Hydrol. Earth Syst. Sci. Discuss., 11, 11489–11531, 2014
www.hydrol-earth-syst-sci-discuss.net/11/11489/2014/
doi:10.5194/hessd-11-11489-2014

Hydrology and Earth System Sciences
Open Access
Discussions

Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge datasets (2002–2012)

O. P. Prat[1] **and B. R. Nelson**[2]

[1]Cooperative Institute for Climate and Satellites-NC (CICS-NC), North Carolina State University, and NOAA/National Climatic Data Center, Asheville, NC, USA
[2]Remote Sensing Applications Division (RSAD), NOAA/NESDIS/NCDC, Asheville, NC, USA

Received: 5 September 2014 – Accepted: 8 September 2014 – Published: 16 October 2014

Correspondence to: O. P. Prat (olivier.prat@noaa.gov)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

We use a suite of quantitative precipitation estimates (QPEs) derived from satellite, radar, and surface observations to derive precipitation characteristics over CONUS for the period 2002–2012. This comparison effort includes satellite multi-sensor datasets (bias-adjusted TMPA 3B42, near-real time 3B42RT), radar estimates (NCEP Stage IV), and rain gauge observations. Remotely sensed precipitation datasets are compared with surface observations from the Global Historical Climatology Network (GHCN-Daily) and from the PRISM (Parameter-elevation Regressions on Independent Slopes Model). The comparisons are performed at the annual, seasonal, and daily scales over the River Forecast Centers (RFCs) for CONUS. Annual average rain rates present a satisfying agreement with GHCN-D for all products over CONUS (±6 %). However, differences at the RFC are more important in particular for near-real time 3B42RT precipitation estimates (−33 to +49 %). At annual and seasonal scales, the bias-adjusted 3B42 presented important improvement when compared to its near real time counterpart 3B42RT. However, large biases remained for 3B42 over the Western US for higher average accumulation ($\geq 5\,\mathrm{mm\,day^{-1}}$) with respect to GHCN-D surface observations. At the daily scale, 3B42RT performed poorly in capturing extreme daily precipitation (> 4 in day-1) over the Northwest. Furthermore, the conditional analysis and the contingency analysis conducted illustrated the challenge of retrieving extreme precipitation from remote sensing estimates.

1 Introduction

Remotely sensed precipitation products are now coming of age where they can be considered for climatological applications. Over the last 30 years, numerous long-term rainfall datasets were developed using rain gauge (RG) precipitation measurements, remotely sensed (ground based radars, satellites) quantitative precipitation estimates (QPE), or combining different sensors, each of which has specific characteristics

HESSD
11, 11489–11531, 2014

and limitations. Extensive information on precipitation measurement methodologies and available precipitation products can be found in Michaelides et al. (2009), Kidd et al. (2010), and Tapiador et al. (2012) among others. One of the limitations in using rain gauge based precipitation datasets lies in the fact that the geographical coverage is not spatially homogeneous. By contrast, multi-sensor satellite-based products: PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks: Sorooshian et al., 2000) and variants PERSIANN-CDR (Climate Data Record: Ashouri et al., 2014), CMORPH (CPC MORPHing technique: Joyce et al., 2004), and TRMM (Tropical Rainfall Measuring Mission) TMPA (TRMM Multisatellite Precipitation Analysis: Huffman et al., 2007) or ground-based radar rainfall estimates: NCEP (National Centers for Environmental Prediction) Stage IV (Lin and Mitchell, 2005) or more recently the National Mosaic and Multi-sensor QPE (NMQ/Q2) (Zhang et al., 2011), provide an opportunity to broach the problem of sparse observations over land and/or ocean. Precipitation datasets at high spatial (typically 4–25 km) and temporal (1–6 h) resolution, allow for assessing annual, seasonal, and daily characteristics of precipitation at local, regional, and continental scales (Huffman et al., 2001; Sorooshian et al., 2002; Nesbitt and Zipser, 2003; Liu and Zipser, 2008; Nesbitt and Anders, 2007; Sapiano and Arkin, 2009; Prat and Barros, 2010; Sahany et al., 2010; Kidd et al., 2012; Prat and Nelson, 2013a, b, 2014 among others).

The purpose of this study is to evaluate the ability of QPE products to describe precipitation patterns and capture precipitation extremes in a climatological perspective. While a lot of studies are available that compared different radar/satellite products on an event-to-event basis, in this work we focus on the long-term perspective (11 years). The objective of this study is to provide a comparison of a suite of common quantitative precipitation estimates derived from satellites, radars, and rain gauges datasets for the period 2002–2012 over CONUS. Our aim is to evaluate the ability of satellite (TMPA 3B42, 3B42RT) and ground based remotely sensed (Stage IV) precipitation products to describe precipitation patterns and provide a long-term picture of the evolution of precipitation over time. In particular, we will investigate how the

HESSD
11, 11489–11531, 2014

different QPE products compare with respect to long-term surface observations and what are the associated uncertainties. The choice of 3B42 is guided by the fact that a monthly accumulation adjustment is performed on the near-real time algorithm 3B42RT and thus provides bias-adjusted precipitation estimates when compared to non-adjusted versions of CMORPH and PERSIANN. Furthermore, there is a fair amount of studies available that compare the respective merit of the datasets described above either against each other or against other datasets used as reference. Those studies often investigate isolated events such as intense precipitation or focus on a time period rather limited in time (day, month, season). Seldom studies that deal with the long-term assessment of precipitation products (annual or multi-annual basis) are available in the scientific literature. Therefore, this paper proposes to evaluate satellite precipitation estimates in the perspective of climate applications. The remotely sensed datasets will be compared against surface observations from the Global Historical Climatology Network) (GHCN-Daily) and estimations from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) which combines surface observations with a digital elevation model to account for the orographic enhancement of precipitation. Both GHCN-D and PRISM will be used as a baseline for QPE products evaluations. The study will analyze eleven years (2002–2012) of rainfall data over CONUS. Although 11 years is not a long climatology, the duration of the study will be informative enough to derive long-term trends, assess systematic biases, and capture year-to-year and seasonal variability. In addition to long-term average precipitation characteristics, we will investigate the ability for each of those QPE products to capture extreme events and how they compare with surface observations.

The paper is organized as follows. In a first section, we present briefly the precipitation datasets used in this study. In a second section, we will present a comparison between precipitation estimates at the annual and seasonal scales. In a third part, we will investigate the impact of differing spatial and temporal resolutions with respect to the datasets' ability to capture extreme precipitation events. Finally, the paper wraps up summarizing the major results of this study.

HESSD
11, 11489–11531, 2014

2 Precipitation datasets and algorithms description

In this section, we provide a brief description of these different precipitation datasets used. The interested reader will refer to the references cited.

2.1 Rain gauge precipitation datasets: GHCN-Daily

Precipitation surface observations are taken from the Global Historical Climatology Network-Daily (GHCN-D). The dataset gathers records from over 80 000 stations over 180 countries. About two-thirds of those stations report total daily precipitation only and other stations include additional information such as maximum and minimum temperature, snowfall, and snow depth (Menne et al., 2012). The entire dataset is routinely quality controlled to ensure basic consistency of the dataset. Figure 1a presents the location of the 8815 surface observations in the GHCN-D database over CONUS. For the current study, only the 4075 rain gauges reporting at least 90 % of the time during the period 2002–2012 are selected to ensure stable statistics (Fig. 1b). Although there is a 50 % decrease from the total number of rain gauges, the remaining rain gauges conserved a comparable spatial distribution than the original network and the removed gauges were evenly distributed throughout CONUS. The surface stations are compared with the nearest pixel of the gridded precipitation estimates derived from the selected datasets (PRISM, Stage IV, 3B42, 3B42RT) described below.

2.2 Rain gauge gridded precipitation datasets: PRISM

The PRISM algorithm (available at http://www.prism.oregonstate.edu/) combines point data with a digital elevation model to generate gridded estimates of precipitation along with a suite of climatological variables such as temperature, snowfall, degree dew point, among others (Daly et al., 1994). Data are available at the daily, monthly, and annual scale and at various spatial resolutions (800 m–4 km). In this work, we use the monthly precipitation estimates at the 4 km nominal spatial resolution (dataset AN81m:

HESSD
11, 11489–11531, 2014

PRISM Technical Note, 2014). The PRISM precipitation estimates incorporate surface data observations from GHCN-D among others and the systematic comparison of point surface observations from GHCN-D and gridded estimates from PRISM will be performed as a consistency check. The PRISM precipitation estimates will be used as a baseline data set to evaluate remotely sensed precipitation products (Stage IV, 3B42, 3B42RT) at the annual and seasonal scale.

2.3 Radar precipitation datasets: the Stage IV analysis

The NCEP Stage IV product, herein referred to as Stage IV, is a near real time product that is generated at NCEP separately from the NWS Precipitation Processing System (PPS) and the NWS River Forecast Center (RFC) rainfall processing. Originally the Stage IV product was intended for assimilation into atmospheric forecast models to improve quantitative precipitation forecasts (QPF) (Lin and Mitchell, 2005). However the length of record, consistency of data availability, and ease of access have made the Stage IV product attractive for many applications. Data are available in GRIB format for hourly, 6 hourly, and daily temporal scales and they are gridded on the Hydrologic Rainfall Analysis Projection (HRAP) (Reed and Maidment, 1995, 1999) at a nominal 4 km spatial resolution. Figure 1c presents the geographical extent of the 12 RFCs and Table 1 reports the number of available rain gauges by RFC. The Stage IV precipitation data are available via: (http://data.eol.ucar.edu/codiac/dss/id=21.093). The reader will find a more detailed description of the Stage IV precipitation estimates generation from the RFC level and up to the final mosaicked product as well as related artifacts and uncertainties in Nelson et al. (2014).

2.4 Satellite precipitation QPE datasets: TMPA 3B42 and 3B42RT

The satellite QPE TMPA 3B42 Version 7 (V7) blends optimally different remotely sensed microwave such as the TRMM Microwave Imager (TMI), the Special Sensor Microwave Imager (SSM/I), the Advanced Microwave Scanning Radiometer (AMSR),

HESSD
11, 11489–11531, 2014

and the Advanced Microwave Sounding Unit (AMSU), along with calibrated IR estimates of rain gauge corrected monthly accumulation (Huffman et al., 2007). TMPA 3B42 provides precipitation estimates for the domain 50° S–50° N at a 3 hourly and quarter degree resolution (0.25° × 0.25°) from which seasonal, daily, and sub-daily precipitation characteristics can be derived. Over the years, the retrieval algorithms of the different products incorporated within 3B42 were modified. The algorithm 3B42 itself had several versions and a major improvement of the precipitation estimates was provided in 2007 to correct for low biases (Huffman et al., 2007). The 3B42 V7 represents substantial improvement when compared to the previous Version 6 (V6). The version 7 incorporates additional satellite products along with the reprocessed versions of the merged algorithms (AMSU, TMI, AMSR, SSM/I). However, the major upgrade consists of the use of a single, uniformly processed surface precipitation gauge analysis from the Global Precipitation Climatology Centre (GPCC) (Huffman and Bolvin, 2013). The use of the GPCC rain gauge analysis explains most of the differences observed between V6 and V7 over land and over coastal areas. A brief comparison of both versions was provided for the period 1998–2009 over North and Central America, which encompasses the current CONUS domain, in Prat and Nelson (2013b, see Fig. B1). In this work, we also use the near real time version of the product (3B42RT), which is produced operationally and does not use the monthly rain gauge correction (GPCC) but incorporates an a-priori climatological correction (Huffman et al., 2007). In addition to products relying on gauge measurements (GHCN-D, PRISM) and incorporating gauge information for bias adjustment purposes (Stage IV, TMPA 3B42), the use of the near real time dataset 3B42RT has a double objective. First it provides a quantification of the biases adjustment in the perspective of long-term climatological applications. Second, it aims to infer the suitability of precipitation products to capture precipitation extremes in near real time.

HESSD
11, 11489–11531, 2014

3 Annual characteristic of precipitation: differences between datasets

3.1 Annual average precipitation

Figure 2 displays the annual average precipitation derived from PRISM (Fig. 2a), Stage IV (Fig. 2b), 3B42 (Fig. 2c), and 3B42RT (Fig. 2d) for the period 2002–2012. All datasets present comparable precipitation patterns with higher rainfall east of 97° W, over the southeast, and over the Pacific Northwest. Precipitation derived from Stage IV displays a closer agreement with PRISM with comparable rainfall over the Northwest, and over the Rockies. The adjusted 3B42 presents a better visual agreement with PRISM and Stage IV than the near real time version 3B42RT. However, rainfall over the Pacific Northwest is noticeably lower than retrieved from PRISM and Stage IV. The effect of monthly accumulation correction between 3B42 and 3B42RT is particularly noticeable over the Northwest, the Rockies, and the Northeast. Over the Northeast, the annual average precipitation differences between 3B42 and 3B42RT are above $+2\,\mathrm{mm\,day^{-1}}$. Annual average precipitation differences between bias-adjusted and non-adjusted datasets are about $+1\,\mathrm{mm\,day^{-1}}$ over the Northeast. Those differences are about $-1.5\,\mathrm{mm\,day^{-1}}$ over the Rockies. CONUS-wide the mean average annual precipitation for the unadjusted 3B42RT is $2.62\,\mathrm{mm\,day^{-1}}$; for 3B42 it is $2.54\,\mathrm{mm\,day^{-1}}$ (3 % difference).

3.2 Comparison with surface observations

Figure 3 displays the scatterplots along with the Q–Q (Quantile–Quantile) plots for annual average precipitation derived from PRISM, Stage IV, 3B42, and 3B42RT when compared to GHCN-D network. Over CONUS (Fig. 3a), we observe a very good agreement between surface observations from the GHCN-D network and PRISM (a = 0.98; R^2 = 0.98) as expected due to the fact that PRISM gridded precipitation estimates incorporate GHCN-D stations. Values for the mean annual average precipitation and the associated standard deviation (σ) are relatively close at $2.42\,\mathrm{mm\,day^{-1}}$

HESSD
11, 11489–11531, 2014

(σ = 1.11 mm day^{-1}) for PRISM and 2.47 mm day^{-1} (σ = 1.14 mm day^{-1}) for GHCN-D. The differences observed toward higher rain rates (R > 6 mm day^{-1}) are due to the algorithm that uses a digital elevation model and incorporates complex precipitation processes such as rain shadows and coastal effects among others. Comparison of Stage IV estimates with GHCN-D displays an overall satisfying agreement (a = 0.93; R^2 = 0.93) with lower precipitation estimates for stage IV for rain rates greater than 4 mm day^{-1}. Stage IV displays a lower mean average precipitation (−6 %) when compared to GHCN-D surface stations (4075 rain gauges). The satellite QPEs (3B42, 3B42RT) display the highest mean annual average precipitation over CONUS when compared to other precipitation estimates (GHCN-D, PRISM, Stage IV) with 2.54 mm day^{-1} for 3B42 and 2.62 mm day^{-1} for 3B42RT along with a lower correlation coefficient (Fig. 3a). However, while the mean annual average precipitation is higher than surface observations, 3B42 and 3B42RT display negative biases in the upper part of the distribution (R > 4 mm day^{-1}) as revealed by the Q–Q plots. In addition, the bias-adjusted 3B42 presents a better agreement with surface observations (a = 1.00; R^2 = 0.83) than the near real time precipitation estimates from 3B42RT (a = 0.99; R^2 = 0.36). Overall, a better agreement is found for Stage IV than for the satellite estimates (3B42, 3B42RT) in the upper part of the distribution.

The differences between surface observations (GHCN-D) and the precipitation datasets (PRISM, Stage IV, 3B42, 3B42RT) vary greatly when considering RFCs separately. For instance, the Lower Mississippi (LM) displays a good agreement regardless of the dataset considered (Fig. 3b). PRISM (3.64 mm day^{-1}) presents the best agreement with GHCN-D (3.75 mm day^{-1}). Little differences are found between 3B42 and 3B42RT precipitation estimates with an average rain rate of 3.87 mm day^{-1} and 3.90 mm day^{-1} respectively. The bias-adjusted 3B42 however presents a narrower distribution (lower σ) than for 3B42RT. Furthermore, 3B42 displays a better agreement with GHCN-D than Stage IV (3.48 mm day^{-1}). Over the Missouri Basin River (MB), Stage IV (1.59 mm day^{-1}) and 3B42 (1.74 mm day^{-1}) display a very good agreement with GHCN-D (1.59 mm day^{-1}), while the near real time 3B42RT present a rain rate

HESSD
11, 11489–11531, 2014

50 % higher (2.37 mm day^{-1}). Another example of the differences observed at the RFC level can be found over Northwest (NW). While Stage IV display a good agreement with GHCN-D (and PRISM), the satellite estimates 3B42 and 3B42RT severely underestimate precipitation above 2 mm day^{-1} with average precipitation respectively 20 % and 25 % lower than GHCN-D (Fig. 3c). The monthly-adjusted accumulation 3B42 presents a substantial improvement when compared with the near real time 3B42RT for accumulation higher than 2 mm day^{-1}. Table 2 summarizes the differences between GHCN-D and the different datasets. For PRISM, the linear regression coefficient when compared to surface observation (GHCN-D) remains within a narrow range (0.97–1.03) for the different RFCs considered. For Stage IV, the variations are greater and indicate a general underestimation with (a) varying between 0.87 and 1. The bias-adjusted 3B42 presents a wider variation range ($0.63 < a < 1.11$), which is noticeably narrower than the coefficient obtained with the near real time precipitation estimates 3B42RT ($0.52 < a < 1.42$). Furthermore, apart from two RFCs (CN and OH), the adjusted 3B42 has a linear regression coefficient (a) closer to unity than 3B42RT as well as a better R^2 coefficient (except for NE).

Figure 4a displays the average annual precipitation derived from all datasets (GHCN-D, PRISM, Stage IV, 3B42, 3B42RT) for the different RFCs. Over CONUS, GHCN-D (2.47 mm day^{-1}) and PRISM (2.42 mm day^{-1}) present comparable rain rates with a relative difference of 2 %. Stage IV presents a slightly lower average annual rain rate (2.31 mm day^{-1}) with a difference of −7 and −4 % when compared to GHCN-D and PRISM respectively. Satellite QPEs exhibit higher rain rates with 2.52 mm day^{-1} for 3B42 and 2.62 mm day^{-1} for 3B42RT. As expected, 3B42 presents a lower difference when compared to surface observation with a difference of +2 and +4 % with GHCN-D and PRISM respectively. For the near real time 3B42RT, the difference with surface observations is +5.7 % (GHCN-D) and +7.6 % (PRISM). The Lower Mississippi River Basin (LM) exhibits the higher average annual rain rate regardless of the dataset (GHCN-D, PRISM, Stage IV, 3B42, 3B42RT) while the Colorado Basin River (CB)

HESSD
11, 11489–11531, 2014

displays the lower average annual rain rate for GHCN-D, Stage IV, and 3B42 (while the minimum for 3B42RT is found for California Nevada (CN).

The differences (%) with respect to GHCN-D surface measurements are presented in Fig. 4b. Over CONUS differences are found between −6.4 % (St. IV) and +6.1 % (3B42RT). For PRISM differences are below 4 % regardless of the RFC considered. The fact that GHCN-D and PRISM agree relatively well allows us to consider PRISM as a baseline for comparison with radar (Stage IV) and remotely sensed satellite (3B42, 3B42RT) quantitative precipitation estimates. Over the Midwestern US (AB, MB, NC), 3B42 displays higher rain rates (+8.3 %). The bias-adjusted 3B42 shows a better agreement with GHCN-D when compared to 3B42RT that shows accumulation +29 % higher than surface observations over the Midwest. Conversely, 3B42 exhibits an underestimation (−16.8 %) over the Western US (CN, NW) when compared to GHCN-D and is comparable with differences observed for 3B42RT (−10 %). Apart from the Ohio River Basin (OH) and the Southeast (SE), precipitation estimates from 3B42 are in better agreement with GHCN-D than 3B43RT as expected due to the bias-corrected accumulation. Over CONUS, the precipitation estimates derived from 3B42 and 3B42RT are relatively close with a slightly lower rain rate for 3B42 (−4 %). The magnitude of the bias adjustment (difference between 3B42RT and 3B42) remains below 7 % over most of the basins (AB: +5 %, CN: −7 %, LM: +0.7 %, NW: −7 %, OH: −3 %, SE: −3 %, WG: +4 %). This can be explained by the fact that 3B42RT uses an a-priori bias adjustment based on climatological for the near real time algorithm (Huffman et al., 2007). Important bias correction is performed over the Midwest (MB) with a +38 % difference between 3B42RT and 3B42, reducing the differences with GHCN-D from +49 % down to +9 % for 3B42RT and 3B42 respectively. This correction is to account for the overestimation of summertime convection by Passive Microwave retrieval that tends to associate important sub-cloud evaporation with precipitation. The Colorado Basin (CB) RFC is the domain that displays the most important difference between 3B43RT and 3B42 (+42 %). For the remaining RFCs, the differences between 3B42RT and 3B42 remains moderate between −21 and +12 % (MA: −17 %, NC:

HESSD
11, 11489–11531, 2014

+12 %, NE: −21 %). Stage IV presents globally lower differences with GHCN-D (−14 to +1 %) and PRISM (−17 to +4 %). Lower differences can be explained by the fact that some RFCs use a combination of radar and rain gauge estimates or rain gauges only over basins where beam blockage from surrounding mountains render delicate precipitation estimates from surface radars (Nelson et al., 2014). Similarly for satellite QPE (3B42, 3B42RT), higher differences are observed for the Western US (−13 % over CN and −14 % over NW) due to the difficulty of precipitation radar retrieval over mountainous areas. Furthermore, Stage IV QPE procedure incorporate different surface observations network such as the near-real time METAR estimates from the ASOS rain gauges sites complemented by estimates obtained from the HADS automated gauges (Lin and Mitchell, 2005). Overall, lower rainfall (−7 %) is observed for Stage IV than for GHCN-D. Results presented here are for an annual basis. Next we investigate the precipitations characteristics at the seasonal scale.

4 Seasonal precipitation

Figure 5 displays the seasonal precipitation for winter (DJF: left) and summer (JJA: right) for PRISM (Fig. 5a), Stage IV (Fig. 5b), 3B42 (Fig. 5c), and 3B42RT (Fig. 5d). A simple comparison shows that PRISM and Stage IV present comparable precipitation patterns regardless of the season. When compared to the near real time 3B42RT, the monthly-adjusted 3B42 displays precipitation patterns that are visually closer to those of PRISM. Differences between near real time 3B42RT and bias adjusted 3B42 are more emphasized on a seasonal basis than observed for the annual basis (Fig. 2). For winter, a negative bias adjustment between 3B42RT and 3B42 (3B42 < 3B42RT) is performed over the Rockies (CB), over the highest latitudes along the US/Canadian border (NC, MB, NW), and East of the Mississippi (LM, SE). Conversely, a positive bias adjustment (3B42 > 3B42RT) is performed along the West coast from Northern California up to the Pacific Northwest (NW, CN). For summer, important negative bias adjustment (3B42 < 3B42RT) is found over the Midwest (MB, NC, AB) and corrects for

HESSD
11, 11489–11531, 2014

the overestimation of summertime convection by PMW sensors that mistake sub-cloud evaporation for precipitation. The portion of the Lower Mississippi domain (LM) located East of the Mississippi displays a negative bias adjustment.

Figure 6a and b present the seasonal rain-rates derived from the different datasets (GHCN-D, PRISM, Stage IV, 3B42, 3B42RT). Between the warm and cold season, average seasonal rain rate derived from surface observations (GHCN-D, PRISM) vary from −95 % (CN) to +270 % (MB). More information on the quantitative differences between datasets can be found in Table 3. Table 3 summarizes the average seasonal rain rate for the different RFCs. For winter, the minimum (maximum) average rain rate is found for MB (LM) with 0.64 mm day^{-1} (3.77 mm day^{-1}). For summer, the minimum (maximum) average rain rate is found for CN (SE) with 0.17 mm day^{-1} (4.53 mm day^{-1}). Seasonal differences between GHCN-D and PRISM remain moderate (−5.9 to +2.1 %) and are comparable to that for the annual basis (Fig. 6c and d). For Stage IV, differences with GHCN-D vary from −18 to −2 % (overall underestimation) for winter and from −28 to +8 % (overall underestimation) for summer. For 3B42, the differences with GHCN-D range from −38 to +25 % (no overall under/overestimation) in winter. For summer, differences between 3B42 and GHCN-D present a narrower range from −2 to +25 % (overall overestimation). Those differences represent a substantial improvement with those observed for the near real time 3B42RT, which vary from −49 to +147 % in winter and from −4 to +92 % in summer (Table 3). The differences between 3B42RT and 3B42 are the most important for MR and CR in winter (+111 and +58 % respectively), or CN and MR in summer (+54 and +49 % respectively). The situations where the highest differences are observed correspond to significant positive biases of 3B42RT (Table 3).

A closer insight at seasonal differences can be seen in Fig. 7 that displays scatterplots and Q–Q plots for the seasonal rain rate for PRISM, Stage IV, 3B42, and 3B42RT with respect to GHCN-D over Northwest (NW). Regardless of the season (winter: Fig. 7a; summer: Fig. 7b), PRISM presents a very good agreement with surface observations with linear regression and R^2 coefficients close to 1 (a = 1.02;

HESSD
11, 11489–11531, 2014

R^2 = 0.96/0.97). The differences in terms of average rain rates over the 242 rain gauges (DJF) and 295 rain gauges (JJA) between GHCN-D and PRISM are about 2–3 %. There is a four-fold difference between the maximum rain rates for winter ($R \approx 12\,\mathrm{mm\,day^{-1}}$) and summer ($R \approx 3\,\mathrm{mm\,day^{-1}}$). For winter, Stage IV displays a moderate underestimation (−12 %) when compared to GHCN-D. TMPA 3B42 presents a rainfall distribution heavily skewed toward lower rain rates ($R < 6\,\mathrm{mm\,day^{-1}}$) when compared with GHCN-D and PRISM ($R > 12\,\mathrm{mm\,day^{-1}}$). Despite performing monthly-corrected accumulation for 3B42, a strong negative bias remains with a mean seasonal averaged rain rate about −30 % lower when compared to GHCN-D and PRISM. For rain rates greater than $4\,\mathrm{mm\,day^{-1}}$, the bias-adjusted 3B42 present a significant improvement when compared to the near real time 3B42RT. However, both 3B42 and 3B42RT display a comparable mean seasonal averaged precipitation ($\approx 2.5\,\mathrm{mm\,day^{-1}}$), which is about −30 % when compared to surface observations. Summer exhibits average rain rates ($\approx 0.82\,\mathrm{mm\,day^{-1}}$) 4 times lower than winter ($3.5\,\mathrm{mm\,day^{-1}}$) as reported by GHCN-D and PRISM (Fig. 7b). Stage IV displays a negative bias for summer when compared to GHCN-D and PRISM (≈ −19 %). Similarly, 3B42 presents a very good agreement with GHCN-D (−1.7 %) and PRISM (−2.4 %) and contrasts with the severe rainfall underestimation observed on the right side of the distribution during winter (Fig. 7a). The real-time 3B42RT displays moderate positive biases when compared to GHCN-D (+2.4 %) and PRISM (+0.4 %).

As can be seen, differences can vary greatly on an annual or seasonal basis. Overall, Stage IV exhibits a better agreement with surface observations than satellite QPEs, although important differences can be found on a RFC basis (Table 3). The differences observed over different RFCs can be a function of climatological characteristic of the domain selected such as precipitation types (stratiform, convective, frozen, snow) and/or precipitation regimes (cyclones, mesoscale convective systems, localized events). Other limitations and uncertainties include technical limitations induced by the topography (beam blocking effect) and the fact that estimates for each individual RFC are processed independently and use different algorithms that are ultimately combined

HESSD
11, 11489–11531, 2014

into a CONUS mosaic (Nelson et al., 2014). Further illustration of the importance of the bias-adjustment 3B42 can be found in Fig. 8. Figure 8 displays the scatterplots over the Missouri Basin River (MB) for winter (Fig. 8a) and summer (Fig. 8b). Once again, GHCN-D and PRISM present an average rain rate difference of about 3 % regardless of the season. For winter, there is a good agreement between Stage IV and GHCN-D (PRISM) with a slight underestimation of −5 % (−2 %) for Stage IV. The bias-adjusted 3B42 presents a moderate average rain rate overestimation with respect to surface observations (+16.7 %) when compared to GHCN-D computed over the 620 RGs selected (90 % time reporting during winter). For summer, Stage IV presents an average rain rate greater than surface observation of +8.4 % when compared to GHCN-D, which is comparable with the differences observed for the average rain rate retrieved from 3B42 (+7.7 %). For both the cold and the warm season, the improvement brought by the 3B42 bias-adjustment is clearly visible. During winter, the near real time 3B42RT exhibits a strong overestimation (+147 %) with a rain rate of 1.57 mm day^{-1} to be compared with the 0.64 mm day^{-1} measured by GHCN-D (Table 3). A closer look at the seasonal rainfall (Fig. 5d) shows that 3B42RT displays a significantly higher rainfall accumulation at higher latitudes and along the edges of the Missouri Basin River RFC when compared to the other datasets (PRISM, Stage IV, 3B42). These differences are due to cold season precipitation (snow, frozen precipitation). For summer, 3B42RT displays a severe overestimation (+61 %) when compared to surface observations. The overestimation of 3B42RT when compared to GHCN-D is found throughout the rain rate spectra. These strong differences are most certainly due to sub-cloud evaporation as mentioned earlier. Regardless of the season, the monthly bias-adjustment performed (3B42) corrects efficiently for both the cold and warm season rainfall accumulation overestimation with an adequate reduction of the average rain rate of about 110 and 50 % between monthly-adjusted and near real time precipitation for winter and summer respectively. Furthermore, while correction is found insufficient over mountainous domains such as over the Northwest (Fig. 7), the bias-adjustment performs very well over the great plains.

HESSD
11, 11489–11531, 2014

5 Daily precipitation, extremes, and conditional analysis

After investigating the ability of the different datasets to describe accurately precipitation patterns, this section investigates their ability to capture intense and extreme precipitation at the daily scale. To do so, different thresholds are selected to define those extremes (Fig. 9). Figure 9a displays the average number of rainy days by year derived from GHCN-D (first column), Stage IV (second column), 3B42 (third column), and 3B42RT (fourth column). For Stage IV, TMPA 3B42, and TMPA 3B42RT the daily accumulation is computed 12Z–12Z. For GHCN-D, the daily accumulation computed depends on the local time and is 07:00–07:00 LST for most of the locations, which correspond to 12Z–12Z on the Eastern US. Therefore, in addition to sensor characteristics, differing spatial and temporal resolution, differences could be explained by accumulation computed over a slightly different time period. Although the number of rainy days appears consistent in term of magnitude for the different observation platforms, there are noticeable differences over specific areas. Although the visual comparison between point data (GHCN-D) and gridded estimates (Stage IV) is delicate due to the scarcity of surface station coverage, both products present a very similar pattern and similar number of rainy days over CONUS. When compared to Stage IV, 3B42 displays a lower number of rainy days over the Northeast (NE), Middle-Atlantic (MA), and Ohio River Basin (OH). Similarly, a lower number of rainy days are observed for 3B42 over the Northwest (NW) when compared to Stage IV. On the other hand, 3B42 displays a higher number of rainy days over the Rockies encompassing part or all of the Missouri Basin River (MB), Colorado Basin River (CB), and California Nevada (CN) when compared to Stage IV. Different sensitivity for light rainfall detection thresholds for each sensor, the ability to retrieve snow/frozen precipitation, beam blockage over the Rockies, and/or the influence of temporal and spatial resolution, can explain the differences. For instance, Stage IV higher spatial resolution could improve the detection of localized events as compared to a coarser resolution. Overall, the rain gauge adjusted radar (Stage IV) and satellite (3B42) datasets display a satisfying

HESSD
11, 11489–11531, 2014

visual agreement over CONUS despite the local differences mentioned above. More important differences are observed with the real-time 3B42RT dataset. Differences between the rain gauge adjusted satellite dataset 3B42 and 3B42RT are particularly important over the Western US (Rocky Mountains) and at higher latitudes with more rainy days for 3B42RT. The seasonal breakdown presented in Fig. 6, shows that for the Rocky Mountains (CB, CN, MB), the differences between 3B42RT and 3B42 correspond to warm season rainfall (Fig. 6d). The differences observed at higher latitudes (MB, NC) are associated with cold season precipitation and are due to the challenge of retrieving snow and frozen precipitation from satellite measurements (Fig. 6c). A conditional analysis performed shows that differences are function of the daily threshold selected. For daily accumulation greater than the wet millimeter days (WMMD: $R > 17.8\,\mathrm{mm\,day^{-1}}$), significant differences are found over the Northwest (NW) and over the Southeastern US (LM, SE) (Fig. 9b). The Wet Millimeter Day threshold corresponds to the precipitation days that exceed the highest daily average over the area considered (Shepherd et al., 2007). For North America, the maximum daily average ($17.8\,\mathrm{mm\,day^{-1}}$) is recorded in Henderson Lake (British Columbia) (source NCDC). Depending on the RFC considered, WMMD events ranges from the 82 to 97 percentiles for LM and CB respectively (Nelson et al., 2014). Both Stage IV and 3B42 display similar distribution patterns of WMMD. The most important differences are found over the Northwest (NW). The differences between GHCN-D and Stage IV are due to the scarcity of station coverage over the Pacific Northwest Coast. For the gridded estimates, Stage IV displays a higher number of WMMD when compared to adjusted satellite estimates 3B42. The biggest differences are observed 3B42RT, which shows a much lower number of rainy days greater than WMMD. This is consistent with the underestimation observed for the daily averages for 3B42RT and to a lesser extent for 3B42 (Figs. 3d and 7a). The rain gauge adjustment performed for 3B42RT increases the number of days with accumulation greater than WMMD to level closer to Stage IV yet remaining below the Stage IV levels. Conversely, the Northeast (NE) that displayed less rainy days for 3B42 and 3B42RT than for Stage IV (Fig. 9a), presents

HESSD
11, 11489–11531, 2014

a higher occurrence of WMMD when compared to Stage IV (Fig. 9b). Similarly, higher counts of WMMD are found over the upper part of the North Central domain (NC) and over the Lower Mississippi (LM). With increasing threshold, the extent of extreme precipitation greater than two inches per day (Karl and Plummer, 1995) i.e. above 50.8 mm day^{-1} (EPD2) decreases and is localized along the Pacific coast (NW, CN), East of 105° W, along with isolated locations over the Rockies and the Southwest (Fig. 9c). Over the Eastern US where rain gauge coverage is denser, GHCN-D and Stage IV display comparable counts for EPD2. Over the Northeast (NE) and the Southeastern US (LM, SE), 3B42 and 3B42RT display a higher number of rainfall above 2 in day^{-1} as compared to Stage IV. Again, the bias-adjusted 3B42 present a higher number of days above 2 in than the real-time 3B42RT but less than for Stage IV. For intense precipitation greater than 4 in day^{-1} (Barlow, 2011) i.e. above 101.6 mm day^{-1} (EPD4), the area is limited to the Pacific Coast and East of 100° W, a domain regularly impacted by tropical cyclones (Prat and Nelson, 2013a, b, 2014a). While the bias-adjusted 3B42 and real-time 3B42RT display a comparable number of EPD4 events over the Southeastern US with a higher occurrence over the LM, 3B42RT display almost no days with accumulation greater than 4 in day^{-1} over the Northwestern US (NW). Although these EPD4 events are relatively infrequent (3 counts or less by year) and roughly correspond to the 0.1–0.5 % top daily events regardless of the RFC considered. Over the Northwest, results indicate a systematic underestimation of intense and extreme precipitation by 3B42RT. The bias-adjustment procedure (3B42) is able to better capture those extreme daily accumulation events (EPD4) that seem to remain underestimated when compared to Stage IV.

More quantitative information can be found in Fig. 10 that displays the proportion of rain gauges (GHCN-D) and the corresponding radar (Stage IV) or satellite (3B42, 3B42RT) pixel experiencing the different daily accumulation thresholds (WMMD, EPD2, EPD4) over the 11 year period. For CONUS (central panel), we note that this proportion of stations/pixels experiencing WMMD, EPD2, and EPD4 are comparable regardless of the platform considered. For instance, all stations/pixels experience WMMD during

HESSD
11, 11489–11531, 2014

the 11 year period (100 %). For EPD2, the ratio remains relatively close regardless of the sensor and varies from 88 % (Stage IV) to 95 % (3B42RT). Similarly for EPD4, the proportion is about 60 % (GHCN, 3B42, 3B42RT) with a slightly lower ratio (54 %) for Stage IV. A few interesting facts can be derived from the isolated RFC figures (border figures). A few RFCs (AB, LM, MA, SE, WG) display comparable ratios regardless of the sensor and the daily accumulation considered. Apart from a couple of RFCs (NC, OH), the ratio of Stage IV pixels experiencing extreme precipitation (EPD4) is relatively close to the ratio of GHCN. When looking for satellite pixels, we observe a relative symmetry for the ratio of stations experiencing EPD2. However, for Western (CN, NW) and Northeast (NE) RFCs we note a strong asymmetry in the ratio of 3B42RT pixels experiencing extreme precipitation (EPD4) when compared to the other sensors. This confirms the fact that over the Western US, the non-adjusted satellite QPE severely underestimates extreme daily precipitation. Interestingly, over the neighboring Colorado Basin River RFC (CB) we note a higher proportion of 3B42RT pixels displaying EPD2 and EPD4 than observed for the other sensors (GHCN, Stage IV, 3B42). Furthermore, regardless of the RFC and daily accumulation considered, the ratio of pixels for 3B42 is very close to that of the GHCN stations, hence providing confidence in the bias-adjustment performed.

Figure 11 provides a count of the total number of rainy days (Fig. 11a), WMMD (Fig. 11b), EPD2 (Fig. 11c), and EPD4 (Fig. 11d) for GHCN-D, Stage IV, 3B42, and 3B42RT over CONUS and for each RFC at the rain gauge location. For the number of rainy days, we note that 3B42 and 3B42RT provide comparable results and display less variability across the RFCs when compared to GHCN-D and Stage IV. For GHCN and Stage IV, the RFCs over the Rockies or located partially West of 95° W (AB, CB, CN, WG) display about half of the rainy days than the Eastern (MA, NC, NE, OH, SE) and Northwest (NE) RFCs. The Colorado Basin (CB) RFC presents consistently the lowest average number of events by active stations regardless of the daily accumulation. On the other hand, the Lower Mississippi (LM) present the highest average number of events regardless of the sensor considered. For selected RFCs, the

HESSD
11, 11489–11531, 2014

differences between 3B42 and 3B42RT are particularly important for EPD2 and EPD4. The biggest differences are found for Northwest (NW) and Missouri Basin River (MB). For the Missouri Basin River, the number of EPD2 and EPD4 events for 3B42RT is about 50 and 130 % higher respectively than for 3B42 and is attributed to summertime convection and sub-cloud evaporation over the Midwest. Over NW, the number of EPD2 and EPD4 events retrieved after bias-adjustment (3B42) is 6- and 3-fold the number of events indicated by 3B42RT due to the difficulty of capturing extreme precipitation in real-time over the area. For the latest case, consider that those events (EPD4) correspond to only a handful of occurrences for the period 2002–2012, and that caution should be made when analyzing those results.

The previous results were provided for the entire period 2002–2012. Figure 12 displays a contingency analysis between daily precipitation from the surface stations (GHCN-D) and the corresponding radar pixel (Stage IV). The number of rainy days ($R > 0\,\mathrm{mm\,day^{-1}}$) observed simultaneously at the rain gauge and the radar pixel is 62 % over CONUS (Fig. 12a). Significant differences are observed between RFCs and vary from 49 % (CB) to 71 % (OH). Events observed only by the radar are 24 % over CONUS, which is higher than the ratio for gauge only events (14 %). A similar trend is observed regardless of the RFC considered ranging from 18 % (NE) to 32 % (MB) for radar only events and from 8 % (AB) to 22 % (CN) for rain gauge only events. With increasing rain rate the event observed simultaneously by the gauge and the radar decrease from 62 % ($R > 0\,\mathrm{mm\,day^{-1}}$: Fig. 2a), to 56 % (WMMD: Fig. 12b), to 43 % (EPD2: Fig. 12c), and to 35 % (EPD4: Fig. 12d). Furthermore, while the ratio of events observed at the radar pixel only remain relatively close around 20 % regardless of the daily rainfall threshold (i.e. between 17 % for WMMD and 24 % for $R > 0\,\mathrm{mm\,day^{-1}}$), the number of events missed by the radar increases from 14 % ($R > 0\,\mathrm{mm\,day^{-1}}$) to 45 % (EPD4). In addition for accumulation greater than $2\,\mathrm{in\,day^{-1}}$, the number of extreme events missed by one or the other sensor is greater than the number of events observed simultaneously by both sensors (Fig. 12c). For accumulation greater than $4\,\mathrm{in\,day^{-1}}$,

HESSD
11, 11489–11531, 2014

the proportion of events missed by the radar becomes more important except for the Arkansas-Red Basin (AB) RFC.

Figure 13 displays the contingency analysis at each station location. Regardless of the daily accumulation ($R > 0\,\mathrm{mm\,day^{-1}}$: Fig. 13a), the Eastern US and West Coast stations present a higher proportion of events observed simultaneously at the rain gauge and radar pixel (median column). There is a strong contrast between the Eastern and Western US with the Eastern US displaying a lower proportion of events observed at the gauge only. With increasing daily accumulation (Fig. 13b–d), the number of rainfall events decreases over the Rockies, the Western and Northern US as described previously (Fig. 8). While the spatial extent of intense precipitation events becomes more and more limited to the Southeastern US with increasing daily accumulation, the ratio of events observed at the radar only (right column) remains relatively constant. For concurrent rainfall events (median column) the ratio decreases significantly for daily accumulation greater than $2\,\mathrm{in\,day^{-1}}$ (Fig. 13c). With increasing daily accumulation, as the number of events become smaller and spatially more localized, the ratio of events missed by the radar increases importantly over Midwest (Fig. 13b–d). However, caution is advised when looking at increasing threshold events in particular over areas where those events become more and more scarce. A closer look shows that most of the events observed at one of the other sensor (NY: left column; YN: right column) for accumulation greater than EPD2 and EPD4 are single occurrence events. For EPD2, the single occurrence events are located West of the -103° W longitude (Fig. 13e). For EPD4, apart from isolated events over the Rockies and the Pacific coast, most of single occurrence events are located at the edge of the Southeastern US, i.e. East of -100° W and North of the 40° N latitude (Fig. 13f).

6 Summary and conclusion

We compared quantitative precipitation estimates from satellite (3B42 and 3B42RT) and radar (Stage IV) with surface observations (GHCN-D) and models (PRISM) over

HESSD
11, 11489–11531, 2014

CONUS for the period 2002–2012. The comparisons were performed at the annual, seasonal, and daily scales over the major river basins within CONUS. The main conclusions of this study are summarized below.

- Over CONUS the different datasets show a satisfying agreement on an annual basis with differences ranging between −6.4 % (St. IV) and +6.1 % (3B42RT). At the RFC level, PRISM displays a difference of ±4 % with GHCN-D. Stage IV presents a tendency to underestimate when compared to surface observations (−14 to +1 %). A bigger spread of the differences is found between 3B42 and GHCN-D (−28 to +9 %). Finally, 3B42RT displays the bigger differences with GHCN-D (−33 to +49 %).

- The bias-adjusted 3B42 represents a significant improvement when compared to the near real time 3B42RT. The 3B42RT biases were particularly important at the seasonal scale over Northwest and the Western US (CN, NW) and at higher latitude over the Midwest (MB, NC) during winter with an important underestimation (−35 %) of the daily accumulation in the first case (CN, NW), and a severe overestimation (+100 %) in the second case (MB, NC). During summer, 3B42RT presents large positive biases (+45 %) over the Midwest (MB, NC). The bias adjustment (3B42) reduces those differences to moderate levels (i.e. from +100 to +22 % in winter and from +45 to +7 % in summer). Over the CNRFC, 3B42RT presents alternatively a severe underestimation for winter (−45 %) and a severe overestimation for summer (+121 %) with an overall annual differences of −23 % with surface observations GHCN-D.

- Despite the bias-adjustment, large biases remained for 3B42 at higher daily average accumulation ($> 5\,\mathrm{mm\,day^{-1}}$) over CONUS. Discrepancies can be explained by the difference between point (RG) and area (satellite, radar) measurements and the difficulty to capture localized, convective, and orographic events due to the coarser resolution. Furthermore, those differences can be more important at the seasonal scale and for selected basins in particular over the

HESSD
11, 11489–11531, 2014

Western part of CONUS (Pacific Northwest, Rocky Mountains) due to the difficulty to retrieve precipitation over mountainous areas.

– Stage IV presents an overall better agreement with surface observations than 3B42. At the seasonal level Stage IV displays the same tendency of rainfall underestimation with respect to surface observations with differences ranging from −18 to −2 % for winter and from −28 to +8 % for summer. Comparatively, 3B42 displays a bigger spread with no particular tendency (−38 to +25 %) for winter and a tendency of rainfall overestimation (−2 to +25 %) for summer.

– At the daily scale, the conditional analysis performed using increasing daily precipitation thresholds (0–4 in day^{-1}), showed that the sensor ability to capture intense and extreme precipitation depended on the domain considered. In particular, the near real time satellite QPE 3B42RT displayed poor performance in capturing intense daily precipitation over the Northwest. The bias-adjusted 3B42 exhibited a significant improvement and level closer to surface station (GHCN-D) and radar statistics (Stage IV) over the 11 year period.

– A contingency analysis performed at the rain gauge location and the corresponding radar pixel, showed that with increasing daily accumulation from greater than 0 to greater than 4 in day^{-1}, the ratio of events observed simultaneously by the gauge and the radar decreased from 62 to 45 %. Furthermore, while the ratio of events observed only by the radar remained close (around 20 %) regardless of the daily accumulation, the number of events measured at the ground but missed by the radar increased from 15 to 45 %. Although caution is required due to the fact that large rainfall events above 2 in day^{-1} (a fortiori events greater than 4 in day^{-1}) are infrequent and geographically limited to the pacific Northwest and Eastern US, results illustrate the challenge of retrieving extreme precipitation (top 1 % percentile) from remote sensing.

HESSD
11, 11489–11531, 2014

A similar study is currently underway for the different satellite QPE products available (CMORPH, PERSIANN). This work is included in a broader effort to evaluate long-term multi-sensor QPEs in the perspective of developing Climate Data Records (CDRs) for precipitation.

Acknowledgements. This research was supported by the NOAA/NCDC Climate Data Records and Science Stewardship Program through the Cooperative Institute for Climate and Satellites – North Carolina under the agreement NA09NES4400006.

References

Ashouri, H, Hsu, K., Sorooshian, S., Braithwaite, D., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: daily precipitation climate data record from multi-satellite observations for hydrological and climate studies, B. Am. Meteorol. Soc., doi:10.1175/BAMS-D-13-00068.1, in press, 2014.

Barlow, M.: Influence of hurricane-related activity on North American extreme precipitation, Geophys. Res. Lett., 38, L04705, doi:10.1029/2010GL046258, 2011.

Chen, M., Xie, P., Janowiak, J. E., and Arkin, P. A.: Global land precipitation: a 50-yr monthly analysis based on gauge observations, J. Hydrometeorol., 3, 249–266, 2002.

Daly, C., Neilson, R. P., and Phillips, D. L.: A statistical-topographic model for mapping climatological precipitation over mountainous terrain, J. Appl. Meteorol., 33, 140–158, 1994.

Huffman, G. J. and Bolvin, D. T.: TRMM and Other Data Precipitation Data Set Documentation, Lab. for Atmos., NASA Goddard Space Flight Cent. and Sci. Syst. and Appl. Inc., available at: ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf (last access: 7 September 2013), 2013.

Huffman, G. J., Adler, R. F., Morrissey, M., Bolvin, D., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global precipitation at one-degree daily resolution from multi-satellite observations, J. Hydrometeorol., 2, 36–50, 2001.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K., Hong, Y., Stocker, E. F., and Wolf, D. B.: The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeorol., 8, 38–55, 2007.

HESSD
11, 11489–11531, 2014

Joyce, R., Janowiak, J., Arkin, P., and Xie, P.: CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, J. Hydrometeorol., 5, 487–503, 2004.

Karl, T. R. and Plummer, N.: Trends in high-frequency climate variability in the twentieth century, Nature, 377, 217–220, 1995.

Kidd, C., Levizzani, V., and Laviola, S., S.: Quantitative precipitation estimation from earth observation satellites. Rainfall: state of the science, Geoph. Monog. Ser., doi:10.1029/2009GM000920, online first, 2010.

Kidd, C., Bauer, P., Turk, J., Huffman, G. J., Joyce, R., Hsu, K.-L., and Braithwaite, D.: Intercomparison of high-resolution precipitation products over northwest Europe, J. Hydrometeorol., 13, 67–83, 2012.

Lin, Y. and Mitchell, K. E.: The NCEP stage II/IV Hourly Precipitation Analyses: Development and Applications, Preprints, 19th Conf. on Hydrology, San Diego, CA, Amer. Meteor. Soc., 9–13 January, Paper 1.2, 2005.

Liu, C. and Zipser, E. J.: Diurnal cycles of precipitation, clouds, and lightning in the tropics from 9 years of TRMM observations, Geophys. Res. Lett., 35, L04819, doi:10.1029/2007GL032437, 2008.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An overview of the global historical climatology network-daily database, J. Atmos. Ocean. Tech., 29, 897–910, 2012.

Michaelides, S., Levizzani, V., Anagnostou, E., Bauer, P., Kasparis, T., and Lane, J. E.: Precipitation: measurement, remote sensing, climatology and modeling, Atmos. Res., 94, 512–533, 2009.

Nelson, B. R., Kim, D., and Seo, D. J.: Multisensor precipitation reanalysis, J. Hydrometeorol., 11, 666–682, 2010.

Nelson, B. R., Prat, O. P., Seo, D. J., and Habib, E.: Assessment and implications of NCEP Stage IV quantitative precipitation estimates, Weather Forecast, in review, 2014.

Nesbitt, S. W. and Anders, A. M.: Very high-resolution precipitation climatologies from the Tropical Rainfall Measuring Mission precipitation radar, Geophys. Res. Lett., 36, L15815, doi:10.1029/2009GL038026, 2009.

Nesbitt, S. W. and Zipser, E. D.: The diurnal cycle of rainfall and convective intensity according to three years of TRMM measurements, J. Climate, 16, 1456–1475, 2003.

HESSD
11, 11489–11531, 2014

Prat, O. P. and Barros, A. P.: Assessing satellite-based precipitation estimates in the Southern Appalachian mountains using rain gauges and TRMM PR, Adv. Geosci., 25, 143–153, doi:10.5194/adgeo-25-143-2010, 2010.

Prat, O. P. and Nelson, B. R.: Precipitation contribution of tropical cyclones in the Southeastern United States from 1998 to 2009 using TRMM precipitation data, J. Climate, 26, 1047–1062, 2013a.

Prat, O. P. and Nelson, B. R.: Mapping the world's tropical cyclone rainfall contribution over land using the TRMM Multi-satellite precipitation analysis, Water Resour. Res., 49, 7236–7254, doi:10.1002/wrcr.20527, 2013b.

Prat, O. P. and Nelson, B. R.: Characteristics of annual, seasonal, and diurnal precipitation in the Southeastern United States derived from long-term remotely sensed data, Atmos. Res., 144, 4–20, 2014.

PRISM Technical Note: Descriptions of PRISM Spatial Climate Datasets for the Conterminous United States, Tech. Note, 14 pp., available at: http://www.prism.oregonstate.edu/, last access: 1 October 2014, 2013.

Reed, S. and Maidment, D.: A GIS Procedure for Merging NEXRAD Precipitation Data and Digital Elevation Models to Determine Rainfall–Runoff Modeling Parameters, Online Report 95-3, Center for Research in Water Resources (CRWR), University of Texas at Austin, 1995.

Reed, S. and Maidment, D.: Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling, J. Hydrol. Eng., 4, 174–182, 1999.

Sahany, S., Venugopal, V., Nanjundiah, R.: Diurnal-scale signatures of monsoon rainfall over the Indian region from TRMM satellite observations, J. Geophys. Res., 115, D02103, doi:10.1029/2009JD012644, 2010.

Sapiano, M. R. P. and Arkin, P. A.: An intercomparison and validation of high-resolution satellite precipitation estimates with 3 hourly gauge data, J. Hydrometeorol., 10, 149–166, 2009.

Shepherd, J. M., Grundstein, A., and Mote, T. L.: Quantifying the contribution of tropical cyclones to extreme rainfall along the coastal southeastern United States, Geophys. Res. Lett., 34, L23810, doi:10.1029/2007GL031694, 2007.

Sorooshian, S., Hsu, K., Gao, X., Gupta, H., Imam, B., and Braithwaite, D.: Evolution of the PERSIANN system satellite-based estimates of tropical rainfall, B. Am. Meteorol. Soc., 81, 2035–2046, 2000.

HESSD
11, 11489–11531, 2014

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., and de Castro, M.: Global precipitation measurement: methods, datasets and applications, Atmos. Res., 104–105, 70–97, 2012.

Zhang, J., Howard, K., Langston, C., Vasiloff, S., Arthur, A., Van Cooten, S., Kelleher, K., Kitzmiller, D., Ding, F., Seo, D. J., Wells, E., and Dempsey, C.: National mosaic and multi-sensor QPE (NMQ) system, B. Am. Meteorol. Soc., 92, 1321–1338, 2011.

HESSD
11, 11489–11531, 2014

Table 1. List of the 12 NWS RFCs and corresponding number of GHCN-D rain gauges.

Nbr	ID	Name	Number of COOP RG	
			Unconditional (Reporting 90 % of time)	Unconditional (Total operational)
0	CONUS	CONUS	4075	8815
1	ABRFC	Arkansas Red Basin	325	637
2	CBRFC	Colorado Basin River	243	521
3	CNRFC	California Nevada	202	537
4	LMRFC	Lower Mississippi	363	728
5	MARFC	Middle Atlantic	166	378
6	MBRFC	Missouri Basin River	654	1385
7	NCRFC	North Central	541	1225
8	NERFC	Northeast	195	462
9	NWRFC	Northwest	260	570
10	OHRFC	Ohio River Basin	378	833
11	SERFC	Southeast	387	781
12	WGRFC	West Gulf	361	758

HESSD
11, 11489–11531, 2014

Table 2. Average rain rate (mm day^{-1}), and comparisons with surface observations (GHCN-D) for annual precipitation estimated derived from PRISM, Stage IV, 3B42, and 3B42RT. The comparison [%|a|R^2] includes the differences (%), and the linear regression coefficients (a; R^2) over CONUS and each RFC. For each QPE datasets, the numbers in bold and italic-bold indicate the upper and lower limits when compared to GHCN-D.

ID	GHCN-D	PRISM	Stage IV	TMPA	
				3B42	3B42RT
CONUS	2.47	2.42 [−2.2|0.98|0.98]	2.31 [−6.4|0.93|0.93]	2.52 [+2.1|1.00|0.83]	2.62 [+6.1|0.99|0.36]
ABRFC	2.13	2.06 [−3.0|0.97|0.99]	**2.14 [+0.7|1.00|0.96]**	**2.27 [+ 6.9|1.06|0.96]**	2.40 [+12.5|1.12|0.94]
CBRFC	0.89	0.90 [+0.7|0.99|0.87]	0.77 [−13.0|0.89|0.86]	0.84 [−6.1|0.86|-0.02]	1.18 [+33.2|1.22|0.02]
CNRFC	1.46	1.46 [+0.3|1.00|0.98]	1.32 [−9.7|0.91|0.92]	***1.05 [−28.0|0.63|0.54]***	***0.98 [−32.2|0.52|−0.10]***
LMRFC	3.75	3.64 [−2.8|0.97|0.84]	3.48 [−7.3|0.93|0.54]	3.87 [+3.2|1.03|0.26]	3.90 [+3.9|1.0|-0.59]
MARFC	3.28	***3.17 [−3.4|0.97|0.89]***	3.12 [−5.0|0.95|0.63]	3.34 [+1.6|1.01|-0.21]	2.78 [−15.2|0.84|-1.49]
MBRFC	1.59	1.54 [−2.6|0.97|0.97]	1.59 [+0.4|1.00|0.87]	1.72 [+8.4|1.08|0.94]	**2.37 [+49.3|1.42|0.56]**
NCRFC	2.42	2.35 [−2.9|0.97|0.90]	2.19 [−9.3|0.91|0.65]	2.63 [+9.0|1.09|0.74]	2.95 [+21.9|1.21|0.17]
NERFC	3.44	3.38 [−1.8|0.98|0.89]	3.14 [−8.7|0.91|0.43]	3.43 [−0.1|0.99|-0.88]	2.73 [−20.7|0.78|-1.42]
NWRFC	2.28	**2.36 [+3.6|1.03|0.97]**	***1.96 [−13.9|0.87|0.93]***	1.80 [−20.8|0.65|0.43]	1.68 [−26.1|0.47|-3.58]
OHRFC	3.28	3.20 [−2.3|0.98|0.89]	3.09 [−5.8|0.94|0.45]	3.44 [+4.9|1.05|0.42]	3.35 [+2.2|1.02|0.13]
SERFC	3.58	3.47 [−3.1|0.97|0.84]	3.36 [−6.2|0.93|0.45]	3.68 [+2.7|1.02|-0.08]	3.56 [−0.7|0.98|-0.08]
WGRFC	2.04	1.98 [−3.0|0.97|0.98]	1.89 [−7.4|0.93|0.94]	2.11 [+3.2|1.03|0.96]	2.19 [+7.4|1.06|0.87]

HESSD
11, 11489–11531, 2014

Table 3. Average rain rate (mm day^{-1}) and differences [%] between GHCN-D and other annual precipitation estimates (PRISM, Stage IV, 3B42, 3B42RT) over CONUS and over each RFC for winter (DJF) and summer (JJA). For each QPE datasets and season, the numbers in bold and italic-bold indicate the upper and lower limits when compared to GHCN-D.

Season	ID	GHCN-D	PRISM	Stage IV	TMPA	
					3B42	3B42RT
DJF	CONUS	2.09	2.03 [−2.8]	1.89 [−9.5]	2.09 [+0.1]	2.31 [+10.7]
	ABRFC	1.17	1.14 [−2.4]	**1.15 [−1.5]**	1.28 [+9.2]	1.31 [+11.6]
	CBRFC	1.01	1.01 [−0.2]	0.86 [−14.8]	0.83 [−17.4]	1.32 [+30.4]
	CNRFC	3.09	3.06 [−1.2]	2.72 [−12.2]	***1.92 [−38.0]***	***1.57 [−49.3]***
	LMRFC	3.77	3.64 [−3.5]	3.37 [−10.6]	4.03 [+6.8]	4.12 [+9.1]
	MARFC	2.60	***2.44 [−5.9]***	2.54 [−2.2]	2.60 [+0.1]	2.32 [−10.4]
	MBRFC	0.64	0.62 [−2.5]	0.60 [−4.9]	**0.74 [+16.7]**	**1.57 [+146.8]**
	NCRFC	1.35	1.30 [−4.0]	1.22 [−10.3]	1.69 [+24.9]	2.44 [+80.0]
	NERFC	2.89	2.76 [−4.4]	2.67 [−7.6]	3.13 [+8.2]	2.59 [−10.3]
	NWRFC	3.54	**3.60 [+1.8]**	3.11 [−12.2]	2.44 [−31.1]	2.51 [−29.0]
	OHRFC	2.86	2.73 [−4.7]	2.73 [−4.7]	3.11 [+8.8]	3.31 [+15.6]
	SERFC	3.13	3.00 [−4.0]	2.78 [−11.2]	3.25 [+3.7]	3.00 [−4.3]
	WGRFC	1.51	1.51 [−0.8]	***1.25 [−18.0]***	1.51 [+0.7]	1.59 [+4.5]
JJA	CONUS	2.73	2.65 [−3.0]	2.64 [−3.3]	2.85 [+4.4]	3.32 [+21.5]
	ABRFC	2.80	2.72 [−2.9]	2.94 [+5.2]	3.00 [+7.3]	3.45 [+23.2]
	CBRFC	0.88	0.86 [−2.6]	0.77 [−13.3]	0.93 [+5.3]	1.24 [+40.0]
	CNRFC	0.17	0.17 [+0.6]	***0.12 [−27.6]***	**0.21 [+24.7]**	**0.33 [+91.8]**
	LMRFC	3.62	***3.48 [−3.8]***	3.46 [−4.4]	3.70 [+2.3]	3.92 [+8.2]
	MARFC	3.63	3.51 [−3.5]	3.38 [−6.9]	3.81 [+4.8]	3.71 [+2.1]
	MBRFC	2.37	2.30 [−3.0]	**2.56 [+8.4]**	2.55 [+7.7]	3.81 [+60.9]
	NCRFC	3.31	3.22 [−2.6]	3.02 [−8.7]	3.51 [+6.1]	4.33 [+31.0]
	NERFC	3.78	3.68 [−2.5]	3.39 [−10.2]	***3.71 [−1.9]***	***3.64 [−3.7]***
	NWRFC	0.82	**0.84 [+2.1]**	0.66 [−19.1]	0.81 [−1.7]	0.84 [+2.4]
	OHRFC	3.45	3.36 [−2.5]	3.20 [−7.2]	3.59 [+4.0]	3.90 [+13.1]
	SERFC	4.53	4.36 [−3.7]	4.41 [−2.6]	4.66 [+2.8]	5.00 [+10.3]
	WGRFC	2.39	2.30 [−3.7]	2.39 [−0.1]	2.48 [+3.8]	2.69 [+12.5]

HESSD
11, 11489–11531, 2014

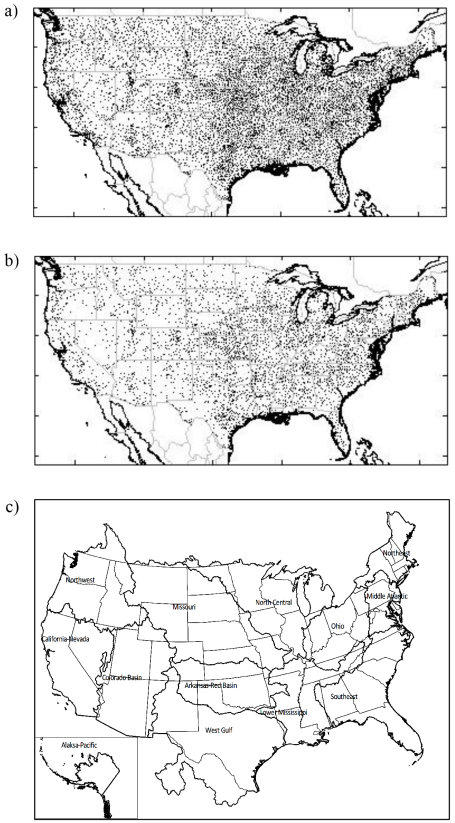

Figure 1. (a) Locations of the GHCN-Daily rain gauges locations over CONUS: **(a)** total 8815 rain gauges, and **(b)** the 4075 rain gauges reporting at least 90 % of the time during the period 2002–2012. **(c)** National Weather Service (NWS) 12 River Forecast Centers (RFCs).

HESSD
11, 11489–11531, 2014

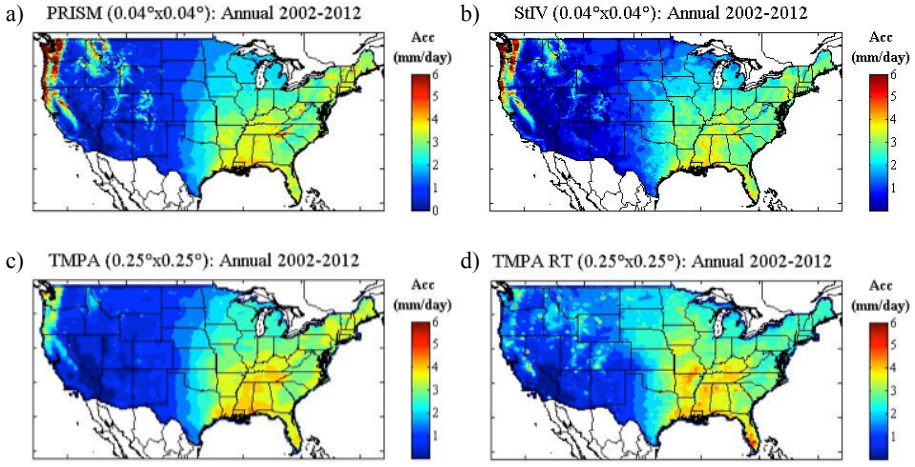

Figure 2. Annual average precipitation derived from: **(a)** PRISM, **(b)** Stage IV, **(c)** TMPA 3B42, and **(d)** TMPA 3B42RT for the period 2002–2012.

HESSD
11, 11489–11531, 2014

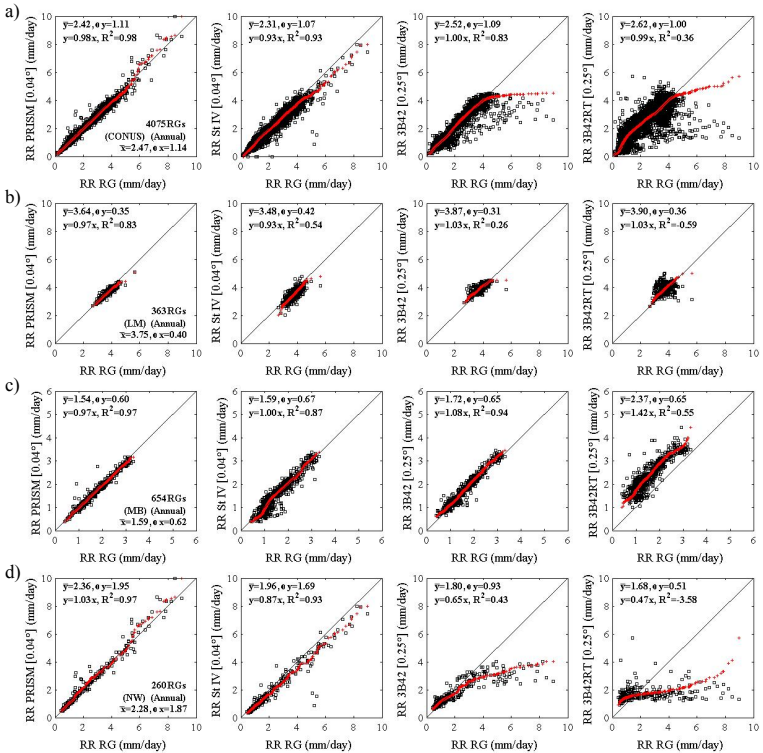

Figure 3. Scatterplots (black) and Q–Q (red) plots for annual precipitation derived from PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT when compared to GHCN-Daily network for: **(a)** CONUS, **(b)** Lower Mississippi River Basin (LM), **(c)** Missouri Basin River (MB), and **(d)** Northwest (NW) for the period 2002–2012. Please note that for row **c)**, the scale is different than for rows **a)**, **b)**, and **d)**.

HESSD
11, 11489–11531, 2014

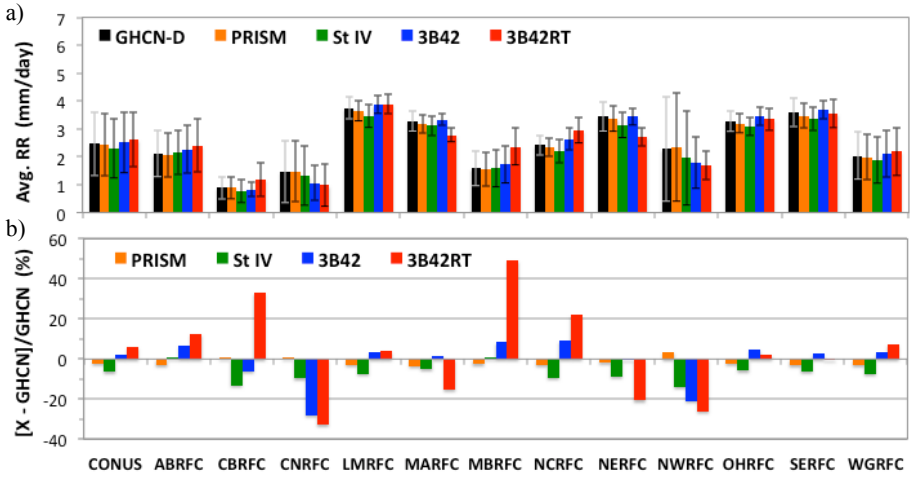

Figure 4. Average annual precipitation derived from GHCN-D, PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for the different RFCs. Differences (%) with respect to GHCN-D surface measurements.

HESSD
11, 11489–11531, 2014

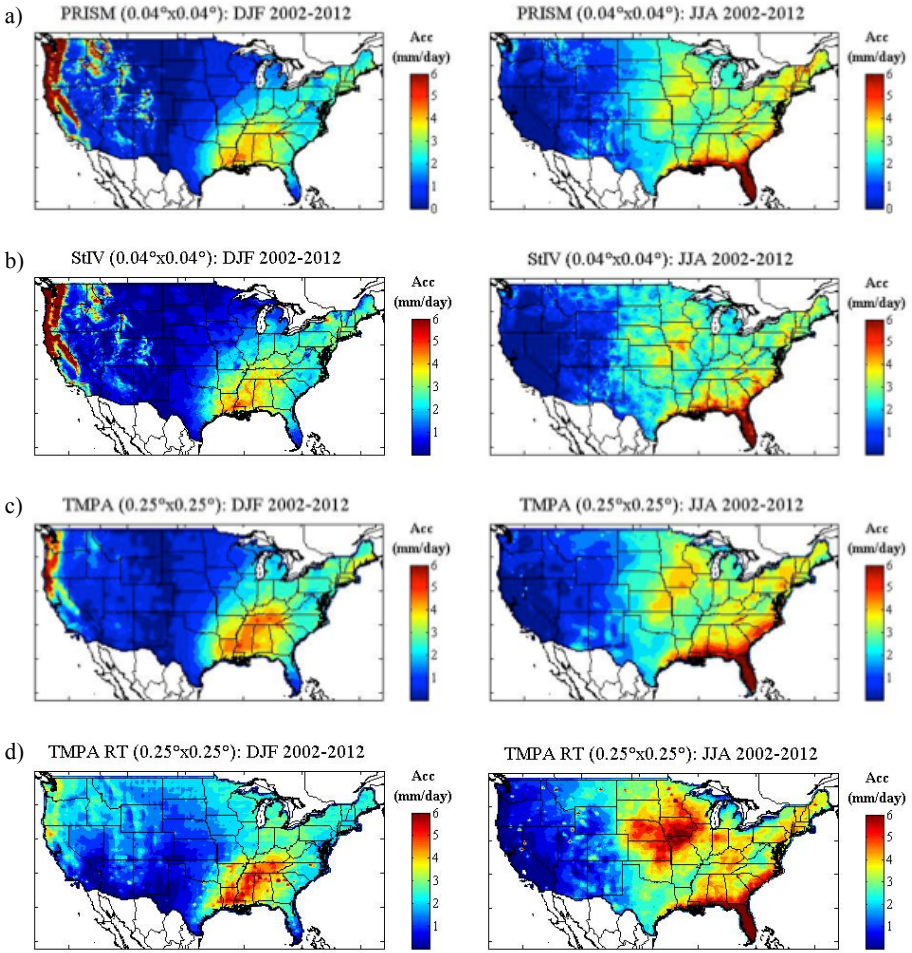

Figure 5. Winter (DJF: left column) and summer (JJA: right column) precipitation derived from: **(a)** PRISM, **(b)** Stage IV, **(c)** TMPA 3B42, and **(d)** TMPA 3B42RT for the period 2002–2012.

HESSD
11, 11489–11531, 2014

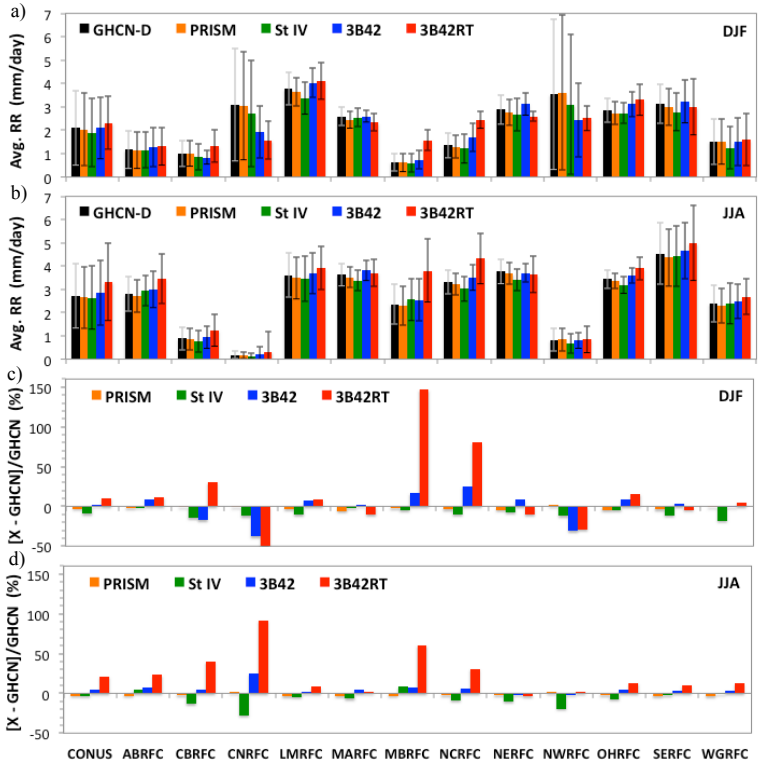

Figure 6. Seasonal rain-rate derived from GHCN-D, PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for each RFC and for: **(a)** winter (DJF), and **(b)** summer (JJA). Differences between GHCN-D and PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for: **(c)** winter (DJF), and **(d)** summer (JJA).

HESSD
11, 11489–11531, 2014

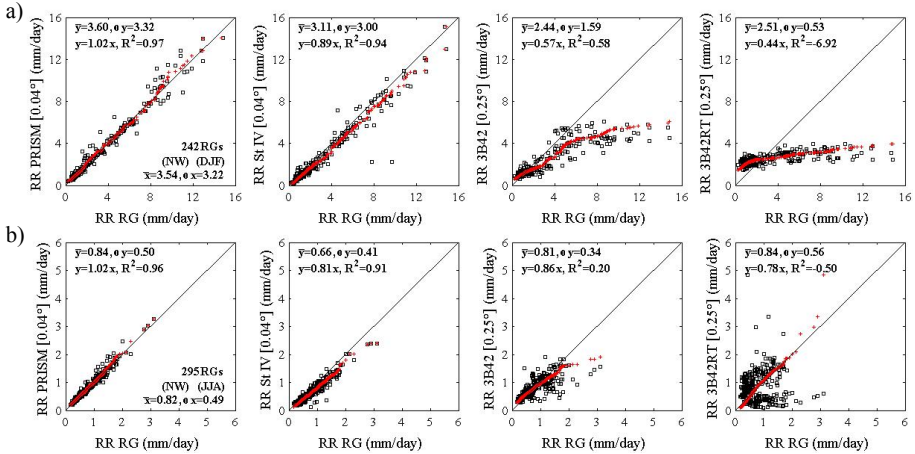

Figure 7. Scatterplots (black) and Q–Q plots (red) for the seasonal rain-rate for PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for: **(a)** winter (DJF), and **(b)** summer (JJA) over Northwest (NW).

HESSD
11, 11489–11531, 2014

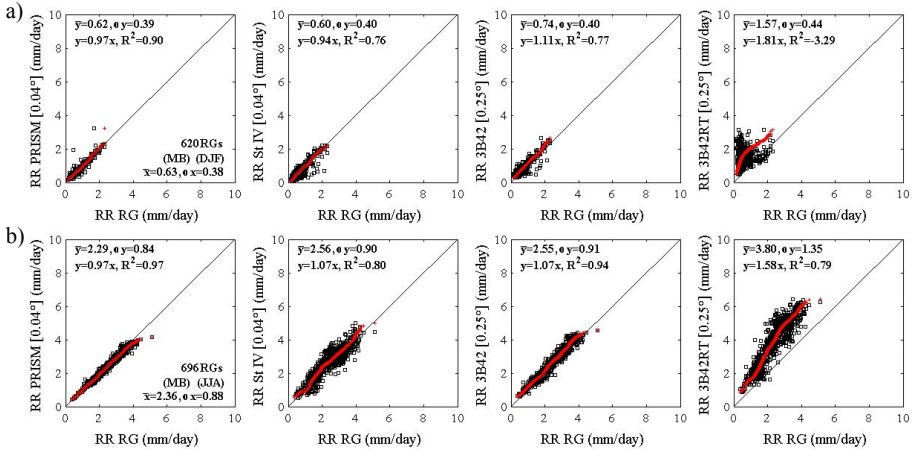

Figure 8. Scatterplots (black) and Q–Q plots (red) for the seasonal rain-rate for PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for: **(a)** winter (DJF), and **(b)** summer (JJA) over the Missouri Basin River (MB).

HESSD
11, 11489–11531, 2014

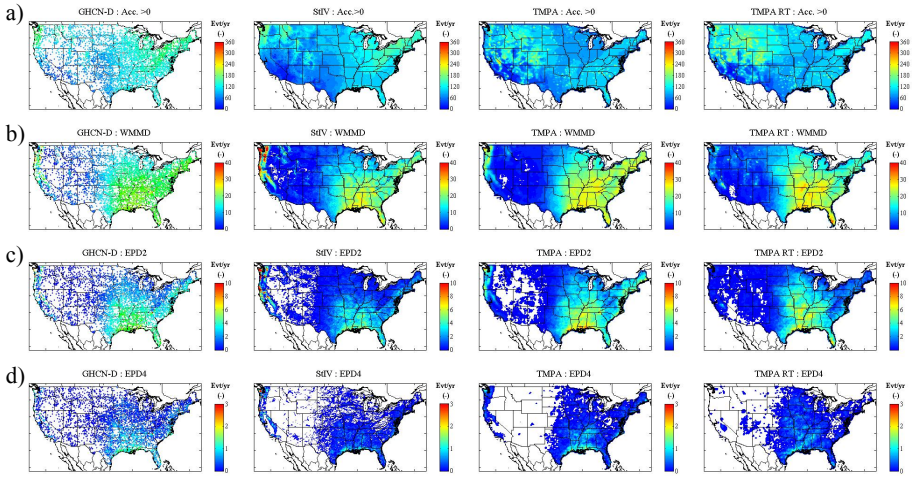

Figure 9. (a) Number of rainy days (RR > 0 mm day^{-1}), **(b)** wet millimeter days (RR > 17.8 mm day^{-1}: WMMD), **(c)** precipitation days with accumulation greater than 2 in day^{-1} (RR > 50.8 mm day^{-1}: EPD2), and **(d)** precipitation days with accumulation greater than 4 in day^{-1} (RR > 101.6 mm day^{-1}: EPD4) for GHCN-D (first column), Stage IV (second column), TMPA 3B42 (third column), and TMPA 3B42RT (fourth column).

HESSD
11, 11489–11531, 2014

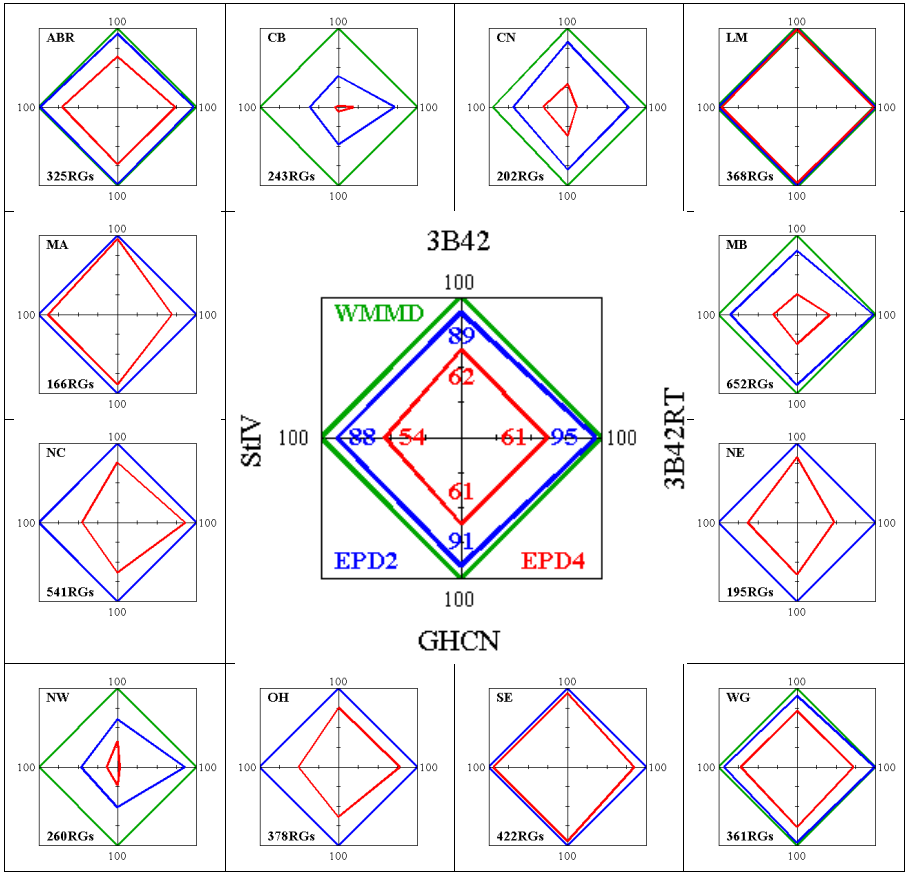

Figure 10. Proportion (%) of stations (GHCN-D) and corresponding pixel (Stage IV, TMPA 3B42, TMPA 3B42RT) experiencing WMMD, EPD2, and EPD4 over CONUS (central figure) and for the 12 River Forecast Centers (border figures).

HESSD
11, 11489–11531, 2014

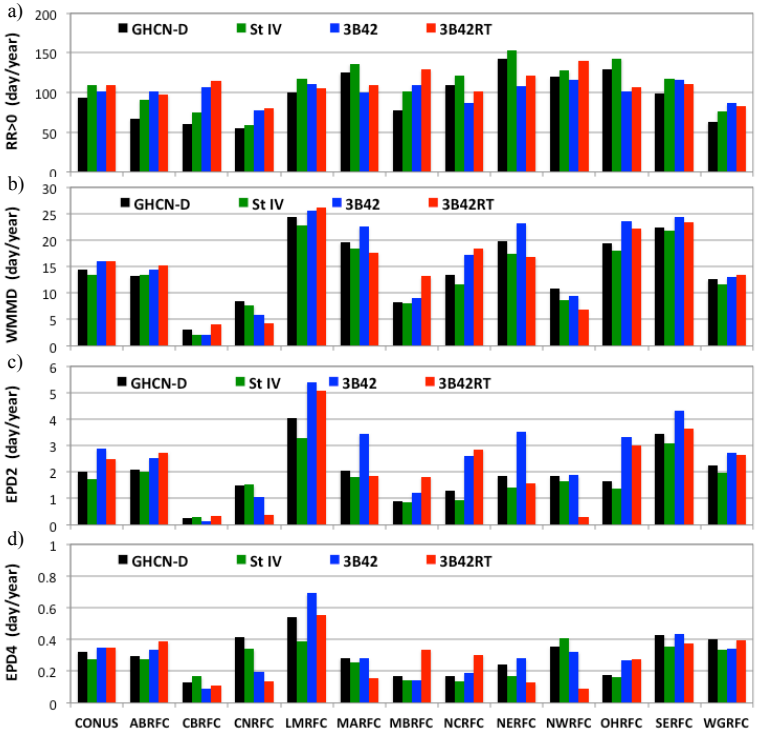

Figure 11. Average number of: **(a)** rainy days, **(b)** wet millimeter days (WMMD) i.e. precipitation days with accumulation greater than 17.8 mm day^{-1}, **(c)** precipitation days with accumulation greater than 2 in day^{-1} (EPD2), and **(d)** precipitation days with accumulation greater than 4 in day^{-1} (EPD4) for GHCN-D, Stage IV, TMPA 3B42, and TMPA 3B42RT over CONUS and for the 12 River Forecast Centers (RFCs). Data are for the period 2002–2012. The average number of days is normalized by the number of locations experiencing at least one event (Fig. 10).

HESSD
11, 11489–11531, 2014

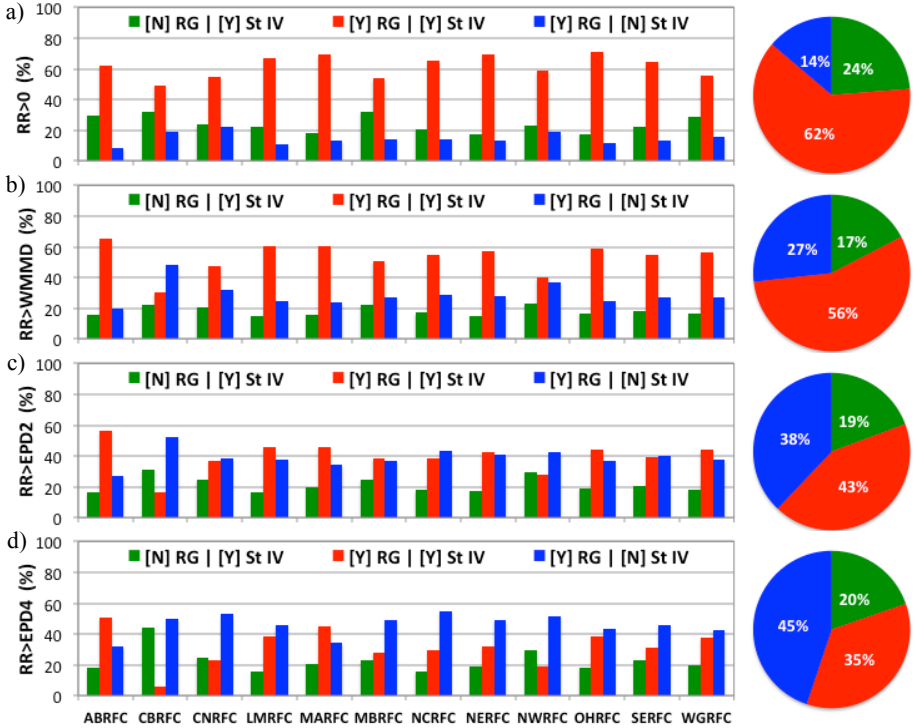

Figure 12. Contingency as a function of the daily threshold selected: **(a)** RR> 0, **(b)** RR >WMMD, **(c)** RR >EPD2, and **(d)** RR >EPD4 for rainfall observed simultaneously at the rain gauge and at the radar [YY: red], and successively at the rain gauge only [YN: blue], or at the radar only [NY: green] over CONUS (circle) and for the 12 RFCs (bars). Data are for the period 2002–2012.

HESSD
11, 11489–11531, 2014

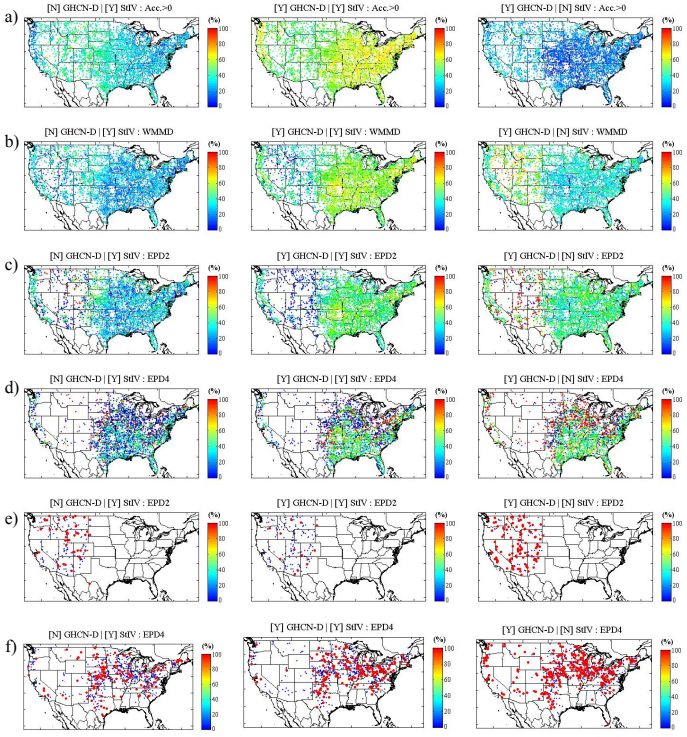

Figure 13. Contingency analysis at the rain gauge site with respect to the daily rainfall accumulation: **(a)** RR > 0, **(b)** RR > WMMD, **(c)** RR > EPD2, and **(d)** RR > EPD4 for rain observed at the radar pixel only (first column), simultaneously at the rain gauge and radar (second column), and at the rain gauge only (third column). **(e)** and **(f)** same as **(c)** and **(d)** but only displaying single event occurrence over the period 2002–2012.

HESSD
11, 11489–11531, 2014