSR always overestimates rainfall. From this analysis, CMORPH and CFSR have a significant BIAS.
3. The errors on cumulative annual rainfall given TAMSAT, 3B42V7 and SM2R-CCI across elevation are rather contrasting. For instance, 3B42V7 and SM2R-CCI slightly overestimate rainfall in high and low elevation while underestimate it in middle elevation. TAMSAT shows high underestimation in the highland station, and progressively overestimate in the lowland station.
4. The 10 year mean cumulative rainfall value of the five SREs differ tremendously. The total mean annual rainfall difference between the SREs is about 600 mm (in Mehal Meda), 2740 mm (in Debre Markos), and 1600 mm (in Assosa).

5.2. Bias correction using *ecdf* mapping

This subsection assesses and discusses the improvements obtained by using the bias correction function generated at the calibration stations (or calibration time series) and applied to the SREs at validation stations (or validation time series). This technique is evaluated i) by comparing the SRE *ecdf* and the gauge rainfall *ecdf*, and (ii) by computing statistics (e.g., BIAS, confusion matrix, and accuracy index) between ground data and the SREs before and after the correction.

Figure 6 shows the variance of SREs (before correction) *ecdf* distribution from the gauge rainfall *ecdf* distribution in

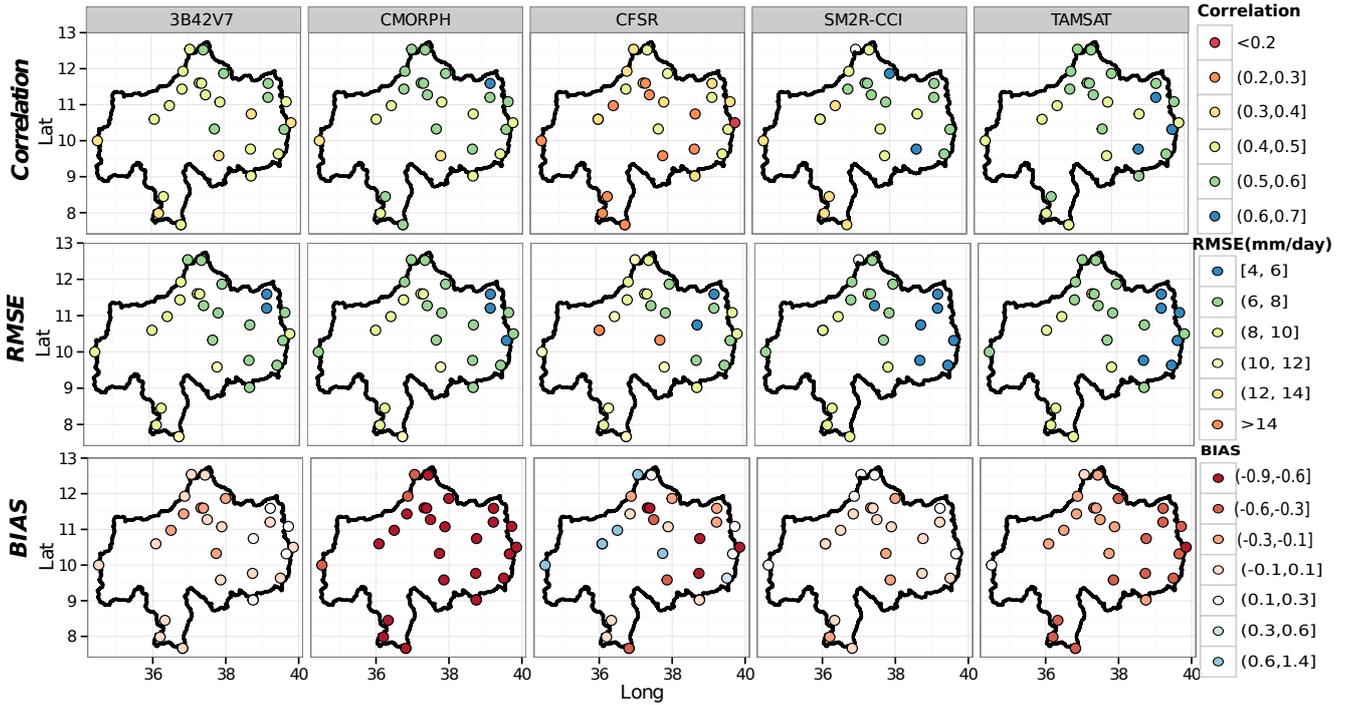


Figure 3: The spatial distribution of GOF values for different SREs: correlation coefficient (first row), RMSE (second row) and Bias (third row).

the case of spatial cross-validation strategy. For all SREs except TAMSAT, it is shown that there is dry-day frequency underestimation. In comparison to the gauge observed rainfall, most SREs tend to overestimate very low (light) rainfall frequency (drizzling-effect Wilcke et al. (2013)), as shown in *ecdf* distribution at figure 6a. The observed drizzling traits for most SREs could be due to its large spatial scale representation in comparison to the point ground-gauge observation. To avoid this kind of effects, the selection of a proper threshold for dry/wet days plays an important role. In figure 6, the zoomed-out plot, gauge observed dry day frequency is used to determine the dry/wet threshold. However, in this approach, a problem arise when SRE dry-day frequency is much greater than the gauge observed rainfall (Themebl et al., 2012; Wilcke et al., 2013), and when dry-day frequency corresponds to high rainfall value for determining SRE threshold. The dry-day frequency matching approach, as shown in figure 6a, is used to define the threshold for CFSR (at 1 mm), 3B42V7 (at 0.5 mm) and CMORPH (at 0.3 mm) in the JJA period. However, for TAMSAT the dry-day frequency is already to high, and threshold is kept to the original 0.1 mm. Contrary to TAMSAT, SM2R-CCI dry-day and low intensity rainfall is much smaller than the gauge observation. For SM2R-CCI the dry-day frequency matching pointed to high threshold (4.5 mm). Since this will cut an important rainfall information, and as the use of this threshold (4.5 mm) reduce the bias correction performances, we also used the original 0.1 mm threshold for SM2R-CCI. Once the threshold information for the SREs is estimated, below these thresholds, the rainfall values are set to be zero, and above the threshold the *ecdf* of SREs are mapped to the *ecdf* of the observed data to construct the correction function. This correction function is

thus used to correct new SREs.

The overall evaluation of SREs improvement from the bias correction application in both independent station (cross-validation) and time series GOF index (mainly bias), confusion matrix, and detection capacity (accuracy) are shown in table 3 and figure 7. The table 3 shows application of the *ecdf* bias correction in a spatial cross-validation and future time series validation improves SREs. The result illustrate that for spatial cross-validation application, except CFSR, most SREs rainfall bias strongly improved. For instance, the highest bias observed was -70% for CMORPH, and the summer season bias correction method improves it to only -4% (table 3), followed by SM2R-CCI from -12% to 1%. The improvement in the detection capacity, however, is very small. This could be due to the use of a single correction function for the summer season as a whole. The construction of monthly correction functions could further improve the results. The results for time series split are very similar. The exceptional in the time series evaluation is 3B42V7 which does not show any improvement in bias.

The overall detection capacity improvement from the correction procedures is also presented in figure 7A and B. Generally, it improves the detection capacities and bias of the SRE products. For instance, for spatial cross-validation, large proportions of CFSR and SM2R-CCI estimates, 15 to 30% and 40 to 60% respectively, are concentrated in the 5-10mm rainfall range. The correction function improves this trend and BIAS for all classes is distributed to all classes of SREs estimations. The derivation of dry/wet rainfall threshold using the *ecdf* matching also improves the detection capacity for the zero (dry) rainfall class, except for TAMSAT, which already has the highest detection capacity for dry days (figure 7). This

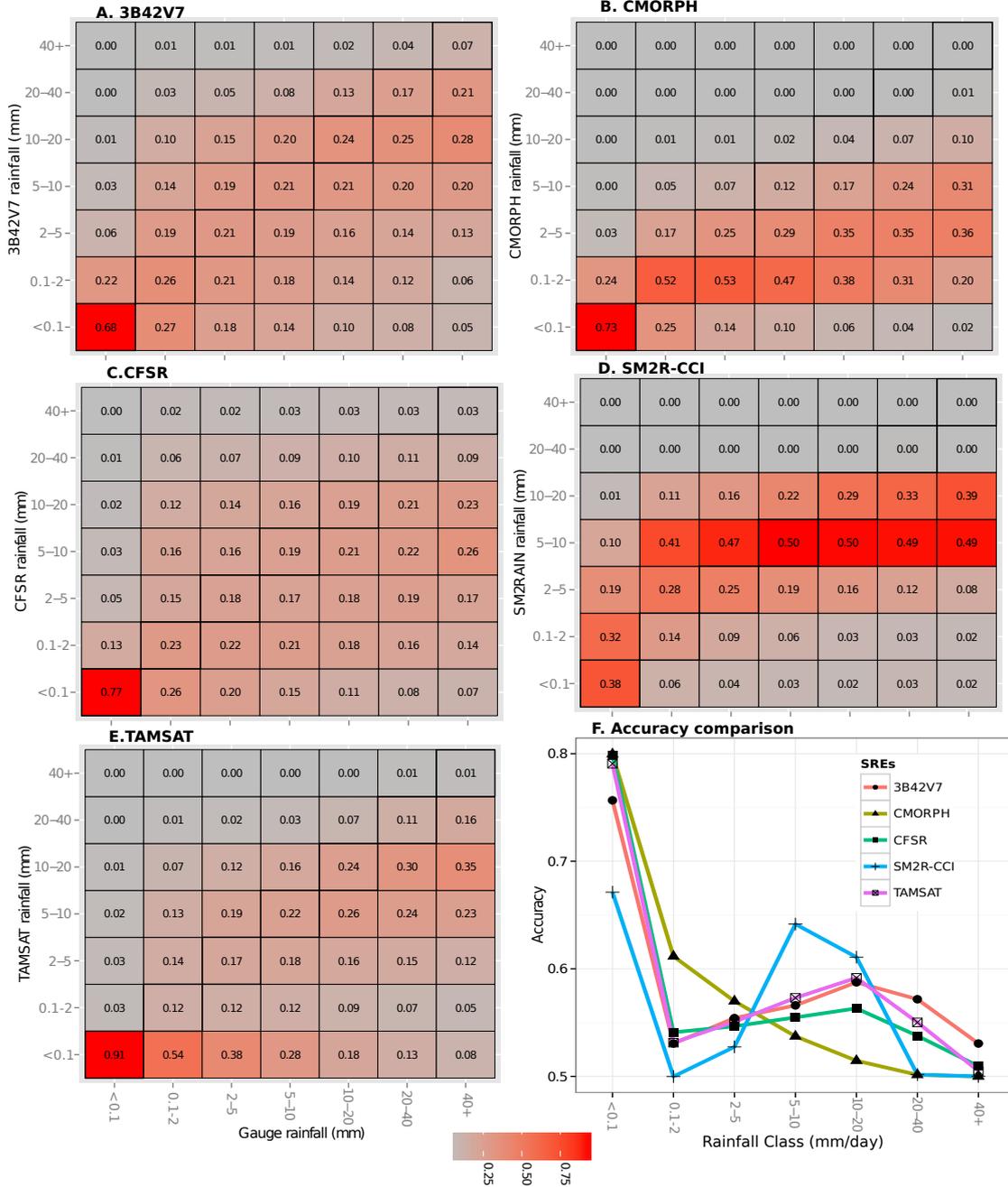


Figure 4: Confusion matrix that shows the proportion of observed rainfall classes in the x-axis reproduced by the different SREs (3B42V7 (a), CMORPH (b), CFSR (c), and SM2R-CCI (d), TAMSAT (e)) in the y-axis. Each value in the confusion matrix is the proportion of the gauge observed rainfall class (along column) that is estimated as a particular SREs class (along row). The comparison of SREs using accuracy index, for different rainfall classes, is shown in figure F. The statistics are based on the daily rain gauge data size of ≤ 0.1 mm (N = 71010), 0.12 mm (N = 11059), 25 mm (N = 8500), 510 mm (N = 8602), 1020 mm (N = 8546), 2040 mm (N=5141) and ≥ 40 mm (N = 1098).

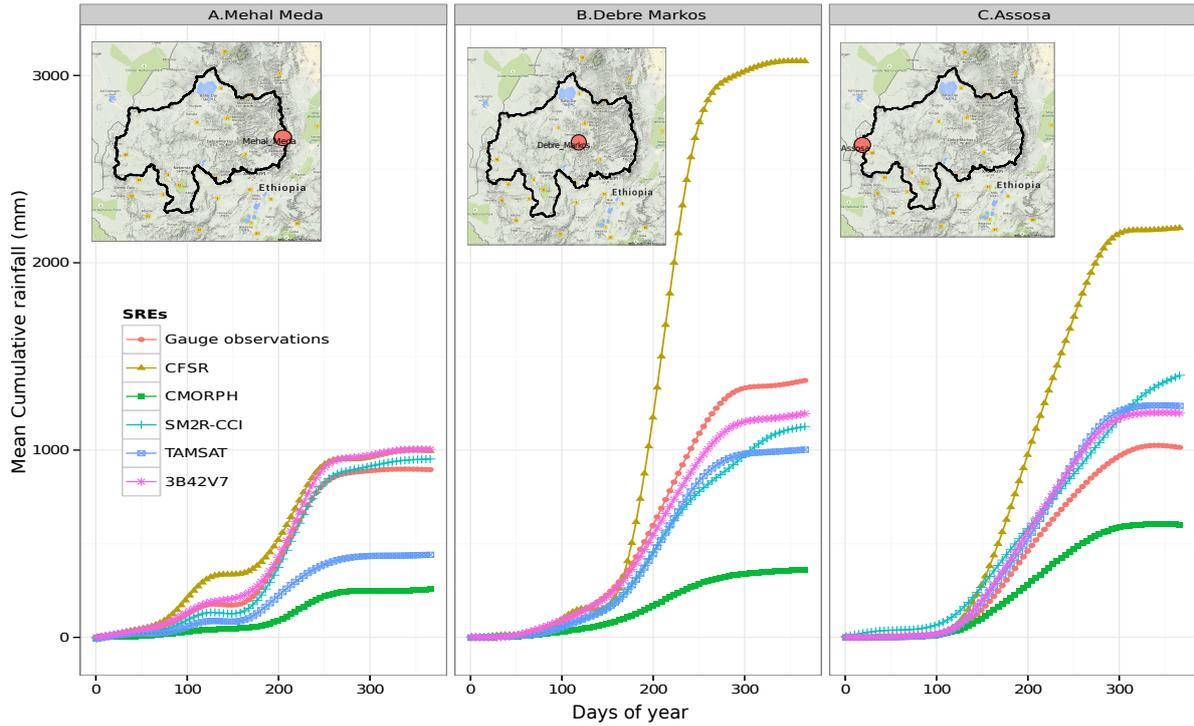


Figure 5: Annual mean cumulative rainfall estimations based on five SREs and gauges data.

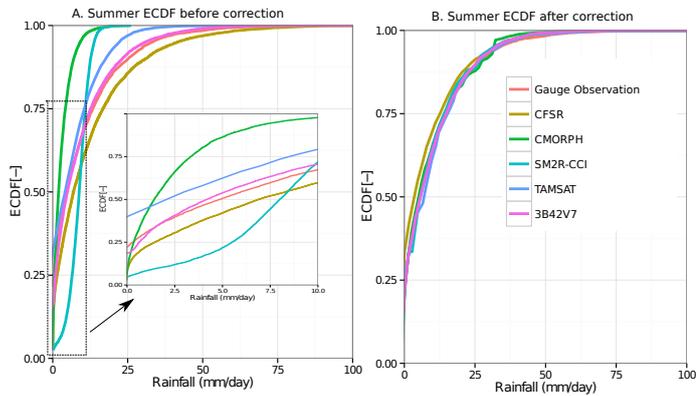


Figure 6: Cumulative distribution functions for different SREs and observed rainfall data before correction, used to construct the *ecdf* correction function (a), and SREs and observed data after the correction function applied for the validation data set (b). The zoomed-in plot in figure a is used to identify the dry and wet days for the SREs.

is important because it gives an objective decision on reducing the small rainfall values by the SREs i.e drizzling effects of the products. In the case of TAMSAT, the correction function did not improve the detection capacity of the first class (class 0) because the products have no "drizzling effect" and instead overestimate dry days in comparison to the observed rainfall data. In this case, since it has the highest detection capacity for zero class (figure 4), the raw SREs are accepted for the first one or two classes and the *ecdf* matching will be applied to high values. The correction in 3B42V7 improves the dry days estimation, however, it increases the missing rate to higher rainfall values. The result of future time series validation shows exactly the same pattern of improvement (figure 4B). The independent station and future time series validation results suggest that simple *ecdf* bias correction procedure can be used to obtain better quality SREs at non-gauge sites and in future realization of UBN basin.

Table 3: The comparison of different SREs using summary statistics (Bias) and detection capacity (accuracy index) against the gauge observed data at validation stations and time series during summer season (JJA) for about 10 years. The effect of bias correction on the other statistics (correlation coefficient and RMSE) is very small, and results are not reported here.

SREs	spatial cross validation				time series validation			
	Before correction		After correction		Before correction		After correction	
	BIAS (%)	Accuracy	BIAS	Accuracy	BIAS (%)	Accuracy	BIAS	Accuracy
CFSR	13.60	0.51	-16.00	0.52	16.40	0.52	15.20	0.51
CMORPH	-70	0.53	-4.0	0.56	-72.20	0.53	-1.10	0.55
SM2R-CCI	-12.70	0.52	-1.40	0.54	-7.70	0.51	-2.60	0.54
TAMSAT	-26.20	0.55	11.30	0.58	-23.80	0.55	11.30	0.57
3B42V7	-9.00	0.54	-1.60	0.55	-12.20	0.54	-12.80	0.55

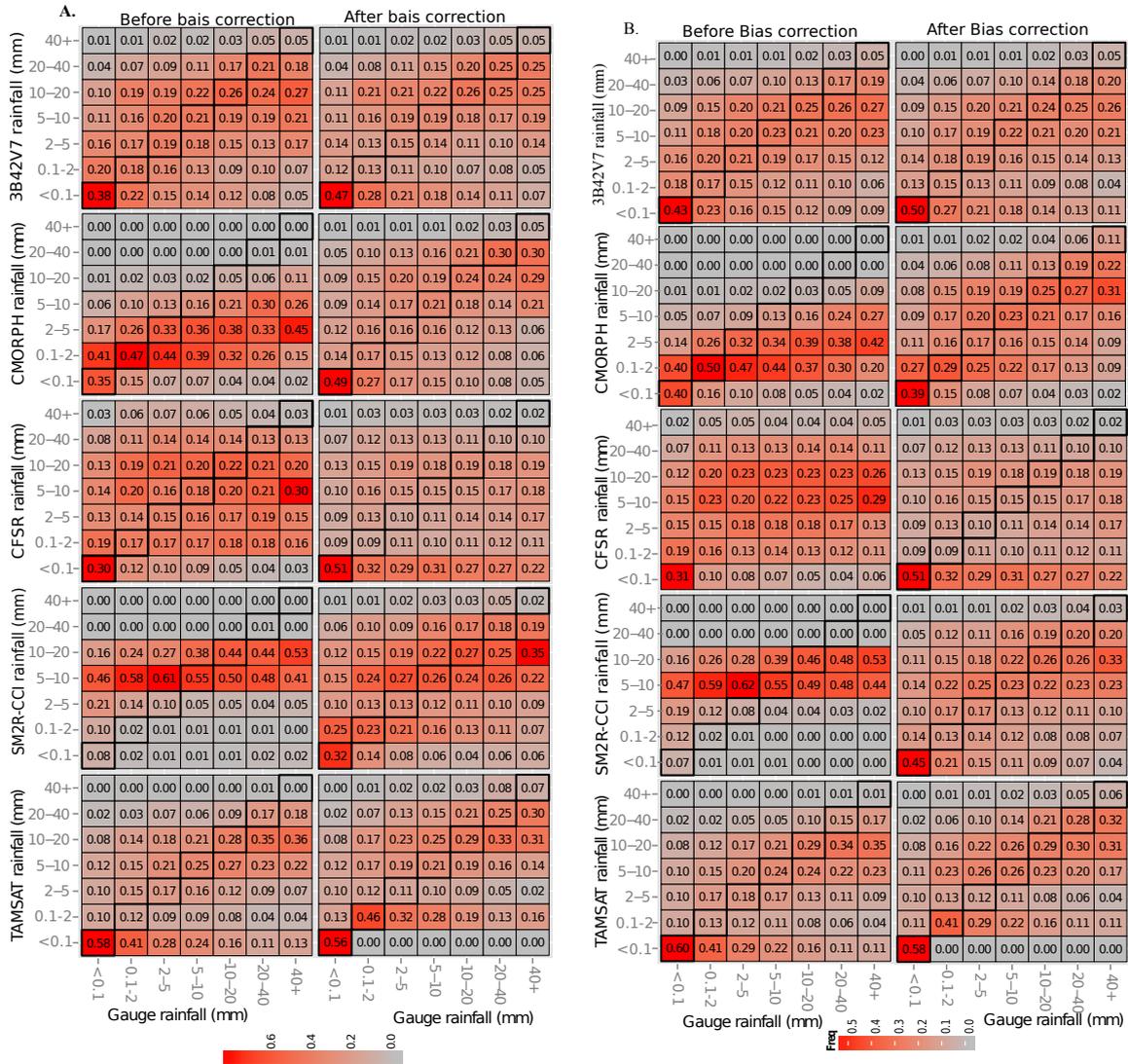


Figure 7: The effects of bias correction on the confusion matrix of different SREs (CFSR, CMORPH, SM2R-CCI, TAMSAT, and 3B42V7 from above to below) during the summer season (June, July, August) applied in independent stations (A) and in future time series (B). The first column shows the detection capacities of SREs before the bias correction and the second column is after bias correction.

6. Conclusions

This paper comparatively evaluates five satellite rainfall datasets over the Upper Blue Nile basin at daily time steps for 10 years, and uses *ecdf* bias correction technique to improve SREs. Different goodness-of-fit statistics (r , RMSE, BIAS) and confusion matrix are utilized to compare SREs and evaluate the bias correction strategy. The major findings are summarized as follows:

- The correlation coefficient based on 10 years of daily data at the whole basin scale shows similar value for TAMSAT (median=0.51), SM2R-CCI (0.5), and CMORPH (0.52), slightly better than 3B42V7 (0.47) and CFSR (0.3) performances.
- CFSR has the highest RMSE (8.9 mm/day) followed by 3B42V7 (7.5 mm/day) and CMORPH (7.3mm/day); TAMSAT (6.8 mm/day) and SM2R-CCI (6.8 mm/day) show relatively lower RMSE values.
- CMORPH has the highest BIAS (-72%), with most of the stations tends to have similar BIAS.
- The spatial distribution of GOF shows that the eastern part of the basin (highlands) has higher performances (lower RMSE and BIAS, and higher correlation) than the western part of the basin (lowlands).
- Generally, the detection skill decreases with increasing rainfall classes. TAMSAT has the highest detection skill for dry days while SM2R-CCI has better detection capacity for medium rainfall intensities (10-20 mm/day).
- The mean annual cumulative rainfall analysis shows, at some stations, that the difference amongst the various SREs is as high as about 2700 mm. Such a huge difference in rainfall is the real concern in using SREs for water balance modelling, and it shows that the choice of SRE product is a topic of real interest.
- The two validation strategy of *ecdf* bias correction, grouped cross validation and time series split validation, show that there is room for improvement with relatively small efforts. The highest and lowest bias improvement following the *ecdf* correction procedures is shown for CMORPH and CFSR, respectively.

In summary, different SREs exhibit different skills. The quantification of different SREs performances and characterization of errors is an important initial step for distributed hydrological model set-up in the basin. As it is not possible to depict the whole picture of SREs impact on the basin hydrological budget, the implementation of SREs and bias correction procedure in hydrological modeling framework in UBN basin is an important issue we would like to address in the next step.

Replicable Research

The datasets used in this study, mainly the five SREs, R codes, and supplementary material will be available for reproducing the results of the paper at <http://ecohydrogeomorphometry.blogspot.it/2015/09/satellite-rainfall-estimation-product.html>.

Acknowledgments

We would like to acknowledge the National meteorological Agency of Ethiopia for providing us the gauge rainfall data. We thank Elena Tarnavsky for providing the TAMSAT product. The Authors also thanks the projects GLOBAQUA and CLIMAWARE that partially support the first Author.

Appendix A. Stations split into calibration and validation

As shown in figure A.8, the total number of rainfall stations is split in two: those used for *ecdf* matching and generating the transfer function and those used to assess the performances of the corrected SREs.

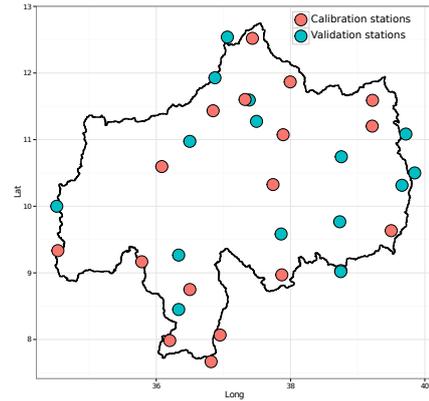


Figure A.8: The spatial distribution of stations used to generate the correction function (calibration stations) and used to evaluate the performances of the *ecdf* mapping method (validation points).

Appendix B. Goodness-of-fit

The three goodness-of-fit indexes used in the paper are defined as follows: the Pearson linear correlation coefficient (r):

$$r = \frac{1}{N} \frac{\sum_{n=1}^N (S_i - \bar{S})(G_i - \bar{G})}{\sqrt{\sum_{n=1}^N (S_i - \bar{S})^2 \sum_{n=1}^N (G_i - \bar{G})^2}}, \quad (\text{B.1})$$

where S and G are the SREs and gauge rainfall estimation respectively, N is the number of observations, and the overbar is the mean operator. The Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (S_i - G_i)^2}, \quad (\text{B.2})$$

and *BIAS*:

$$BIAS = \frac{\sum_{i=1}^N (S_i - G_i)}{\sum_{i=1}^N G_i}, \quad (B.3)$$

r and *BIAS* are dimensionless, and *RMSE* is in mm. The optimal values are 1, for *r*, 0 for *RMSE*, and 0 for *BIAS*.

References

- Abu-Zeid, M. A., Biswas, A. K., 1996. River basin planning and management. Oxford University Press.
- AghaKouchak, A., Mehran, A., Norouzi, H., Behrangi, A., 2012. Systematic and random error components in satellite precipitation data sets. *Geophysical Research Letters* 39 (9).
- Bennett, J. C., Grose, M. R., Corney, S. P., White, C. J., Holz, G. K., Katzfey, J. J., Post, D. A., Bindoff, N. L., 2014. Performance of an empirical bias-correction of a high-resolution climate dataset. *International Journal of Climatology* 34 (7), 2189–2204.
- Bewket, W., Sterk, G., 2005. Dynamics in land cover and its effect on stream flow in the chemoga watershed, blue Nile basin, Ethiopia. *Hydrological Processes* 19 (2), 445–458.
- Bitew, M. M., Gebremichael, M., Ghebremichael, L. T., Bayissa, Y. A., 2012. Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. *Journal of Hydrometeorology* 13 (1), 338–350.
- Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., Levizzani, V., 2014. Soil as a natural rain gauge: estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research: Atmospheres* 119 (9), 5128–5141.
- Brocca, L., Moramarco, T., Melone, F., Wagner, W., 2013. A new method for rainfall estimation through soil moisture observations. *Geophysical Research Letters* 40 (5), 853–858.
- Cai, Y., Jin, C., Wang, A., Guan, D., Wu, J., Yuan, F., Xu, L., 2015. Spatio-temporal analysis of the accuracy of tropical multisatellite precipitation analysis 3b42 precipitation data in mid-high latitudes of China. *PLoS one* 10 (4).
- Chadwick, R., Grimes, D., Saunders, R., Francis, P., Blackmore, T., 2010. The Tamora algorithm: Satellite rainfall estimates over West Africa using multi-spectral seviri data. *Advances in Geosciences* 25 (25), 3–9.
- Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Puca, S., Rinollo, A., Gabellani, S., Wagner, W., 2015. Integration of satellite soil moisture and rainfall observations over the Italian territory. *Journal of Hydrometeorology* (2015).
- Cohen Liechti, T., Matos, J., Boillat, J.-L., Schleiss, A., 2012. Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi river basin. *Hydrology and Earth System Sciences* 16 (2), 489–500.
- Conway, D., 1997. A water balance model of the upper blue Nile in Ethiopia. *Hydrological Sciences Journal* 42 (2), 265–286.
- Conway, D., 2000. The climate and hydrology of the upper blue Nile river. *Geographical Journal*, 49–62.
- Conway, D., 2005. From headwater tributaries to international river: observing and adapting to climate variability and change in the Nile basin. *Global Environmental Change* 15 (2), 99–114.
- Conway, D., Hulme, M., 1993. Recent fluctuations in precipitation and runoff over the Nile sub-basins and their impact on main Nile discharge. *Climatic Change* 25 (2), 127–151.
- del Jesus, M., Rinaldo, A., Rodríguez-Iturbe, I., 2015. Point rainfall statistics for ecohydrological analyses derived from satellite integrated rainfall measurements. *Water Resources Research* 51 (4), 2974–2985.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S., Ropelewski, C., 2007. Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing* 28 (7), 1503–1526.
- Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S., Ropelewski, C., 2008. Validation of high-resolution satellite rainfall products over complex terrain. *International Journal of Remote Sensing* 29 (14), 4097–4110.
- Dorigo, W., Gruber, A., De Jeu, R., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R., et al., 2015. Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment* 162, 380–395.
- Duan, Z., Bastiaanssen, W., 2013. First results from version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure. *Remote Sensing of Environment* 131, 1–13.
- Ebert, E. E., Janowiak, J. E., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society* 88 (1), 47–64.
- Ferraro, R. R., 1997. Special sensor microwave imager derived global rainfall estimates for climatological applications. *Journal of Geophysical Research: Atmospheres* (1984–2012) 102 (D14), 16715–16735.
- Fuka, D. R., Walter, M. T., MacAlister, C., Degaetano, A. T., Steenhuis, T. S., Easton, Z. M., 2014. Using the climate forecast system reanalysis as weather input data for watershed models. *Hydrological Processes* 28 (22), 5613–5623.
- Gao, Y., Liu, M., 2013. Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau. *Hydrology and Earth System Sciences* 17 (2), 837.
- Gebremichael, M., Bitew, M. M., Hirpa, F. A., Tesfay, G. N., 2014. Accuracy of satellite rainfall estimates in the Blue Nile basin: Lowland plain versus highland mountain. *Water Resources Research* 50 (11), 8775–8790.
- Gebremichael, M., Hossain, F., 2010. Satellite rainfall applications for surface hydrology. Springer.
- Grimes, D., Pardo-Iguzquiza, E., Bonifacio, R., 1999. Optimal areal rainfall estimation using rain gauges and satellite data. *Journal of Hydrology* 222 (1), 93–108.
- Habib, E., ElSaadani, M., Haile, A. T., 2012. Climatology-focused evaluation of CMORPH and TMPA satellite rainfall products over the Nile basin. *Journal of Applied Meteorology and Climatology* 51 (12), 2105–2121.
- Haile, A. T., Habib, E., Rientjes, T., 2013. Evaluation of the climate prediction center (CPC) Morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile river. *Hydrological Processes* 27 (12), 1829–1839.
- Haile, A. T., Yan, F., Habib, E., 2015. Accuracy of the CMORPH satellite-rainfall product over Lake Tana basin in Eastern Africa. *Atmospheric Research* 163, 177–187.
- Hirpa, F. A., Gebremichael, M., Hopson, T., 2010. Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *Journal of Applied Meteorology and Climatology* 49 (5), 1044–1051.
- Hossain, F., Huffman, G. J., 2008. Investigating error metrics for satellite rainfall data at hydrologically relevant scales. *Journal of Hydrometeorology* 9 (3), 563–575.
- Huffman, G. J., Bolvin, D. T., 2013. TRMM and other data precipitation data set documentation. NASA, Greenbelt, USA, 1–40.
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., Stocker, E. F., 2007. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology* 8 (1), 38–55.
- Hwang, S., Graham, W. D., 2013. Development and comparative evaluation of a stochastic analog method to downscale daily GCM precipitation. *Hydrology and Earth System Sciences* 17 (11), 4481–4502.
- Iizumi, T., Nishimori, M., Dairaku, K., Adachi, S. A., Yokozawa, M., 2011. Evaluation and intercomparison of downscaled daily precipitation indices over Japan in present-day climate: Strengths and weaknesses of dynamical and bias correction-type statistical downscaling methods. *Journal of Geophysical Research: Atmospheres* (1984–2012) 116 (D1).
- Jakob Themeßl, M., Gobet, A., Leuprecht, A., 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology* 31 (10), 1530–1544.
- Jarraud, M., 2008. Guide to meteorological instruments and methods of observation (WMO-no. 8). World Meteorological Organisation: Geneva, Switzerland.
- Jobard, I., Chopin, F., Bergès, J. C., Roca, R., 2011. An intercomparison of 10-day satellite precipitation products during West African monsoon. *International journal of remote sensing* 32 (9), 2353–2376.
- Joyce, R. J., Janowiak, J. E., Arkin, P. A., Xie, P., 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology* 5 (3), 487–503.
- Kebede, S., Travi, Y., Alemayehu, T., Marc, V., 2006. Water balance of Lake Tana and its sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia.

- Journal of hydrology 316 (1), 233–247.
- Kim, U., Kaluarachchi, J. J., Smakhtin, V. U., 2008. Generation of monthly precipitation under climate change for the upper blue Nile river basin, Ethiopia. *Journal of Hydrology* 316 (1), 233–247.
- Kisaka, M. O., Mucheru-Muna, M., Ngetich, F., Mugwe, J., Mugendi, D., Mairura, F., 2015. Rainfall variability, drought characterization, and efficacy of rainfall data reconstruction: Case of eastern Kenya. *Advances in Meteorology* 2015.
- Kumar, P., Kishitawal, C., Pal, P., 2014. Impact of satellite rainfall assimilation on weather research and forecasting model predictions over the Indian region. *Journal of Geophysical Research: Atmospheres* 119 (5), 2017–2031.
- Kummerow, C., Barnes, W., Kozu, T., Shiue, J., Simpson, J., 1998. The tropical rainfall measuring mission (trmm) sensor package. *Journal of Atmospheric and Oceanic Technology* 15 (3), 809–817.
- Kummerow, C., Simpson, J., Thiele, O., Barnes, W., Chang, A., Stocker, E., Adler, R., Hou, A., Kakar, R., Wentz, F., et al., 2000. The status of the tropical rainfall measuring mission (trmm) after two years in orbit. *Journal of Applied Meteorology* 39 (12), 1965–1982.
- Laurent, H., Jobard, I., Toma, A., 1998. Validation of satellite and ground-based estimates of precipitation over the Sahel. *Atmospheric Research* 47, 651–670.
- Lázaro, R., Rodrigo, F., Gutiérrez, L., Domingo, F., Puigdefábregas, J., 2001. Analysis of a 30-year rainfall record (1967–1997) in semi-arid Spain for implications on vegetation. *Journal of Arid Environments* 48 (3), 373–395.
- Liu, Z., Ostrenga, D., Teng, W., Kempler, S., 2012. Tropical rainfall measuring mission (trmm) precipitation data and services for research and applications. *Bulletin of the American Meteorological Society* 93 (9), 1317–1325.
- Liu, Z., Rui, H., Teng, W. L., Chiu, L. S., Leptoukh, G., Vicente, G. A., 2007. Online visualization and analysis: A new avenue to use satellite data for weather, climate, and interdisciplinary research and applications. In: *Measuring Precipitation From Space*. Springer, pp. 549–558.
- Maidment, R. I., Grimes, D., Allan, R. P., Tarnavsky, E., Stringer, M., Hewison, T., Roebeling, R., Black, E., 2014. The 30 year tamsat African rainfall climatology and time series (tarcat) data set. *Journal of Geophysical Research: Atmospheres* 119 (18), 10–619.
- Maidment, R. I., Grimes, D. I., Allan, R. P., Greatrex, H., Rojas, O., Leo, O., 2013. Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda. *Meteorological Applications* 20 (3), 308–317.
- Maurer, E. P., Pierce, D. W., 2014. Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean. *Hydrology and Earth System Sciences* 18 (3), 915–925.
- Mellander, P.-E., Gebrehiwot, S. G., Gardenas, A. I., Bewket, W., Bishop, K., 2013. Summer rains and dry seasons in the upper blue Nile basin: the predictability of half a century of past and future spatiotemporal patterns. *PLoS ONE* 8 (7), 1932–6203.
- Michelangeli, P.-A., Vrac, M., Loukos, H., 2009. Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophysical Research Letters* 36 (11).
- Mishra, A., Hata, T., 2006. A grid-based runoff generation and flow routing model for the upper blue Nile basin. *Hydrological Sciences Journal* 51 (2), 191–206.
- Mishra, A., Hata, T., Abdelhadi, A., 2004. Models for recession flows in the upper blue Nile river. *Hydrological Processes* 18 (15), 2773–2786.
- Moon, S.-E., Ryoo, S.-B., Kwon, J.-G., 1994. A Markov chain model for daily precipitation occurrence in South Korea. *International Journal of Climatology* 14 (9), 1009–1016.
- Panofsky, H. A., Brier, G. W., Best, W. H., 1958. Some application of statistics to meteorology.
- Porcù, F., Milani, L., Petracca, M., 2014. On the uncertainties in validating satellite instantaneous rainfall estimates with rain gauge operational network. *Atmospheric Research* 144, 73–81.
- Prakash, S., Mitra, A. K., Pai, D., 2015. Comparing two high-resolution gauge-adjusted multisatellite rainfall products over India for the southwest monsoon period. *Meteorological Applications*.
- Rientjes, T., Haile, A., Kebede, E., Mannaerts, C., Habib, E., Steenhuis, T., 2011. Changes in land cover, rainfall and stream flow in upper Gilgel Abay catchment, blue Nile basin—Ethiopia. *Hydrology and Earth System Sciences* 15 (6), 1979–1989.
- Romilly, T. G., Gebremichael, M., 2010. Evaluation of satellite rainfall estimates over Ethiopian river basins. *Hydrology and Earth System Sciences Discussions* 7, 7669–7694.
- Romilly, T. G., Gebremichael, M., 2011. Evaluation of satellite rainfall estimates over Ethiopian river basins. *Hydrology and Earth System Sciences* 15 (5), 1505–1514.
- URL <http://www.hydro1-earth-syst-sci.net/15/1505/2011/>
- Rozante, J. R., Moreira, D. S., de Goncalves, L. G. G., Vila, D. A., 2010. Combining TRMM and surface observations of precipitation: technique and validation over South America. *Weather and Forecasting* 25 (3), 885–894.
- Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., et al., 2010. The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society* 91 (8), 1015–1057.
- Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., et al., 2014. The NCEP climate forecast system version 2. *Journal of Climate* 27 (6), 2185–2208.
- Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van den Dool, H., Pan, H.-L., Moorthi, S., Behringer, D., et al., 2006. The NCEP climate forecast system. *Journal of Climate* 19 (15), 3483–3517.
- Salio, P., Hobouchian, M. P., Skabar, Y. G., Vila, D., 2015. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmospheric Research* 163, 146–161.
- Schmidli, J., Frei, C., 2005. Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century. *International Journal of Climatology* 25 (6), 753–771.
- Seleshi, Y., Zanke, U., 2004. Recent changes in rainfall and rainy days in Ethiopia. *International Journal of Climatology* 24 (8), 973–983.
- Shaw, E. M., Beven, K. J., Chappell, N. A., Lamb, R., 2010. *Hydrology in practice*. CRC Press.
- Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H. V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society* 81 (9), 2035–2046.
- Steenhuis, T. S., Collick, A. S., Easton, Z. M., Leggesse, E. S., Bayabil, H. K., White, E. D., Awulachew, S. B., Adgo, E., Ahmed, A. A., 2009. Predicting discharge and sediment for the Abay (Blue Nile) with a simple model. *Hydrological Processes* 23 (26), 3728–3737.
- Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R. P., Stringer, M., Chadwick, R., Kayitakire, F., 2014. Extension of the Tamsat satellite-based rainfall monitoring over Africa and from 1983 to present. *Journal of Applied Meteorology and Climatology* 53 (12), 2805–2822.
- Tekleab, S., Uhlenbrook, S., Mohamed, Y., Savenije, H., Temesgen, M., Wenginger, J., 2011. Water balance modeling of upper blue Nile catchments using a top-down approach. *Hydrology and Earth System Sciences*, 15(7).
- Thiemeßl, M. J., Gobiet, A., Heinrich, G., 2012. Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change* 112 (2), 449–468.
- Thorne, V., Coakley, P., Grimes, D., Dugdale, G., 2001. Comparison of Tamsat and CPC rainfall estimates with rain gauges, for southern Africa. *International Journal of Remote Sensing* 22 (10), 1951–1974.
- Tian, Y., Peters-Lidard, C. D., Choudhury, B. J., Garcia, M., 2007. Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *Journal of Hydrometeorology* 8 (6), 1165–1183.
- Uhlenbrook, S., Mohamed, Y., Gragne, A., 2010. Analyzing catchment behavior through catchment modeling in the Gilgel Abay, upper blue Nile river basin, Ethiopia. *Hydrology and Earth System Sciences* 14 (10), 2153–2165.
- Wale, A., Rientjes, T., Gieske, A., Getachew, H., 2009. Ungauged catchment contributions to Lake Tana's water balance. *Hydrological Processes* 23 (26), 3682–3693.
- Wang, W., Xie, P., Yoo, S.-H., Xue, Y., Kumar, A., Wu, X., 2011. An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Climate Dynamics* 37 (7-8), 1601–1620.
- Ward, E., Buytaert, W., Peaver, L., Wheatler, H., 2011. Evaluation of precipitation products over complex mountainous terrain: A water resources perspective. *Advances in Water Resources* 34 (10), 1222–1231.
- Wilcke, R. A. I., Mendlik, T., Gobiet, A., 2013. Multi-variable error correction of regional climate models. *Climatic Change* 120 (4), 871–887.
- Worqlul, A. W., Maathuis, B., Adem, A. A., Demissie, S. S., Langan, S., Steenhuis, T. S., 2014. Comparison of rainfall estimations by TRMM 3B42, MPEP and CFSR with ground-observed data for the Lake Tana basin in Ethiopia. *Hydrology and Earth System Sciences* 18 (12), 4871–4881.
- Xue, Y., Huang, B., Hu, Z.-Z., Kumar, A., Wen, C., Behringer, D., Nadiga, S., 2011. An assessment of oceanic variability in the NCEP climate forecast system reanalysis. *Climate Dynamics* 37 (11-12), 2511–2539.

- Young, M. P., Williams, C. J., Chiu, J. C., Maidment, R. I., Chen, S.-H., 2014. Investigation of discrepancies in satellite rainfall estimates over ethiopia. *Journal of Hydrometeorology* 15 (6), 2347–2369.
- Zhao, L., Weng, F., 2002. Retrieval of ice cloud parameters using the advanced microwave sounding unit. *Journal of Applied Meteorology* 41 (4), 384–395.