Authors' answers to Reviewer's #1 comments

We would like to thank the Reviewer for the comments and suggestions that helped improving the quality of the manuscript. Below is an item-by-item reply to the Reviewer's comments and suggestions. The Reviewer's comments are reported in italic (1), the answers to the Reviewer are reported in blue (2), and the modifications to the manuscript are reported in green (3). The corrections will be inserted in the revised manuscript as indicated.

The authors evaluate relatively long (2002-2012) records of precipitation estimates from different sensors and devices (satellite, radar, rain gages). The focus is on the continental United States and analyses are performed at different temporal scales. The work and results seem solid and fit nicely within the line of research that has been pursued by the two authors. With that said, I have some major and minor issues with this study, as detailed below.

Major comments:

1. A big problem I have with this study has to do with its motivations. I found asking myself "why this study" quite a few times. If the focus is on a climatological analysis, then the satellite is not your best choice given that you have rain gages for a much longer period of time. The authors wrote, "this paper proposes to evaluate satellite precipitation estimates in the perspective of climate applications." Why? I would use 30+ years of rain gage measurements rather than 13 years of satellite estimates to accomplish that.

As indicated in the title, the primary goal of this paper is to evaluate different precipitations remotely sensed products like radar and satellite. This evaluation was performed over the period of record over which the different products are available. Stage IV is available from 2002 to present, while the Tropical Rainfall Measuring Mission (TRMM) suite of product is available starting in 1998.

If the longer-term perspective of this work is to use those remotely sensed estimates in the perspective of climatic application, a thorough evaluation of these products needs yet to be performed. The present study aims to serve as a benchmark in determining the difference and uncertainties between the different sensors suitability over CONUS for which we have extensive means of observation (in-situ, radar, satellites). According to our knowledge, there are no studies providing a comparison for in-situ, radar, and satellite measurements over an 11-year period.

To account for the Reviewer's comment, we modified the Introduction section and removed all source of confusion.

We will remove (p. 11490, ll. 22-23) *"Remotely sensed precipitation products are now coming of age where they can be considered for climatological applications.*"

We will replace (p. 11491, l. 23) *"Over the last 30 years …*" By: "*Over the last decades …*"

We will replace (p. 11491, l. 21) *"... and capture precipitation extremes in a climatological perspective.*" By: "*...and capture precipitation extremes over a multi-annual time frame.*"

We will remove (p. 11491, ll. 28-29) *"... and provide a long-term picture of the evolution of precipitation over time.*"

We will remove (p. 11492, ll. 11-12) *"Therefore, this paper proposes to evaluate satellite precipitation estimates in the perspective of climate applications.*"

We will modify (p. 11491, ll. 19-21) *"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."* By: "*The duration of the study will allow to assess systematic biases and capture year-to-year and seasonal variability.*"

2. I don't think that the United States is not the right place to perform a study of this kind because of the large rain gage network. I would also argue that these results cannot be generalized to other regions of the world for two main reasons: 1) different climatology, synoptic conditions and types of event; 2) lack of a dense rain gage network (compared to the United States) to bias-correct 3B42. A study of this kind would have made more sense to me if the focus had been global and/or if analyses had been performed at the sub-daily scale (even though I would still argue that in the United States there are rain gages providing data at the hourly/sub-hourly scale for a period of time longer than the satellite).

Indeed, the generalization of this study cannot be generalized to a global scale. A similar study would be necessary to assess those biases over different climatological zones. However, the goal of this study was to evaluate radar and satellite products with respect to in-situ observations over CONUS. We agree that a global evaluation of satellite products would be pertinent but this will be the object of a future communication. A long-term global analysis of the most common satellite QPE products including adjusted datasets (3B42, CMORPH-ADJ, PERSIANN-CDR) and their unadjusted counterparts (3B42, CMORPH, PERSIANN) is currently being conducted.

Regarding the precipitation analysis at the hourly/sub-daily scale. The satellite products used in this study have a three-hourly temporal resolution. While the Stage IV is available at the hourly, 6-hourly, and daily scales, the hourly Stage IV precipitation estimates should be used with caution especially when using them to compare to other data sets due to the automated quality control at the RFC level that cannot identify bad rain gauge reports (See Fig. 3 in Nelson et al. 2015). For more details about the products used, see the revised version of the manuscript.

Reference:

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, 2015.

3. There are a number of additional rain gage gridded rainfall products with high space-time resolution (e.g., see data by Ed Maurer at Santa Clara University).

As mentioned by the reviewer, there are other rain gauge gridded rainfall datasets available. In addition to the one named before, we could also mention the gridded estimates from GPCP (Global Precipitation Climatology Project), or GPCC (Global Precipitation Climatology Centre). In this work, we selected the GHCN-D dataset (more precisely the US-COOP subset of the GHCN-D dataset: See the answers to the second Reviewer for more details on the selection of the in-situ data). Regarding the dataset mentioned by the Reviewer, the dataset seems to be available for the period 1949-2010. Furthermore, it uses the same (or a portion) COOP stations used in this study (Maurer et al. 2002). Our goal in this study was to use directly the point stations for the direct comparison with the satellite and radar data. For consistency purpose, we used the gridded PRISM (Parameter-elevation Regressions on Independent Slopes) Model that incorporates GHCN-D in situ observations for annual and seasonal precipitation characteristics. By using point measurements (GHCN-D), (1) we avoid uncertainties and processing artifacts generated by the gridding procedure, and (2) we are able to select the rain gauges that were in operation over the entire period of record (we selected the rain gauges that reported at least 90% of the time over the period 2002-2012 (about 50% of the stations).

Reference:

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier D. P., and Nijssen, B.: A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States, *J. Climate* 15, 3237-3251, 2002.

4. I would assume that the vertical bars in Figures 4 and 6 are standard deviations. If that's the case, I doubt that the differences in mean are statistically significant. Please test this formally.

The vertical bars indicate indeed the standard deviations. We tested the statistical significance of the differences between the different datasets (PRISM, Stage IV, 3B42, 3B42RT) when compared to GHCN-D. The table below reports the results of the significance test for the annual (Fig. 4: Table 2) and seasonal (Fig. 6: Table 3) at the 5% significance level (Y=significance at the 5% level).

ID	YEAR				DJF				JJA			
	PRISM	St.IV	3B42	3B42RT	PRISM	St.IV	3B42	3B42RT	PRISM	St.IV	3B42	3B42RT
CONUS	Y	Y	Y	Y	-	Y	-	Y	Y	Y	Y	Y
ABRFC	-	-	Y	Y	-	-	-	-	-	Y	Y	Y
CBRFC	-	Y	-	Y	-	Y	Y	Y	-	Y	-	Y
CNRFC	-	-	Y	Y	-	-	Y	Y	-	Y	-	Y
LMRFC	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	Y
MARFC	Y	Y	-	Y	Y	-	-	Y	Y	Y	Y	-
MBRFC	-	-	Y	Y	-	-	Y	Y	-	Y	Y	Y
NCRFC	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y
NERFC	-	Y	-	Y	Y	Y	Y	Y	-	Y	-	Y
NWRFC	-	Y	Y	Y	-	-	Y	Y	-	Y	-	-
OHRFC	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y
SERFC	Y	Y	Y	-	Y	Y	-	-	-	-	-	Y
WGRFC	-	Y	-	Y	-	Y	-	-	-	-	-	Y

In the revised version, we will add a sign '' next to the values that are statistically different. The legend for Tables 2 and 4 will be completed by: "The asterisk indicates that the datasets are* *statistically different at the 5% significance level with respect to the surface observations." The revised sections 3 and 4, will include the result of the statistical tests.*

We will add (p. 11498, l. 10):

The differences are statistically significant at the 5% significance for about half of the RFCs (5 over 12).

We will add (p. 11498, l. 11):

… and statistically significant differences for 9 of the RFCs.

5. Another element that is not discussed and that could affect the evaluation of the products is related to the fact part of the study region is outside of the orbit of TRMM (35N/S).

The products 3B42RT and 3B42 are available from 50S-50N and are here used over CONUS for the domain 24N-40N. The TRMM Multisatellite Precipitation Analysis (TMPA) uses input datasets from low earth orbit satellites (TMI: The Microwave Imager; SSM/I: Special Sensor Microwave Imager; AMSR-E: Advanced Micro-wave Scanning Radiometer-Earth Observing System; AMSU-B: Advanced Microwave Sounding Unit-B). Apart from TMI, which has a footprint covering the band 38S-38N, all the other sensors (SSM/I, AMSR-E, AMSU-B) have coverage beyond the band 50S-50N of TMPA. While the sensor differing coverage might have an impact of the quality of the precipitation estimates for latitude above 38N (for the band 38N-40N), this point is beyond the scope of the study. More detail on the algorithm 3B42RT and 3B42 can be found in Huffman et al. (2007) and Huffman and Bolvin (2013).

Reference:

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., 2013. [Available at: ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf, accessed 28 Jan. 2015].

We synthetized this point in the section 2.4 related to 3B42 and 3B42RT product description.

6. Section 5 requires some additional work. I don't think it makes sense to compare "extreme" rainfall at fine resolution (Stage IV or even worse rain gages) with respect to a _625 km^2 pixel. Analyses of this kind should have been performed by both regridding Stage IV and interpolating the rain gages, or by using thresholds associated with the rainfall distribution for each pixel/location and products (e.g., 95th or 99th percentile).

The goal of this study being to use the different products at their native resolution, we do not see the necessity of regridding the different datasets from point measurement (in-situ) or 4-km (Stage IV) to the 25-km resolution of TMPA 3B42. However, we agree that this point, also mentioned by the other Reviewer requires clarification. The pertinent question is whether the correlation distance (defined as the distance at which the pixel-to-pixel correlation decreases below 1/e≈0.37) is greater than the native resolution of the sensors that is 4-km for Stage IV and 25-km for TMPA.

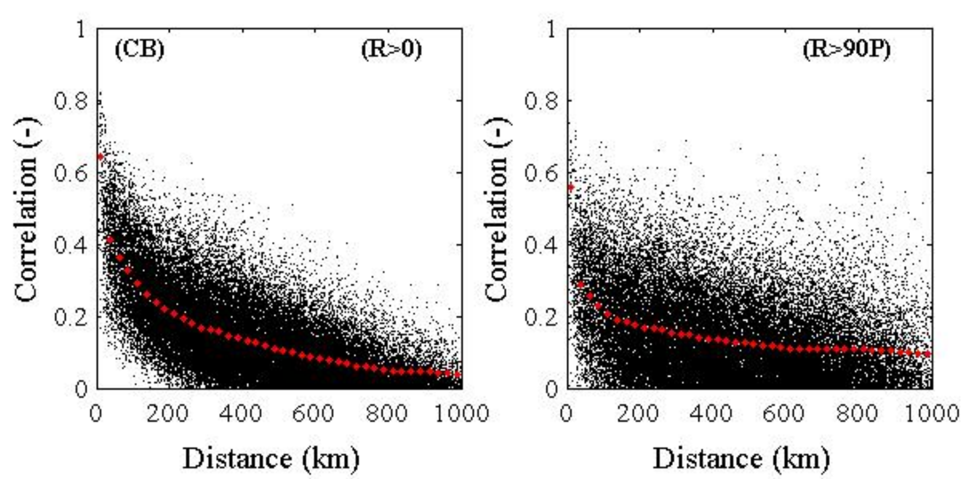
For unconditional rain rates, the correlation distance depends on the time scale considered. When it comes to extreme precipitation events (i.e. conditional analysis), this distance depends on the threshold considered (value or percentiles). The Table below shows the 90th percentiles of daily precipitation for the different RFCs at the seasonal (DJF, MAM, JJA, SON) and annual (YEA) scale derived from Stage IV (Nelson et al. 2015). As can be seen the there is an important variation between the seasons and the different RFCs. Roughly, the 90th percentile corresponds to the Wet Millimeter Day (WMMD: 18.8 mm/day).

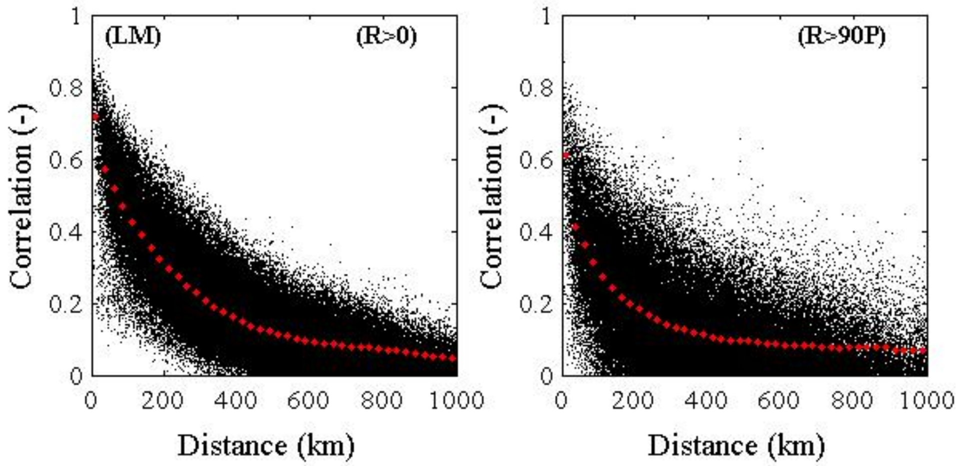
Percentile	Season	Rain rate (mm/day)											
		RFCs											
		NW	CN	CB	MB	AB	WG	NC	OH	LM	NE	MA	SE
90th	DJF	15.1	21.9	9.4	*6.4*	15.1	16.5	9.7	16.5	**28.6**	16.5	19.5	26.2
	MAM	11.3	13.5	*7.4*	14.1	23.0	23.4	16.5	21.4	**30.7**	17.1	18.8	28.5
	JJA	9.3	*6.3*	7.3	17.9	24.0	21.6	20.9	22.3	24.1	19.4	20.9	**26.0**
	SON	13.0	13.8	*10.1*	12.9	22.8	24.3	17.1	22.1	**30.5**	23.8	26.1	28.8
	YEA	12.5	15.9	***8.3***	13.9	22.0	21.8	16.8	**30.4**	**28.4**	19.1	21.2	27.0

**For each threshold and season, the max value is reported in bold and the min value is reported in italic. On average, the CBRFC (Colorado Basin) reports the lower values, while the LMRFC (Lower Mississippi) displays the higher values. The values reported in red are the values tested for the determination of the correlation distances for the 90th percentile.*

For each station with precipitation corresponding to different daily conditions (R>0, R>90th percentile), we computed the spatial correlation of daily precipitation events regardless of the daily values of the other stations for the CB and LM RFCs. The red dotted lines represent the average values obtained for 25-km intervals.

The correlations decrease rapidly with increasing distances between stations. The decrease is sharper for the CB RFC than the LM RFC and with increasing percentile of the daily value. For the 90th percentile, the average correlation distance (at 1/e) is about 30-80-km. Typically, the average correlation distance is greater than the representative footprint of each sensor: 4-km for Stage IV and 25-km for TMPA.





In the revised version, we synthetize those results. See also the answer to the other Reviewer. We added in the text related to Fig. 10 (p. 11507, l. 18)

However those results have to be interpreted with caution as they present a count of the daily events over the 11-year period. The number of events decreases with increasing rain rate and the WMMD correspond roughly to the 90th percentile precipitation events regardless of the RFC (Nelson et al. 2015). A test was performed to determine the interstation correlation of daily precipitation events corresponding to the 90th percentile (not shown). For each station, the correlation was computed using the daily events greater than the 90th percentile regardless of the values of the other stations. Results showed that for those high-intensity events, the average correlation distance was about 30-80-km which is comparable with the satellite footprint.

We also added (p. 11508, l. 13):

We will assume that the rain gauge is representative of the grid-averaged rainfall for Stage IV. The computation of the interstation correlation for daily events indicated that the correlation distance was greater than the 4-km spatial resolution of the radar (not shown).

Minor comments:

7. While the manuscript is generally well written, there are few typos here and there [e.g., pg. 11492, line 14 (remove parenthesis); pg. 11497, line 6 (stage IV); pg. 11499, line 15 (3B43RT)].

This will be corrected in the revised version.

8. Pg. 11505, line 9: rain gages have troubles measuring solid precipitation as well.

We added this point at the end of the revised section 4.

9. Pg. 11509, line 12: why "radar-only" if rain gages are used as well?

Here we refer to the radar pixel location that is defined as the closest radar pixel to a given rain gauge. This corresponds to a "false alarm" where precipitation is detected only at the radar pixel and nothing is detected by the rain gauge. For clarity purpose, we added the word "pixel" between "radar" and "only" as well as in the Figure 12 legend.

10. Please include confidence intervals in the qq-plots on Figures 3, 7, and 8.

The multi-paneled qq-plots and scatterplots are already delicate to decipher due to the number of points. Therefore it is doubtful that this would be of any interest to add the confidence interval on the figures. An option would be to replace the scatterplots by the confidence interval. However, we feel that the scatterplots provide a more pertinent illustration of the differences with respect to surface observations especially between the adjusted and un-adjusted versions of 3B42.

11. Figure 10: are the proportions different at the 5% level? Similar question for Figure 11.

We are not sure what the Reviewer means by that. The Figure 10 represents the number of stations or corresponding pixel that experience precipitation above a given condition (R>WMMD, EPD2, EPD4: 1 for yes; 0 for no) for each RFC.

12. Are the 3B42RT data rerun every time there was a change in the number of satellites providing data? If not, it is hard to make comparisons.

The 3B42RT product is the near real time satellite QPE and has a lag time of about 6-hrs. The product incorporates various satellite measurements. For a given time stamp (available at 00Z, 03Z, 06Z, 09Z, 12Z, 15Z, 18Z, 21Z), the precipitation estimate include the best available information that falls within the time window of +/- 1.5hr centered around the nominal time. The number of satellite providing data can be different for each time stamp and therefore differences in quality can be expected from one time stamp to another. However, this point is beyond the scope of this study. For more details we will refer to Huffman et al. (2007) and Huffman and Bolvin (2013).

The description of the satellite products 3B42 and 3B42RT that includes a discussion related to uncertainties was extended in the revised version. We added (p. 11495, l. 5):

The quality of the blended precipitation estimates depends on the number of satellite estimates available at a given time stamp and on the sensor characteristics.

Authors' answers to Reviewer's #2 comments

We would like to thank the Reviewer for the comments and suggestions that helped improving the quality of the manuscript. Below is an item-by-item reply to the Reviewer's comments and suggestions. The Reviewer's comments are reported in italic (1), the answers to the Reviewer are reported in blue (2), and the modifications to the manuscript are reported in green (3).

The aim of this paper is to assess the characteristics of a suite a ground-based (GHCN-Daily, PRISM and Stage IV) and space-borne (TMPA) QPE products over the CONUS over a 10-years period at various time scales. Rain gauge observations and PRISM are taken as a reference for the evaluation of remote sensed products at the annual, seasonal, and daily scales over River Forecast Centers (RFCs). It is shown that if all products present similar annual average accumulation over the CONUS, discrepancies appear at the RFC scale in particular with 3B42RT. The gauge correction in 3B42 often mitigates these biases compared to 3B42RT at the annual and seasonal scale except for the Western US. Extreme daily precipitation is shown to be challenging to retrieve for all remote sensing products.

The research area of using ground and satellite data to obtain good quality distributed rain estimations at the global scale recovers a real need for a number of applications ranging from climate analysis to prediction of floods.

The paper is clear and structured, but the first part is hard to follow with numerous statistical numbers and too few interpretation of the results. The novelty of the work needs to be better highlighted. The results are not always or poorly explained. Keys aspects in the comparison methodology are not well described. Recommendation is "major revision". I would recommend that the authors address carefully the points mentioned below for the manuscript to be ultimately accepted for publication.

1. Title: the authors should include the time scales they are considering (daily to annual). One important value of TMPA and Stage IV is that these datasets are available at sub-daily scale. This is not evaluated here.

We welcome the suggestion. The title will be changed into:

Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge datasets at daily to annual scales (2002-2012)

2. As a basis for this analysis, the GHCN-Daily gauges are is used as a reference to remote sensed precipitation products. It is probably not suitable since some of these products (3B42, Stage IV) ingest gauge correction over the CONUS and are not mutually independent. This is particularly highlighted with some extremely high correlation coefficients (e.g. p.11496 l. 24, p.11501 l.29). In this context what is the contribution of this analysis, and can it be better explicated?

At the locations specified by the Reviewer (p. 11496, l.24, p. 11501, l. 29), the correlations observed between GHCN-D and PRISM are indeed mentioned. For PRISM, this high correlation was "*expected due to the fact that PRISM gridded precipitation estimates incorporate GHCN-D stations*". Earlier in the text, we also noted that (p. 11494, l. 1-4) that "*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.*" and that (p. 11494, l. 4-6) "*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.*".

If we think, that for PRISM this point was sufficiently developed, for the other products that are being evaluated (3B42, Stage IV), we welcome the request for a better clarification.

For 3B42, the monthly gauge adjustment uses the precipitation gauge analysis from the Global Precipitation Climatology Centre (GPCC) (Huffman and Bolvin 2013). The gauge analysis used is the GPCC Monitoring Product at 1 deg grid resolution (Schneider et al., 2011). This specific analysis uses SYNOP (synoptic weather observation reports) and CLIMAT reports that are received near-real time from 7000-8000 automated stations worldwide. This number of 7000-8000 stations worldwide (GPCC Monitoring product) is to be compared with approximately over 20000 GHCN-D stations for CONUS alone (Menne et al. 2012). Furthermore, while GHCN-D incorporates surface observations from different sources (see Table 2 in Menne et al. 2012), we selected the subset from the US-Cooperative Observing network (US-COOP), which represented about 9000 stations. The US-COOP network includes first order stations (1600 manual and automatic synoptic stations) and stations from volunteer observers. We further reduced this number by selecting only the stations that reported at least 90% of the time during the period 2002-2012 (4075 stations). In that context, if it is possible that some of the first order automated synoptic stations included in GHCN-D are also used in the GPCC gauge analysis (SYNOP), most of the GHCN-D stations used for evaluation are not a part of the GPCC Monitoring Product used in the 3B42 adjustment (please note that GPCC proposes Full Data Reanalysis Products which might possibly incorporate more of the GHCN-D stations including the US-COOP stations from the volunteer observers).

For Stage IV, the in-situ data incorporation depends on each River Forecast Center (RFC). Briefly, Stage IV represents the final stage of the process that combines mosaicked estimates from the 12 RCFs (Lin and Mitchel 2005). Stage IV analysis are produced for hourly, and 6-hourly totals. The 6-hourly totals are combined into daily accumulation (12Z-12Z). The gauges used at the RFC level include available hourly rain gauges such as HADS (Hydrometeorological Automated Data System) gauges, ASOS (Automated Surface Observing System), and AWOS (Automated Airport Weather Stations) reports (Hou et al. 2014). Furthermore, some of the Western RFCs (California-Nevada: CNRFC, Northwestern: NWRFC) do not use radar estimates due to poor coverage over mountainous areas. The RFCs use an automated analysis of rain gauge observations (Mountain Mapper) that incorporates the gridded monthly precipitation climatology from PRISM (Parameter-elevation Regressions on Independent Slopes Model). Changes in gauge-adjustments procedures are always possible at the RFC level by using the best possible in-situ data observations. It is therefore possible that some GHCN-D stations could be incorporated in the bias-adjustment but it is not known for certain. What we do know is that the SERFC does not use any of the GHCN-D stations (Nelson et al. 2010). In that context, except PRISM that incorporates GHCN-D data including COOP stations, it is reasonable to assume that the HADS, ASOS, and AWOS observations used in the Stage IV bias-adjustment procedure are different from the GHCN-D stations used for the remotely sensed products evaluation.

In both cases (Stage IV, 3B42) it is virtually impossible to track down and identify the stations that are or aren't used in the bias-adjustment procedure. Therefore, a high correlation might or might not indicate that a particular or a group of station is included in that procedure. While we are fully aware, that this point might constitutes a limitation of the present study, there aren't many alternative at hands due to the lack of an alternate consistent long-term dense rain gauge network that could be used in the evaluation of the different precipitation products. However, being aware of this challenge, we performed an assessment of Stage IV using the US Climate Reference Network (USCRN). The USCRN consists in 114 surface stations for the contiguous US with about a decade (since 2002) worth of observations for the first installed stations (Diamond et al. 2013). This analysis is presented in Nelson et al. (2015). Results showed that for the unconditional rainfall rate (mm/day), comparison of Stage IV with USCRN in situ data displayed satisfying to strong statistic (R^2>0.85) for the different RFCs.

To summarize, the fraction of rain gauge intervening both in the evaluation (GHCN-D) and the adjustment procedure (3B42, Stage IV) appear limited and shouldn't prevent GHCN-D and the COOP subset used to be suitable to perform the radar/satellite products evaluation over the period 2002-2012.

To account for this point, we will add the following, in the datasets descriptions in section 2 (GHCN-D, Stage IV, 3B42). For GHCN-D (p. 11493, l. 10):

The GHCN-D datatset incorporates surface observations from different sources (see Table 2 in Menne et al. 2012). We selected the subset from the US-Cooperative Observing network (US-COOP), which represented about 9000 stations. The US-COOP network includes first order stations (1600 manual and automatic synoptic stations) and stations from volunteer observers.

For Stage IV (p. 11494, l. 17):

Stage IV represents the final stage of the process that combines mosaicked estimates from the 12 RCFs. The gauges used at the RFC level for bias-adjustment include available hourly rain gauges such as HADS (Hydrometeorological Automated Data System) gauges, ASOS (Automated Surface Observing System), and AWOS (Automated Airport Weather Stations) reports (Hou et al. 2014). Furthermore, some of the Western RFCs (Colorado Basin River: CBRFC, California-Nevada: CNRFC, Northwestern: NWRFC) do not use radar estimates due to poor coverage over mountainous areas. Those RFCs use an automated analysis of rain gauge observations (Mountain Mapper) that incorporates the gridded monthly precipitation climatology from PRISM. Although changes in gauge-adjustments procedures are always possible at the RFC level including the incorporation of the best available in-situ observations, it is reasonable to assume that the in-situ observations used in the Stage IV bias-adjustment procedure (HADS, ASOS, AWOS) are different from the US-COOP subset of GHCN-D used for the evaluation. For the Western RFCs however, the incorporation of the PRISM climatology that uses GHCN-D in-situ data will have to be kept in mind for any analysis.

For 3B42 and 3B42RT (p. 11495, l. 14):

The gauge analysis used is the GPCC Monitoring Product at 1 deg grid resolution (Schneider et al., 2010, 2011). This specific analysis uses SYNOP (synoptic weather observation reports) and CLIMAT reports that are received near-real time from 7000-8000 automated stations worldwide. While it is possible that some of the first order automated synoptic stations included in GHCN-D are also used in the GPCC gauge analysis (SYNOP), most of the US-COOP subset of the GHCN-D stations used for evaluation are not a part of the GPCC Monitoring Product used in the 3B42. Being virtually impossible to track down and identify the automated stations that are or aren't used in the bias-adjustment procedure for 3B42, we are confident that this number remaining relatively low will not compromise the independent assessment of the 3B42 dataset.

References:

Diamond, H. J., Karl, T. R., Palecki, M. A., Baker, B. C., Bell, J. E., Leeper, R. D., Easterling, D. R., Lawrimore, J. H., Meyers, T. P., Helfert, M. R., Goodge, G., and Thorne, P. W.: U.S. Climate Reference Network after one decade of operations: status and assessment. *Bulletin of the American Meteorological Society*, 94, 424-498, 2013.

Hou, D., Charles, M., Luo, Y., Toth, Z., Zhu, Y., Krzysztofowicz, R., Lin, Y., Xie, P., Seo, D.-J., Pena, M., and Cui, B.: Climatology-Calibrated Precipitation Analysis at Fine Scales: Statistical Adjustment of Stage IV toward CPC Gauge-Based Analysis. *Journal of Hydrometeorology*, 15, 2542-2557, 2014.

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., 2013. [Available at: ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf, accessed 28 Jan. 2015].

Nelson, B. R., Seo, D.-J., and Kim, D.: Multisensor Precipitation Reanalysis. *Journal of Hydrometeorology*, 11, 666-682, 2014.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., and Ziese, M.: GPCC Monitoring Product: Near Real-Time Monthly Land-Surface Precipitation from Rain-Gauges based on SYNOP and CLIMAT data. DOI: 10.5676/DWD_GPCC/MP_M_V4_100; http://dx.doi.org/10.5676/DWD_GPCC/MP_M_V4_100, 2011.

Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre, DWD, 12 pp., 2010. [Available at ftp://ftp-anon.dwd.de/pub/data/gpcc/PDF/GPCC_intro_products_2008.pdf, accessed 28 Jan. 2015]

3. "Seldom studies that deal with the long-term assessment of precipitation products (annual or multi-annual basis) are available in the scientific literature" (p. 11492 ll.9-11): could you please cite some studies? E.g. Chen, S., et al. (2013), Evaluation of the successive V6 and V7 TRMM multisatellite precipitation analysis over the Continental United States, Water Resour. Res., 49, 8174–8186, doi:10.1002/2012WR012795.

The reference will be added.

Chen, S., Hong, Y., Gourley, J. J., Huffman, G. J., Tian, Y., Cao, Q., Yong, B., Kirstetter, P.-E., Hu, J., Hardy, J., Li, Z., Khan, S. I., and Xue, X.: Evaluation of the successive V6 and V7

TRMM multisatellite precipitation analysis over the Continental United States, *Water Resour. Res.*, 49, 8174–8186, doi:10.1002/2012WR012795, 2013.

4. The impact of the resolution of these various products on the comparison results needs more discussion. The limitations of point rain gauge measurements for evaluation of area-averaged precipitation estimates have been documented for years by Ciach and Krajewski (1999), Ciach et al, (2003), Habib et al (2004) to cite a few. This issue is acknowledged in recent studies using gauges to evaluate satellite precipitation estimates (e.g. Kirstetter et al. 2013).

References:

Ciach, G.J., Habib, E., and Krajewski, W.F., 2003. Zero-covariance hypothesis in the error variance separation method of radar rainfall verification. Adv. Water Resour., 26, 573–580.

Ciach, G.J., and Krajewski, W.F., 1999. On the estimation of rainfall error variance. Adv. Water Resour., 2, 585–595.

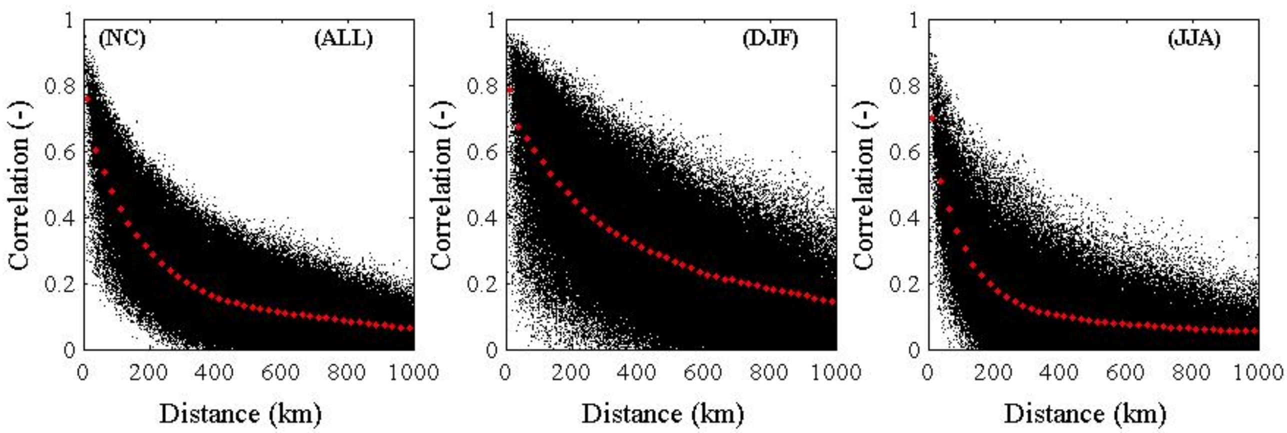
Ciach, G.J., Krajewski, W.F., and Villarini, G., 2007. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. J. Hydrometeorol., 8, 1325–1347.

Kirstetter, P.-E., Viltard, N. and Gosset, M. (2013), An error model for instantaneous satellite rainfall estimates: evaluation of BRAIN-TMI over West Africa. Q.J.R. Meteorol. Soc., 139: 894–911. doi: 10.1002/qj.1964

The difference in spatial resolution of the different products makes the comparison between sensors of different resolution delicate. Numerous studies have investigated the limitations of using rain gauge point measurements to evaluate area-averaged rainfall retrieved from sensors with coarser spatial resolution (Ciach and Krajewski 1999, Ciach et al. 2003, Habib et al. 2004, Ciach et al. 2007). Those differences can be due to the small-scale variability of rainfall averaged over the radar pixel and to deferring resolution between the gauge and the radar (Ciach and Krajewski 1999). For shorter accumulation period (from 5-min to several hours), the random sampling errors due to resolution differences can be dominant (Ciach and Krajewski 1999). With increasing accumulation period, the sampling errors decrease and are significantly reduced at the daily scale and beyond.

The comparisons presented in the paper at the annual (Figs 3-4) and seasonal scales (Figs 6-8) correspond to annual and seasonal accumulation over the 11-year period considered. Therefore, they correspond to long-term precipitation accumulation where the impact of the differing resolution is not significant. Previous studies have shown that for longer accumulation period (10-days) the correlation distance, that is the maximum distance between stations beyond which the correlations become insignificant, is of the order of several hundred kilometers (Gutowski et al. 2003). Those distances are several order of magnitude greater than the sensors spatial resolution.

However, when looking at the daily accumulations (section 5), the impact of the different resolutions needs to be considered. To investigate the impact of the resolution of each sensor, we computed for each RFC the interstation correlation of daily precipitation events. For each station, we selected the daily values greater than zero and computed the correlations regardless of the daily values of the other stations for the different RFCs. The correlations were computed for the annual and seasonal scales (DJF, JJA). The average correlation as a function of the distance was averaged over bins of 25-km. The figure below displays an example for the NCRFC, which represents a median situation in term of average rain rates (2.42mm/day). The red dotted lines represent the average values obtained for 25-km intervals.



The table below reports the values of the correlation distance (or distance beyond which the correlations become insignificant and defined as the distance corresponding to an average correlation of 1/e ≈ 0.37).

RFC	Average correlation distance (km) (distance at 1/e)		
	YEA	DJF	JJA
ABRFC	110	260	60
CBRFC	*60*	*140*	*30*
CNRFC	180	210	130
LMRFC	160	230	40
MARFC	210	380	80
MBRFC	110	210	60
NCRFC	130	310	80
NERFC	230	380	130
NWRFC	130	160	80
OHRFC	160	310	60
SERFC	110	210	40
WGRFC	80	210	60

We note that the correlation distance varies greatly from RFC to RFC and from season to season. At the daily scale considered, the averaged correlation distance is important and is greater than the spatial resolution of the radar (4-km) and the satellite (25-km). The longest correlation distances are found during winter compatible with the fact that cold precipitation is characterized by more widespread stratiform precipitation while warm season precipitation is characterized by more localized convective precipitation events. This seasonal dependency is comparable with results obtained by others over the central US and other time scales (Gutowski et al. 2003, Ciach et al. 2007). The CBRFC (Colorado Basin River RFC) is the one that displays the shorter correlation distance comparable with the satellite footprint (25-km), while the NERFC (Northeast RFC) displays the longer correlation distances.

The above discussion will be synthetized in the revised manuscript. We will add (p. 11496, l. 20):

To compare the different estimates for the annual average precipitation, we make the assumption that each rain gauge represents with sufficient accuracy the area-averaged rainfall over the native resolution of the different products evaluated: PRISM and Stage IV (4x4-km^2) and 3B42 and 3B42RT (4x4-km^2). While, there are well known limitations of using rain gauge point measurements to evaluate area-averaged rainfall retrieved from sensors with coarser spatial resolution (Ciach and Krajewski 1999, Ciach et al. 2003, Habib et al. 2004, Ciach et al. 2007), the random sampling errors due to differing resolutions are mostly dominant at the sub-daily scales (Ciach and Krajewski 1999). For accumulation period of several days the correlation distance (maximum distance between stations beyond which the correlations become insignificant), is of the order of several hundred kilometers (Gutowski et al. 2003). Those distances are several order of magnitude greater than the sensors spatial resolution.

For the daily precipitation section (section 5), we will add (p. 11508, l. 13):

We will assume that the rain gauge is representative of the grid-averaged rainfall for Stage IV. The computation of the interstation correlation for daily events indicated that the correlation distance was greater than the 4-km spatial resolution of the radar (not shown).

We will also add the references mentioned (Ciach and Krajewski 1999, Ciach et al. 2003, 2007, Gutowski et al. 2003, Habib et al. 2004). See also answer to Reviewer #1.

References:

Gutowski Jr., W. J., Decker, S. G., Donavon, R. A., Pan, Z., Arritt, R. W, and Takle E. S.: Temporal–Spatial Scales of Observed and Simulated Precipitation in Central U.S. Climate. *Journal of Climate*, 16, 3841-3847, 2003.

Habib, E., Ciach, G. J, and Krajewski, W. F.: A method for filtering out raingauge representativeness errors from the verification distributions of radar and raingauge rainfall. *Adv. Water Resour.*, 27, 967–980, 2004.

5. p. 11494 ll.21-22: "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)". The present paper is dealing with such artifacts and uncertainties, so would you mind remind them even briefly here?

The following will be added (p. 11494, l. 19):

The reader will find a more detailed description of the Stage IV precipitation retrievals from the RFC level and up to the final mosaicked product in Nelson et al. (2015). In addition to radar only reflectivity scanning and processing (beam blockage, hot and cold biases, bright-band contamination, anomalous propagation, cone of silence), the final mosaicked estimates present biases that are visible in the long-term averages. The fact that not all the RFCs use the same *precipitation estimation algorithm generates radar-to-radar and RFC-to-RFC discontinuities (Nelson et al. 2015).*

6. Could the authors provide some explanations for the seasonal bias adjustment on 3B42 described on p.11500 ll.22-26?

We realize that the sentence was confusing. In this case, the "bias-adjustment" is referring to the differences between 3B42RT (multi-satellite product) and the 3B42 (satellite-gauge product) and the incorporation of the monthly gage analysis (see the cited reference: Huffman and Bolvin 2013). Therefore, the section p. 11500 l. 22 to p. 11501 l. 3 will be modified into:

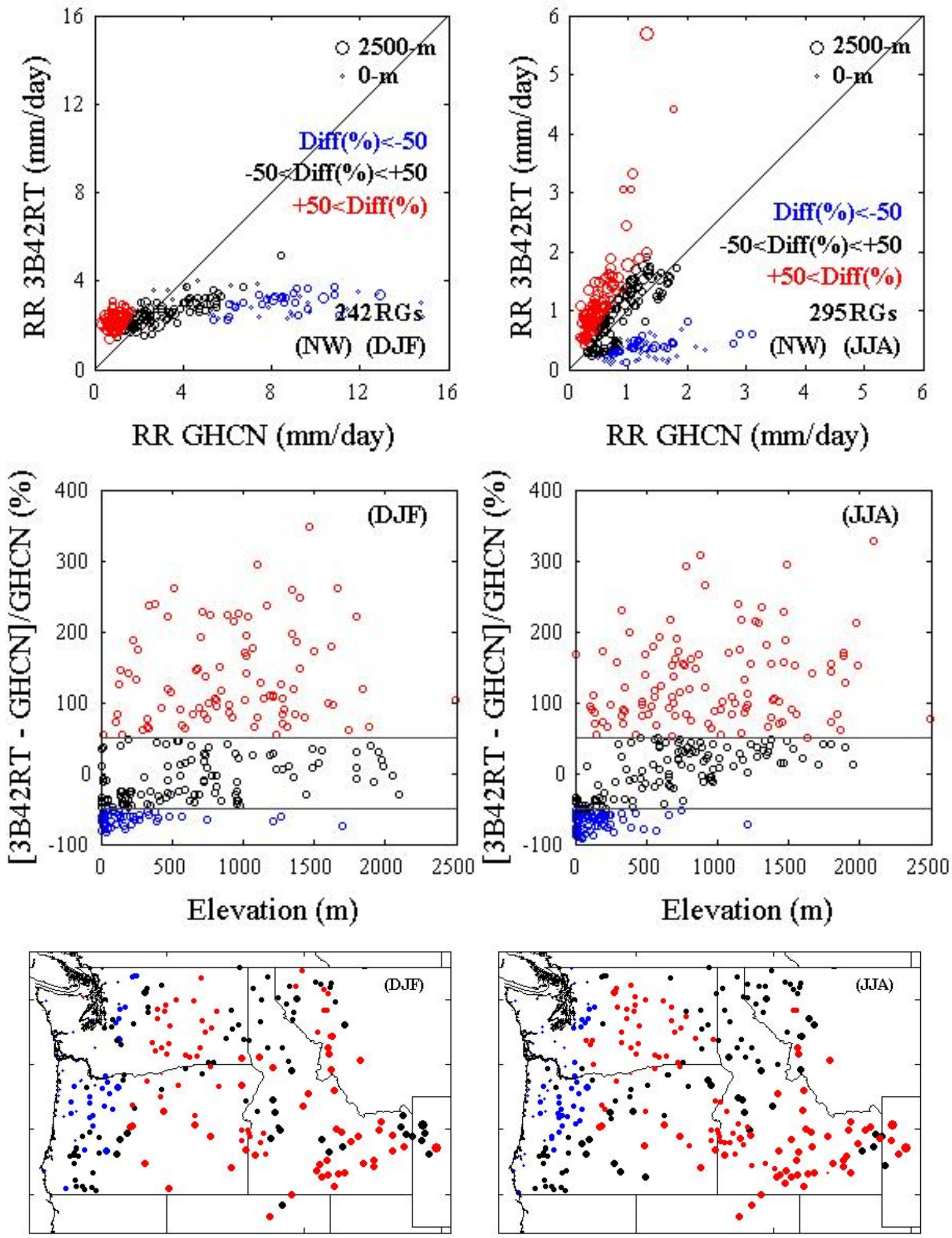
For winter, the bias-adjusted 3B42 precipitation estimates are lower than 3B42RT (3B42<3B42RT) over the Rockies (CB), over the highest latitudes along the US/Canadian border (NC, MB, NW), and East of the Mississippi (LM, SE). Conversely, the 3B42 estimates are found higher than the near-real time 3B42RT (3B42>3B42RT) along the West coast from Northern California up to the Pacific Northwest (NW, CN). For summer, the 3B42 estimates are found very significantly lower than 3B42RT (3B42<3B42RT) over the Midwest (MB, NC, AB). The rain gauge adjustment performed retrospectively corrects the possible overestimation of summertime convection by PMW sensors that mistake sub-cloud evaporation for precipitation (Dinku et al. 2010, 2011, Ochoa et al. 2014). Similarly, for the area of the Lower Mississippi domain (LM) located East of the Mississippi, 3B42 estimates are lower than 3B42RT.

7. Sections 3 and 4 mainly describe the results. Explanations for the observed discrepancies between the products, when provided, are simply listed randomly as in p.11502 ll. 24-29. More constructive and structured analysis and interpretation is necessary. What contribution does these sections bring to the state of knowledge?

We have modified substantially the sections 3 and 4 for clarity purposes and their respective contribution is better emphasized in the revised version. The revised sections included answers from comments #4, #6, #7, #9, #10, and #11.

8. On p. 11502 ll. 19-20: "The real-time 3B42RT displays moderate positive biases when compared to GHCN-D (+2.4%) and PRISM (+0.4%)". The scatterplot on Fig. 7b for 3B42RT shows much more discrepancies that compensate each other. Why is that?

Indeed, we noticed this point. While the bias-adjusted version 3B42 display points homogeneously scattered along the diagonal y=x, the real time 3B42RT presents more discrepancy (strong under/over estimation) that compensate (Q-Q plot along x=y).



Top) Scatterplot for the seasonal rain-rate for 3B42RT for winter (left) and summer (right) over Northwest (NW) displaying underestimation (blue) and overestimation (red) between 3B42RT and GHCN-D. Medium) Differences between 3B42RT and GHCN with respect to the elevation for DJF (left) and JJA (right). Bottom) Location of the points within the NWRFC for DJF (left) and JJA (right).

Results show that the 3B42RT pixels that display a strong underestimation (<-50%) with respect to GHCN are located West of the Cascades mountain range and for low to moderate elevation (h<500m) regardless of the season. Comparatively, the 3B42RT pixels that display a strong overestimation (>50%) are found East of the Cascades mountain range regardless of the elevation (200m<h<2500m). While, the average rain rate remains relatively constant East of the Cascades mountain range (R≈2mm/day), the seasonal differences are more important West of the Cascades with average rain rates less than 2mm/day in Summer to be compared with more than 5mm/day in Winter (and up to 15mm/day). Those difference being important (Diff. <-50%) regardless of the season, they illustrate the difficulties for satellite to capture orographic precipitation (DJF, JJA) and cold season precipitation (DJF). While the bias-adjustment (3B42) provides a good agreement with surface observation for summer (Fig. 7b), the bias-adjustment for winter (Fig. 7a) is more delicate due to the possible presence of cold precipitation.

To account for this point, we will add in the text (p. 11502, l. 19):

The real-time 3B42RT displays small biases when compared to GHCN-D (+2.4%) and PRISM (+0.4%). However, rather than the indication of a good performance, results show that the locations with overestimation are compensated by those with underestimations as can be seen with the Q-Q plot aligning along the diagonal. A closer look indicates that the 3B42RT pixels displaying the strongest underestimation (<-50%) with respect to GHCN-D are located West of the Cascades mountain range for low to moderate elevation (<500m) (not shown). Comparatively, the pixels that display the strongest overestimation (>50%) are found East of the Cascades regardless of the elevation. However, while the average rain rate remains relatively constant East of the Cascades throughout the year (R≈2mm/day), the seasonal differences are more important West of the Cascades with average rain rates of less than 2mm/day in Summer to be compared with more than 5mm/day in Winter. The important underestimation by 3B42RT West of the Cascades regardless of the season, illustrates the difficulties for satellite to capture orographic and cold season precipitation. We also note that despite the bias-adjustment, underestimation remains for 3B42 in winter due to uncertainties related to cold season precipitation measurements mentioned earlier or by rain gauge locations that cannot fully capture orographic effects that can be observed over distances smaller than the satellite resolution (Prat and Barros 2010a).

9. On p.11500 ll.26: "For summer, important negative bias adjustment (3B42<3B42RT) is found over the Midwest (MB, NC, AB) and corrects for the overestimation of summertime convection by PMW sensors that mistake sub-cloud evaporation for precipitation". Is there any citation for this statement?

Yes, there are few references documenting this point. In particular:

Dinku, T., Ruiz, F., Connor, S. J., and Ceccato, P.: Validation and Intercomparison of Satellite Rainfall Estimates over Colombia, *J. Appl. Meteorol. Climatol.*, 49, 1004–1014, doi:10.1175/2009JAMC2260.1, 2010.

Dinku, T., Ceccato, P., and Connor, S. J.: Challenges of satellite rainfall estimation over mountainous and arid parts of east Africa, *Int. J. Remote Sens.*, 30, 5965–5979, doi:10.1080/01431161.2010.499381, 2011.

Ochoa, A., Pineda, L., Crespo, P., and Willems, P.: Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific–Andean region of Ecuador and Peru. *Hydrol. Earth Syst. Sci.*, 18, 3179–3193, 2014.

The references will be added in the text and in the list of references (see also comment #6).

10. The gauge stations have notorious issues to quantify precipitation during the Winter season, caused e.g. by icing, underestimation related to snow drift, etc. This propagates to precipitation products making use of this information (Stage IV, 3B42). The authors need to address this point in order to provide critical insight on the precipitation estimates over CONUS.

Further discussion on rain gauge uncertainties in quantifying cold season precipitation will be added in the revised version. Several studies have documented the uncertainties in measuring solid precipitation with rain gauges. Among the rain gauges that are used either in the adjustment procedure of precipitations products or in the evaluation procedure of remotely sensed products, we find different types of sensors including heated automated tipping bucket rain gauges (ASOS, AWOS, HADS) or procedures relying on human intervention (GHCN-D). For solid precipitation, an exhaustive assessment of those uncertainties can be found in Goodison et al. (1998), while other publication summarize uncertainties related to point measurements of precipitation (Sevruk, et al. 2009, McMillan et al. 2012 among others). Among the systematic errors in measuring frozen precipitation, are evaporation, chimney effect, wind field deformation, wetting losses (McMillan et al. 2012), or possible delayed tips due to snow melting in the funnel (Prat and Barros 2010a) among others. Wind effects can lead to underestimation of frozen precipitation by 50% or more (Groisman et al. 1999). A recent publication (Leeper et al. 2015) provides a good overview of the methodology used to quantify winter precipitation and the related uncertainties for the GHCN-D dataset and the COOP stations subset.

We added in the text (p. 11501, l. 16):

In addition to fundamental limitations in radar and satellite measurement for snow and mixed precipitation events, additional uncertainties are introduced from in-situ data that are used either in the adjustment (HADS, ASOS, AWOS) or in the evaluation (GHCN-D) of remotely sensed products. Among the systematic errors in measuring frozen precipitation, are evaporation, chimney effect, wind field deformation, wetting losses, delayed tips due to snow melting in the funnel, or uncertainties due to human intervention in the measurement procedure (Goodison et al. 1998, Groisman et al. 1999, Sevruk et al. 2009, Prat and Barros 2010b, McMillan et al. 2012, Leeper et al. 2015). Although, this point is beyond the scope of this study, we note that the differences observed between remotely sensed and in-situ data for the higher latitude RFCs (CB, MB, NC, NE, NW, OH) are within the range of that observed for the other RFCs experiencing cold precipitation less frequently (-4.7% to -14.8% vs. -1.5% to -18.0% for Stage IV and -31.1% to +16.7% vs. -38.0% to +16.7% for 3B42) (Table 3).

References:

Goodison, B. E., Louie, P. Y. T., and Yang, D.: WMO solid precipitation measurement intercomparison. WMO Instruments and Observing Methods Rep., WMO/TD-872, 212 pp., 1998.

Groisman, P. Y., Peck, E. L., and Quayle, R. G.: Intercomparison of Recording and Standard Nonrecording U.S. Gauges , *J. Atmos. Ocean. Tech.*, 16, 602-609, 1999.

Leeper, R. D., Rennie, J., and Palecki, M. A.: Observational Perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: Temperature and Precipitation Comparison, *J. Atmos. Ocean. Tech.*, In press, 2015.

McMillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: rainfall, river discharge, and water quality, *Hydrol. Process.*, 26, 4078-4111, 2012.

Prat, O. P., and Barros, A. P.: Ground observations to characterize the spatial gradients and vertical structure of orographic precipitation - Experiments in the inner region of the Great Smoky Mountains, *J. Hydrol.*, 391, 143-158, doi: 10.1016/j.jhydrol.2010.07.013, 2010b.

Sevruk, B., Ondras, M., and Chvila, B.: The WMO precipitation measurement intercomparison, *Atmos. Res.*, 192, 376-380, 2009.

11. p. 11503 ll.27-29: Any explanation why the correction is insufficient in mountainous areas?

The main reason is the difficulty in capturing orographic precipitation over complex terrain. Important orographic rainfall enhancement can typically be observed over distances smaller than the satellite resolution (Prat and Barros, 2010b). Another reason is the difficulty in measuring cold precipitation both by the sensor and the rain gauge (see previous comment). Finally, the locations of the rain gauges

We will add (p. 11503, l. 27):

Overall, the bias-adjusted 3B42 performed very well over the Great Plains (MB, NC) to correct for the overestimation of summertime convection (Table 3) with comparable results than Stage IV (Table 3). Important differences remained however for low daily rainfall (< 1mm/day) and for the western RFCs during wintertime mostly due to the difficulty in capturing orographic precipitation and uncertainties in retrieving cold precipitation by the satellite (Chen et al. 2013, Huffman and Bolvin 2013) but also and by the rain gauges used in the bias-adjustment and the evaluation (Goodison et al. 1998, Groisman et al. 1999, Leeper et al. 2015).

Reference:

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, 2010b.

12. On p. 11505 ll.8-10. Chen et al. (2013) mention this issue. Chen, S., et al. (2013), Evaluation of the successive V6 and V7 TRMM multisatellite precipitation analysis over the Continental United States, Water Resour. Res., 49, 8174–8186, doi:10.1002/2012WR012795.

We thank the Reviewer for this suggestion. We added the reference at the end of the revised section 4.

Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge datasets **at daily to annual scales** (2002-2012)

Olivier P. Prat[1,*], and Brian 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

[2]Remote Sensing Applications Division (RSAD), NOAA/NESDIS/NCDC, Asheville, NC

Revised Version Submitted to Hydrology and Earth System Sciences

Special Issue: Precipitation: measurement and space-time variability

March 2015

*Corresponding author:
Dr. Olivier P. Prat
Cooperative Institute for Climate and Satellites-NC (CICS-NC)
North Carolina State University and NOAA/National Climatic Data Center
151 Patton Ave.
Asheville, NC 28801, USA
Phone: +1-828-257-3141
Email: olivier.prat@noaa.gov

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 (≥ 5 mm/day) 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 a contingency analysis conducted illustrated the challenge in retrieving extreme precipitation from remote sensing estimates.

1. Introduction

Over the last **decades**, 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 **have** specific characteristics 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. 2015), 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 (1-6h) 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**a**, Sahany et al. 2010, Kidd et al. 2012, Prat and Nelson 2013a,b, 2014 among others).

Olivier Prat 2/27/2015 10:04 AM
Deleted: Remotely sensed precipitation products are now coming of age where they can be considered for climatological applications.

Olivier Prat 2/27/2015 10:05 AM
Deleted: 30 years

Olivier Prat 3/1/2015 9:39 PM
Deleted: having

Olivier Prat 2/27/2015 11:13 AM
Deleted: 4

The purpose of this study is to evaluate the ability of QPE products to describe precipitation patterns and capture precipitation extremes **over a multi-annual time frame**. 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. In particular, we will investigate how the 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 **are** 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 **that's limited by day, month, or season. It is seldom that** studies that deal with the long-term assessment of precipitation products (annual or multi-annual basis) are available in the scientific literature **(Chen et al. 2013)**. 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

Olivier Prat 2/27/2015 10:06 AM
Deleted: in a climatological perspective

Olivier Prat 2/27/2015 10:07 AM
Deleted: and provide a long-term picture of the evolution on precipitation over time

Olivier Prat 3/1/2015 9:53 PM
Deleted: is

Olivier Prat 3/1/2015 9:55 PM
Deleted: rather limited in time (day, month, season).

Olivier Prat 3/1/2015 9:55 PM
Deleted: Seldom

Olivier Prat 2/27/2015 10:07 AM
Deleted: Therefore, this paper proposes to evaluate satellite precipitation estimates in the perspective of climate applications.

Olivier Prat 2/27/2015 10:35 AM
Deleted:)

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. **The duration of the study will allow the assessment of** 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.

Olivier Prat 3/1/2015 10:01 PM
Deleted: Although 11 years is not a long climatology, the duration of the study will be informative enough to derive long-term trends, assess

The paper is organized as follows: In **the** first section, we present briefly the precipitation datasets used in this study. In **the** second section, we will present a comparison between precipitation estimates at the annual and seasonal scales. In **the** 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 **summarizes** the major results of this study.

Olivier Prat 3/1/2015 10:01 PM
Deleted: a

Olivier Prat 3/1/2015 10:01 PM
Deleted: a

Olivier Prat 3/1/2015 10:01 PM
Deleted: a

Olivier Prat 3/1/2015 10:02 PM
Deleted: wraps up summarizing

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. **The GHCN-D datatset incorporates surface observations from different sources (see Table 2 in Menne et al. 2012). We selected the subset from the US-Cooperative Observing network (US-COOP), which represented about 9000 stations. The US-COOP network includes first order stations (1600 manual and automatic synoptic stations) and stations from volunteer observers.** 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 (Figure 1b). Although there is a 50% decrease **in** 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.

Olivier Prat 3/1/2015 10:03 PM
Deleted: of the dataset

Olivier Prat 3/1/2015 10:03 PM
Deleted: ,

Olivier Prat 3/1/2015 10:03 PM
Deleted: from

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, **and** degree dew point among others (Daly et al. 1994). Data are available at the daily, monthly, and annual scale and at various spatial resolution**s** (800m to 4km). In this work, we use the monthly precipitation estimates at the 4-km nominal spatial resolution (dataset AN81m:

PRISM Technical Note 2014). The PRISM precipitation estimates incorporate surface data observations from GHCN-D among others. **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.

Olivier Prat 3/1/2015 10:05 PM
Deleted: and the

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 has 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. **Stage IV represents the final stage of the process that combines mosaicked estimates from the 12 RCFs. The gauges used at the RFC level for bias-adjustment include available hourly rain gauges such as HADS (Hydrometeorological Automated Data System) gauges, ASOS (Automated Surface Observing System), and AWOS (Automated Airport Weather Stations) reports (Hou et al. 2014). Furthermore, some of the Western RFCs (Colorado Basin River:**

CBRFC, California-Nevada: CNRFC, Northwestern: NWRFC) do not use radar estimates due to poor coverage over mountainous areas. Those RFCs use an automated analysis of rain gauge observations (Mountain Mapper) that incorporates the gridded monthly precipitation climatology from PRISM. Although changes in gauge-adjustments procedures are always possible at the RFC level including the incorporation of the best available in-situ observations, it is reasonable to assume that the in-situ observations used in the Stage IV bias-adjustment procedure (HADS, ASOS, AWOS) are different from the US-COOP subset of GHCN-D used for the evaluation. For the Western RFCs however, the incorporation of the PRISM climatology that uses GHCN-D in-situ data will have to be kept in mind for any analysis. 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 retrievals from the RFC level and up to the final mosaicked product in Nelson et al. (2015). In addition to radar only reflectivity scanning and processing (beam blockage, hot and cold biases, bright-band contamination, anomalous propagation, cone of silence), the final mosaicked estimates present biases that are visible in the long-term averages. The fact that not all the RFCs use the same precipitation estimation algorithm generates radar-to-radar and RFC-to-RFC discontinuities (Nelson et al. 2015).**

Olivier Prat 2/27/2015 12:17 PM
Deleted: 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 microwave **datasets from low earth orbit satellites (TMI: The Microwave Imager; SSM/I: Special Sensor Microwave Imager; AMSR-E: Advanced Micro-wave Scanning Radiometer-Earth Observing System; AMSU-B: Advanced Microwave Sounding Unit-B)**, along with calibrated IR estimates of rain gauge corrected monthly accumulation (Huffman et al. 2007). TMPA 3B42 provides precipitation estimates for the domain 50ºS–50ºN at a 3-hourly and quarter degree resolution (0.25ºx0.25º) from which seasonal, daily, and sub-daily precipitation characteristics can be derived. **The quality of the blended precipitation estimates depends on the number of satellite estimates available at a given time stamp and on the sensor characteristics.** 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 (V6). The version 7 incorporates additional satellite products along with the reprocessed versions of the merged algorithms (**TMI, SSM/I, AMSR-E, AMSU-B**). 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 gauge analysis used is the GPCC Monitoring Product at 1-deg grid resolution (Schneider et al., 2010, 2011). This specific analysis uses SYNOP (synoptic weather observation reports) and CLIMAT reports that are received near-real time from 7000-8000 automated stations worldwide. While it is possible that some of the first order automated synoptic stations included in GHCN-**

Olivier Prat 2/27/2015 3:24 PM
Deleted: remotely sensed

Olivier Prat 2/27/2015 3:24 PM
Deleted: such as the TRMM Microwave Imager (TMI), the Special Sensor Microwave Imager (SSM/I), the Advanced Microwave Scanning Radiometer (AMSR), and the Advanced Microwave Sounding Unit (AMSU)

Olivier Prat 2/27/2015 3:19 PM
Deleted: V

Olivier Prat 2/27/2015 3:19 PM
Deleted: 6

Olivier Prat 2/27/2015 3:42 PM
Deleted: AMSU, TMI, AMSR, SSM/I

D are also used in the GPCC gauge analysis (SYNOP), most of the US-COOP subset of the GHCN-D stations used for evaluation are not a part of the GPCC Monitoring Product used in the 3B42. Being virtually impossible to track down and identify the automated stations that are or aren't used in the bias-adjustment procedure for 3B42, we are confident that this number remaining relatively low will not compromise the independent assessment of the 3B42 dataset. 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 Figure 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 systematic biases and the adjustment performed with respect to surface observations. Second, it aims to examine the suitability of satellite precipitation products to capture precipitation patterns and extremes in near real time.

3. Annual precipitation: differences between datasets

3.1 – Annual average precipitation

Olivier Prat 2/27/2015 3:33 PM
Deleted: in the perspective of long-term climatological applications.

Olivier Prat 3/1/2015 10:13 PM
Deleted: infer

Olivier Prat 3/1/2015 12:38 AM
Deleted: characteristic of

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 97ºW, 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 **that** retrieved from PRISM and Stage IV. The effect of monthly accumulation correction between 3B42 and 3B42RT is particularly **conspicuous** over the Northeast, the Rockies, and the Northeast. Over the Northeast, the annual average precipitation differences between 3B42 and 3B42RT are above +2 mm/day. Annual average precipitation differences between bias-adjusted and non-adjusted datasets are about +1 mm/day over the Northeast. Those differences are about -1.5 mm/day over the Rockies. CONUS-wide the mean average annual precipitation for the unadjusted 3B42RT is 2.62 mm/day; for 3B42 it is 2.54 mm/day (3% difference).

Olivier Prat 3/1/2015 10:15 PM
Deleted: noticeable

3.2 – Comparison with surface observations

To compare the different estimates for the annual average precipitation, we make the assumption that each rain gauge represents with sufficient accuracy the area-averaged rainfall over the native resolution of the different products evaluated: PRISM and Stage IV (4x4-km^2) and 3B42 and 3B42RT (4x4-km^2). While, there are well known limitations of using rain gauge point measurements to evaluate area-averaged rainfall retrieved from sensors with coarser spatial resolution (Ciach and

Krajewski 1999, Ciach et al. 2003, Habib et al. 2004, Ciach et al. 2007), the random sampling errors due to differing resolutions are mostly dominant at the sub-daily scales (Ciach and Krajewski 1999). For accumulation period of several days the correlation distance (maximum distance between stations beyond which the correlations become insignificant), is of the order of several hundred kilometers (Gutowski et al. 2003). Those distances are several order of magnitude greater than the sensors spatial resolution.

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. Over CONUS (Fig. 3a), we observe a very good agreement between **GHCN-D** surface observations 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 mm/day (σ = 1.11 mm/day) for PRISM and 2.47 mm/day (σ = 1.14 mm/day) for GHCN-D. The differences observed toward higher rain rates (R> 6 mm/day) 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 **(Daly et al. 1994)**. 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 **S**tage IV for rain rates greater than 4 mm/day. 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 for 3B42 and 2.62 mm/day for 3B42RT along with a lower correlation coefficient (Fig. 3a). However,

Olivier Prat 2/27/2015 2:50 PM
Deleted: network

Olivier Prat 2/27/2015 2:51 PM
Deleted: from the GHCN-D network

Olivier Prat 2/27/2015 10:37 AM
Deleted: s

Olivier Prat 2/28/2015 7:21 PM
Deleted: Stage IV displays a lower mean average precipitation (-6%) when compared to GHCN-D surface stations (4075 rain gauges).

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) 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) presents the best agreement with GHCN-D (3.75 mm/day). Little differences are found between 3B42 and 3B42RT **in terms of** average rain rate of 3.87 mm/day and 3.90 mm/day respectively **with however a narrower distribution (lower σ) for the bias-adjusted 3B42 than for 3B42RT.** Over the Missouri Basin River (MB), **important differences are observed between 3B42 and 3B42RT with respect to GHCN-D. The bias-adjusted 3B42 displays a satisfying agreement with surface observations that contrasts with the overestimation displayed by the near real time 3B42RT observations (Fig. 3c). Over Northwest (NW), the bias-adjusted 3B42 presents a substantial improvement when compared to 3B42RT but severe underestimation remains for precipitation above 4 mm/day (Fig. 3d). Although closer to surface observations, Stage IV displays a similar underestimation at higher rain rates.** Table 2 summarizes the differences between GHCN-D and the **other** datasets. For PRISM, the linear regression coefficient when compared to surface observation (GHCN-D) remains within a narrow range (0.97

Olivier Prat 2/28/2015 7:18 PM
Deleted:

Olivier Prat 2/28/2015 10:10 PM
Deleted: precipitation estimates with an

Olivier Prat 2/28/2015 4:37 PM
Deleted: 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).

Olivier Prat 2/28/2015 5:01 PM
Deleted: Stage IV (1.59 mm/day) and 3B42 (1.74 mm/day) display a very good agreement with GHCN-D (1.59 mm/day), while the near real time 3B42RT present a rain rate 50% higher (2.37 mm/day).

Olivier Prat 2/28/2015 6:03 PM
Deleted: 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 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.

Olivier Prat 2/28/2015 6:04 PM
Deleted: different

to 1.03) for the different RFCs considered. **The differences are statistically significant at the 5% significance for about half of the RFCs (5 over 12).** For Stage IV, the variations are greater and indicate a general underestimation with (a) varying between 0.87 and 1 **and statistically significant differences for 9 of the RFCs.** 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).

Figure 4a displays the average annual precipitation derived from all datasets for the different RFCs. The Lower Mississippi River Basin (LM) exhibits the higher average annual rain rate regardless of the dataset**.** **The** Colorado Basin River (CB) displays the lower average annual rain rate for GHCN-D, **PRISM,** Stage IV, and 3B42**.** **For** 3B42RT **the minimum** is found for California Nevada (CN). The differences with respect to GHCN-D are presented in Figure 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. CONUS**-wide**, 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**, Huffman and Bolvin 2013**). 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

Olivier Prat 2/28/2015 6:20 PM
Deleted: 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).

Olivier Prat 2/28/2015 7:02 PM
Deleted: (GHCN-D, PRISM, Stage IV, 3B42, 3B42RT)

Olivier Prat 2/28/2015 7:03 PM
Deleted: Over CONUS, GHCN-D (2.47 mm/day) and PRISM (2.42 mm/day) present comparable rain rates with a relative difference of 2%. Stage IV presents a slightly lower average annual rain rate (2.31 mm/day) 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 for 3B42 and 2.62 mm/day 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).

Olivier Prat 2/28/2015 7:05 PM
Deleted: (GHCN-D, PRISM, Stage IV, 3B42, 3B42RT) while the

Olivier Prat 2/28/2015 7:06 PM
Deleted: (while the minimum for

Olivier Prat 2/28/2015 7:12 PM
Deleted:

Olivier Prat 2/28/2015 7:12 PM
Deleted: (%)

Olivier Prat 2/28/2015 7:12 PM
Deleted: surface measurements

Olivier Prat 2/28/2015 10:13 PM
Deleted: 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

is to account for the overestimation of summertime convection by Passive Microwave retrieval that tends to associate important sub-cloud evaporation with precipitation **as we will see in the next section**. The Colorado Basin (CB) RFC is the domain that displays the most important difference between 3B42RT and 3B42 (+42%). For the remaining RFCs, the differences between 3B42RT and 3B42 remains moderate between -21% and +12% (MA: -17%, NC: +12%, NE: -21%). **Stage IV presents globally lower differences with GHCN-D (-14% to +1%) and PRISM (-17% to +4%) than the satellite estimates 3B42 (-28% to +7%) and 3B42RT (-32% to +49%). At the RFC level, Stage IV almost systematically underestimates precipitation except for two RFCs (AB, MB) with a lower rainfall of -7% when compared to GHCN-D CONUS-wide (Table 2). For the western RFCs (CB, NW), Stage IV presents a better agreement with surface observations (-12%) than 3B42 (-23%) and 3B42RT (-25%). The lower differences can be explained by the fact that the western RFCs use the Mountain Mapper approach and gauge-only estimates (Hou et al. 2014, Nelson et al. 2015).**

4. Seasonal precipitation

4.1 – Seasonal precipitation patterns

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). PRISM and Stage IV present **similar** precipitation patterns regardless of the season. **By comparison with** 3B42RT, the monthly-adjusted 3B42 displays precipitation patterns visually closer to those of PRISM. **The** differences between 3B42RT and 3B42 are more

Olivier Prat 2/27/2015 10:37 AM
Deleted: 3

Olivier Prat 3/1/2015 12:26 AM
Deleted: 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.

Olivier Prat 3/1/2015 12:46 AM
Deleted: A simple comparison shows that

Olivier Prat 3/1/2015 12:44 AM
Deleted: comparable

Olivier Prat 3/1/2015 12:57 AM
Deleted: When compared to the near real time

Olivier Prat 3/1/2015 1:00 AM
Deleted: that are

Olivier Prat 3/1/2015 1:01 AM
Deleted: near real time

Olivier Prat 3/1/2015 1:01 AM
Deleted: bias adjusted

emphasized on a seasonal basis than observed for the annual basis (Fig. 2). **For winter, the bias-adjusted 3B42 precipitation estimates are lower than 3B42RT (3B42<3B42RT) over the Rockies (CB), over the highest latitudes along the US/Canadian border (NC, MB, NW), and East of the Mississippi (LM, SE). Conversely, the 3B42 estimates are found higher than the near-real time 3B42RT (3B42>3B42RT) along the West coast from Northern California up to the Pacific Northwest (NW, CN). For summer, the 3B42 estimates are found to be very significantly lower than 3B42RT (3B42<3B42RT) over the Midwest (MB, NC, AB). The rain gauge adjustment performed retrospectively corrects the possible overestimation of summertime convection by PMW sensors that mistake sub-cloud evaporation for precipitation (Dinku et al. 2010, 2011, Ochoa et al. 2014). Similarly, for the area of the Lower Mississippi domain (LM) located East of the Mississippi, 3B42 estimates are lower than 3B42RT.**

4.2 – Comparison with surface observations

Figures 6a-b present the seasonal rain-rates derived from the different datasets. Between the warm and cold season, average seasonal rain rate derived from surface observations (GHCN-D, PRISM) vary from -95% **for** CN to +270% **for** MB **(Table 3)**. For winter, the minimum (maximum) average rain rate is found for MB (LM) with 0.64 mm/day (3.77 mm/day). For summer, the minimum (maximum) average rain rate is found for CN (SE) with 0.17 mm/day (4.53 mm/day). **We note that the seasonal minima and maxima are observed for the same RFCs regardless of the dataset except for the winter minimum for 3B42RT (Table 3).** Seasonal differences between

Olivier Prat 2/27/2015 12:21 PM
Deleted: 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 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.

Olivier Prat 3/1/2015 8:35 AM
Deleted: (GHCN-D, PRISM, Stage IV, 3B42, 3B42RT)

Olivier Prat 3/1/2015 8:35 AM
Deleted: (

Olivier Prat 3/1/2015 8:35 AM
Deleted:)

Olivier Prat 3/1/2015 8:35 AM
Deleted: (

Olivier Prat 3/1/2015 8:35 AM
Deleted:)

Olivier Prat 3/1/2015 8:36 AM
Deleted: 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.

GHCN-D and PRISM remain moderate (-5.9% to +2.1%) and comparable to that for the annual basis (Fig. 6c-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. **In addition to fundamental limitations in radar and satellite measurement for snow and mixed precipitation events, additional uncertainties are introduced from in-situ data that are used either in the adjustment (HADS, ASOS, AWOS) or in the evaluation (GHCN-D) of remotely sensed products. Among the systematic errors in measuring frozen precipitation, are evaporation, chimney effect, wind field deformation, wetting losses, delayed tips due to snow melting in the funnel, or uncertainties due to human intervention in the measurement procedure (Goodison et al. 1998, Groisman et al. 1999, Sevruk et al. 2009, Prat and Barros 2010b, McMillan et al. 2012, Leeper et al. 2015). Although, this point is beyond the scope of this study, we note that the differences observed between remotely sensed and in-situ data for the higher latitude RFCs (CB, MB, NC, NE, NW, OH) are within the range of that observed for the other RFCs experiencing cold precipitation less frequently (-4.7% to -14.8% vs. -1.5% to -18.0% for Stage IV and -31.1% to +16.7% vs. -38.0% to +16.7% for 3B42) (Table 3).** For summer, differences between 3B42 and GHCN-D present a narrower range from -2% to +25% (overall overestimation). Those differences represent a substantial improvement **over** 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 M**B** and CR in winter (+111% and +58%

Olivier Prat 3/1/2015 10:25 PM
Deleted: with

Olivier Prat 3/1/2015 9:00 AM
Deleted: R

respectively), or CN and M**B** 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 **into** seasonal differences can be seen in Figure 7 that displays scatterplots and Q-Q plots for the seasonal rain-rate over Northwest (NW). Regardless of the season (winter: Fig. 7a; summer: Fig. 7b), PRISM presents a very good agreement with surface observations with **differences of 2-3% for the average rain-rate regardless of the season.** There is a four-fold difference between the maximum rain rates for winter (R ≈ 12 mm/day) and summer (R ≈ 3 mm/day). 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 mm/day) when compared with GHCN-D (R > 12 mm/day). 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 mm/day, the bias-adjusted 3B42 present a significant improvement when compared to the near real time 3B42RT. However, **a strong negative bias remains for 3B42 that displays a comparable mean seasonal precipitation (≈ 2.5 mm/day) than 3B42RT,** which is about -30% when compared to surface observations. Summer exhibits average rain rates (≈ 0.82 mm/day) 4 times lower than winter (3.5 mm/day) (Fig. 7b). Stage IV displays a negative bias for summer when compared to GHCN-D and PRISM (≈ -19%) **and comparable with the bias observed for winter (-12%). Differences remain significant despite the fact that Stage IV uses the PRISM/Mountain Mapper algorithm that combines automated rain gauge observations and PRISM monthly**

Olivier Prat 3/1/2015 9:00 AM
Deleted: R

Olivier Prat 3/1/2015 10:25 PM
Deleted: at

Olivier Prat 3/1/2015 9:02 AM
Deleted: for PRISM, Stage IV, 3B42, and 3B42RT with respect to GHCN-D

Olivier Prat 3/1/2015 9:06 AM
Deleted: linear regression and R^2 coefficients close to 1 (a=1.02; 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%.

Olivier Prat 3/1/2015 9:14 AM
Deleted: and PRISM

Olivier Prat 3/1/2015 9:18 AM
Deleted: both 3B42 and 3B42RT display a comparable mean seasonal averaged precipitation (≈ 2.5 mm/day),

Olivier Prat 3/1/2015 9:22 AM
Deleted: as reported by GHCN-D and PRISM

precipitation climatology (Hou et al. 2014, Nelson et al. 2015). Conversely, 3B42 presents a very good agreement with GHCN-D (-1.7%) and PRISM (-2.4%) and contrast**s** with the severe underestimation observed on the right side of the distribution during winter (Fig. 7a). **The real-time 3B42RT displays small biases when compared to GHCN-D (+2.4%) and PRISM (+0.4%). However, rather than the indication of a good performance, results show that the locations with overestimation are compensated by those with underestimations as can be seen with the Q-Q plot aligning along the diagonal. A closer look indicates that the 3B42RT pixels displaying the strongest underestimation (<-50%) with respect to GHCN-D are located West of the Cascades mountain range for low to moderate elevation (<500m) (not shown). Comparatively, the pixels that display the strongest overestimation (>50%) are found East of the Cascades regardless of the elevation. However, while the average rain rate remains relatively constant East of the Cascades throughout the year (R≈2mm/day), the seasonal differences are more important West of the Cascades with average rain rates of less than 2mm/day in Summer to be compared with more than 5mm/day in Winter. The important underestimation by 3B42RT West of the Cascades regardless of the season, illustrates the difficulties for satellite to capture orographic and cold season precipitation. We also note that despite the bias-adjustment, underestimation remains for 3B42 in winter due to uncertainties related to cold season precipitation measurements mentioned earlier or by rain gauge locations that cannot fully capture orographic effects that can be observed over distances smaller than the satellite resolution (Prat and Barros 2010a).**

Olivier Prat 3/1/2015 9:25 AM
Deleted: Similarly

Olivier Prat 3/1/2015 9:26 AM
Deleted: rainfall

Olivier Prat 2/27/2015 12:28 PM
Deleted: *The real-time 3B42RT displays moderate positive biases when compared to GHCN-D (+2.4%) and PRISM (+0.4%).*

Further illustration of the importance of the bias-adjustment 3B42 can be found in Figure 8 that displays comparisons over the Missouri Basin River (MB). Again, GHCN-D and PRISM present an average rain rate difference of about 3% regardless of the season. Similarly, Stage IV presents a good agreement with surface observations with a small underestimation of -5% for winter (Fig. 8a) and a moderate overestimation of +8.4% for summer (Fig. 8b). For both the cold and the warm season, the improvement brought by the 3B42 bias-adjustment is clearly visible. For winter, the near real time 3B42RT exhibits a severe overestimation of +147% with respect to surface observation, which is reduced to +16.7% for the bias-adjusted 3B42 (Table 3). A closer look at the wintertime precipitation (Fig. 5d) indicates higher rainfall accumulation for 3B42RT at higher latitudes and along the edges of the MB RFC when compared to the other datasets (Fig. 5a-c). These differences are certainly associated with cold season precipitation and are due to the challenge of measuring falling snow and frozen precipitation and precipitation over snow and ice-covered areas by sensors (SSM/I, AMSU-B) used in the near-real time 3B42RT (Huffman and Bolvin 2013). For summer, the bias-adjusted 3B42 exhibits moderate differences of +7.7% with respect to GHCN-D to be compared with an overestimation of +61% for 3B42RT (Table 3). The overestimation of 3B42RT is found consistently throughout the rain rate spectra (Fig. 8b). As mentioned earlier, the 3B42RT overestimation is due to uncertainties in PMW retrievals that associate summertime sub-cloud evaporation with precipitation (Dinku et al. 2010, 2011, Ochoa et al. 2014). The monthly bias-adjustment (3B42) corrects efficiently for both the cold and warm season 3B42RT rainfall overestimation with a reduction of the

Olivier Prat 3/1/2015 1:13 PM
Deleted: 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 block ... [1]

Olivier Prat 3/1/2015 10:05 AM
Deleted: Figure 8 displays the scatterplots over the Missouri Basin River (MB) for ... [2]

Olivier Prat 3/1/2015 10:18 AM
Deleted: For winter, there is a good agreement between Stage IV and GHC ... [3]

Olivier Prat 3/1/2015 10:42 AM
Deleted: During winter, the near real time 3B42RT exhibits a strong overestimati ... [4]

Olivier Prat 3/1/2015 12:04 PM
Deleted: seasonal rainfall

Olivier Prat 3/1/2015 12:04 PM
Deleted: shows that

Olivier Prat 3/1/2015 12:05 PM
Deleted: displays a significantly higher rainfall accumulation

Olivier Prat 3/1/2015 12:05 PM
Deleted: Missouri Basin River

Olivier Prat 3/1/2015 12:05 PM
Deleted: PRISM, Stage IV, 3B42

Olivier Prat 3/1/2015 12:11 PM
Deleted: due to

Olivier Prat 3/1/2015 12:22 PM
Deleted: (snow, frozen precipitation).

Olivier Prat 3/1/2015 11:09 AM
Deleted: For summer, 3B42RT displays a severe overestimation (+61%) when co ... [5]

Olivier Prat 3/1/2015 11:10 AM
Deleted: when compared to GHCN-D

Olivier Prat 3/1/2015 11:56 AM
Deleted: These strong differences are due to uncertainties in PMW retrieval in prese ... [6]

Olivier Prat 3/1/2015 1:38 PM
Deleted: performed

Olivier Prat 3/1/2015 1:39 PM
Deleted: accumulation

Olivier Prat 3/1/2015 1:39 PM
Deleted: an adequate

average rain rate of about 110% and 50% for winter and summer respectively. **Overall, the bias-adjusted 3B42 performed very well over the Great Plains (MB, NC) to correct for the overestimation of summertime convection (Table 3) with comparable results than Stage IV (Table 3). Important differences remained however for low daily rainfall (< 1mm/day) and for the western RFCs during wintertime mostly due to the difficulty in capturing orographic precipitation and uncertainties in retrieving cold precipitation by the satellite (Chen et al. 2013, Huffman and Bolvin 2013) but also and by the rain gauges used in the bias-adjustment and the evaluation (Goodison et al. 1998, Groisman et al. 1999, Leeper et al. 2015).**.

5. Daily precipitation

5.1 – Conditional analysis and extreme precipitation

After investigating the ability of the different datasets to describe precipitation patterns, this section investigates their ability to capture intense and extreme precipitation at the daily scale. **A conditional analysis was conducted using different thresholds for daily accumulation (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 7:00LST-7:00LST for most of the locations, which corresponds to 12Z-12Z on the Eastern US. Therefore, **some uncertainties could arise from computing daily accumulation** over a slightly different time period. Although the number of rainy days appears consistent in term of magnitude for the

Olivier Prat 3/1/2015 1:39 PM
Deleted: between monthly-adjusted and near real time precipitation

Olivier Prat 2/27/2015 1:56 PM
Deleted: 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.

Olivier Prat 3/1/2015 1:14 AM
Deleted: , extremes, and conditional analysis

Olivier Prat 3/1/2015 12:24 PM
Deleted: accurately

Olivier Prat 3/1/2015 12:41 PM
Deleted: in addition to sensor characteristics, differing spatial and temporal resolution, differences could be explained by accumulation computed

different observation platforms, there are noticeable differences over specific areas. Despite a delicate visual comparison between point data measurements from GHCN-D and Stage IV gridded estimates due to the scarcity of station coverage, both products present a very similar pattern and a comparable number of rainy days throughout 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 the satellite's coarser resolution. Overall, the rain gauge adjusted radar (Stage IV) and satellite (3B42) datasets display a satisfying 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. For daily accumulation greater than the Wet Millimeter Days (WMMD: R > 17.8 mm/day), 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

Olivier Prat 3/1/2015 12:42 PM
Deleted: Although the visual

Olivier Prat 3/1/2015 12:45 PM
Deleted: (GHCN-D)

Olivier Prat 3/1/2015 12:45 PM
Deleted: (Stage IV) is delicate

Olivier Prat 3/1/2015 12:46 PM
Deleted: surface

Olivier Prat 3/1/2015 12:46 PM
Deleted: similar

Olivier Prat 3/1/2015 12:47 PM
Deleted: over

Olivier Prat 3/1/2015 12:51 PM
Deleted: a

Olivier Prat 3/1/2015 11:22 PM
Deleted: The seasonal breakdown presented in Figure 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.

highest daily average over the area considered (Shepherd et al. 2007). For North America, this maximum daily average (17.8 mm/day) is recorded in Henderson Lake (British Columbia) (Source NCDC). 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 bias-adjusted satellite estimates 3B42. The biggest differences are observed with 3B42RT that shows a much lower number of rainy days greater than WMMD **(Fig. 9b)**. This is consistent with the underestimation observed for the daily averages for 3B42RT and to a lesser extent for 3B42 (Fig. 3d, Fig. 7a-b). The bias-adjustment increases the number of WMMD **of 3B42 closer to Stage IV levels.** Conversely, **3B42 and 3B42RT** that displayed less rainy days than Stage IV (Fig. 9a), presents a higher occurrence of WMMD when compared to Stage IV **over the Northeast (NE), the upper part of the North Central (NC) domain, and the Lower Mississippi (LM)** (Fig. 9b). **For daily accumulation greater than 2 in day^{-1} (> 50.8 mm/day: Karl and Plummer 1995; hereafter EPD2), GHCN-D and Stage IV display comparable counts for EPD2 over the Eastern US where rain gauges coverage is denser (Fig. 9c).** Over the Northeast (NE) and the Southeastern US (LM, SE), 3B42 and 3B42RT display a higher number of **days with** rainfall above 2 in day^{-1} as compared to Stage IV. **Daily** precipitation greater than 4 in day^{-1} (**> 101.6 mm/day:** Barlow 2011**; hereafter EPD4**) is limited to the Pacific Coast and East of 100°W, a domain regularly impacted by tropical cyclones (Prat and Nelson 2013a,b, 2014a) **(Fig. 9d)**. **These EPD4 events are relatively infrequent (3 counts or less by year) and roughly correspond to**

Olivier Prat 3/1/2015 11:45 PM
Deleted: the

Olivier Prat 3/1/2015 11:45 PM
Deleted: Depending on the RFC considered, WMMD events ranges from the 82- to 97-percentiles for LM and CB respectively (Nelson et al. 2014).

Olivier Prat 3/1/2015 11:52 PM
Deleted: (Fig. 9b)

Olivier Prat 3/2/2015 12:40 AM
Deleted: rain gauge

Olivier Prat 3/2/2015 12:40 AM
Deleted: performed for 3B42RT

Olivier Prat 3/2/2015 12:22 AM
Deleted: days with accumulation greater than

Olivier Prat 3/2/2015 12:40 AM
Deleted: to level closer to Stage IV yet remaining below the Stage IV levels.

Olivier Prat 3/2/2015 12:27 AM
Deleted: the Northeast (NE)

Olivier Prat 3/2/2015 12:27 AM
Deleted: for 3B42 and 3B42RT

Olivier Prat 3/2/2015 12:27 AM
Deleted: for

Olivier Prat 3/2/2015 12:46 AM
Deleted: 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 (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.

Olivier Prat 3/2/2015 1:26 AM
Deleted: Again, the bias-adjusted 3B42 present a higher number of days above 2 inches than the real-time 3B42RT but less than for Stage IV. For

Olivier Prat 3/2/2015 1:40 AM
Deleted: intense

Olivier Prat 3/2/2015 1:44 AM
Deleted: i.e., above 101.6 mm/day (EPD4), the area

the 0.1-0.5% top daily events regardless of the RFC considered. The bias-adjusted 3B42 and real-time 3B42RT display a comparable number of EPD4 events over the Southeastern US. The maximum occurrences are observed over the Lower Mississippi (LM) RFC and are higher than the daily counts for Stage IV. Over the Northwest, the bias-adjusted (3B42) is able to better capture those extreme daily accumulation events (EPD4) with respect to 3B42RT that displays almost no days with rainfall above 4 in day^{-1}.

More quantitative information can be found in Figure 10 that displays the proportion of rain gauges (GHCN-D) and the corresponding radar (Stage IV) or satellite (3B42, 3B42RT) pixels 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 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. Some 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 at 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

Olivier Prat 3/2/2015 1:58 AM
Deleted: While the

Olivier Prat 3/2/2015 2:10 AM
Deleted: with a

Olivier Prat 3/2/2015 1:49 AM
Deleted: higher occurrence

Olivier Prat 3/2/2015 2:16 AM
Deleted: LM

Olivier Prat 3/2/2015 2:23 AM
Deleted: , 3B42RT displays 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.

Olivier Prat 3/2/2015 2:18 AM
Deleted: results indicate a systematic underestimation of intense and extreme precipitation by 3B42RT.

Olivier Prat 3/2/2015 2:18 AM
Deleted: T

Olivier Prat 3/2/2015 2:18 AM
Deleted: ment procedure

Olivier Prat 3/2/2015 2:21 AM
Deleted: that seem to remain underestimated when compared to Stage IV.

Olivier Prat 3/1/2015 12:53 PM
Deleted: A few

Olivier Prat 3/1/2015 10:30 PM
Deleted: for

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. **However those results have to be interpreted with caution as they present a count of the daily events over the 11-year period. The number of events decreases with increasing rain rate and the WMMD correspond roughly to the 90th percentile precipitation events regardless of the RFC (Nelson et al. 2015). A test was performed to determine the interstation correlation of daily precipitation events corresponding to the 90th percentile (not shown). For each station, the correlation was computed using the daily events greater than the 90th percentile regardless of the values of the other stations. Results showed that for those high-intensity events, the average correlation distance was about 30-80-km which is comparable with the satellite footprint.**

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) presents consistently the lowest average number of events by active stations regardless of the daily accumulation **(WMMD, EPD2, EPD4)**. On the other hand, the Lower Mississippi (LM) present the highest average number of events regardless of the sensor considered. For selected RFCs, the differences between 3B42 and 3B42RT are particularly important for EPD2 and EPD4. The biggest differences are found for Northwest (NW) and **the** Missouri Basin River (MB). For the **latest**, 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 **mentioned earlier**. 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 caution **is necessary** when analyzing those results.

5.2 – Contingency analysis between Stage IV and GHCN-D

The previous results were provided for the entire period 2002-2012. Figure 12 displays a contingency analysis between daily precipitation from the **GHCN-D stations** and the corresponding **Stage IV** radar pixel. **We will assume that the rain gauge is representative of the grid-averaged rainfall for Stage IV. The computation of the interstation correlation for daily events indicated that the correlation distance was greater than the 4-km spatial resolution of the radar (not shown).** The number of

Olivier Prat 3/1/2015 9:09 PM
Deleted: RFC

Olivier Prat 3/1/2015 9:12 PM
Deleted: Missouri Basin River

Olivier Prat 3/1/2015 12:56 PM
Deleted: that

Olivier Prat 3/1/2015 12:57 PM
Deleted: should be made

Olivier Prat 3/1/2015 12:48 PM
Deleted: surface stations (GHCN-D)

Olivier Prat 3/1/2015 12:48 PM
Deleted:

Olivier Prat 3/1/2015 12:48 PM
Deleted: (Stage IV)

rainy days (R > 0 mm/day) 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 **events at the** radar **pixel** only and from 8% (AB) to 22% (CN) for rain gauge only events. With increasing rain rate the **number of** events observed simultaneously by the gauge and the radar decreased from 62% (R > 0 mm/day: 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 mm/day), the number of events missed by the radar increases from 14% (R > 0 mm/day) to 45% (EPD4). In addition for accumulation greater than 2 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 in day^{-1}, 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 mm/day: 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 he 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

Olivier Prat 2/27/2015 10:44 AM
Deleted: 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 **pixel** only (right column) remains relatively constant. For concurrent rainfall events (median column) the ratio decreases significantly for daily accumulation greater than 2 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 **or** 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).

Olivier Prat 3/1/2015 10:46 PM
Deleted: of

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 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 mm/day) over CONUS. Discrepancies can be explained by the difference between point (RG) and area (satellite, radar) measurements and the difficulty **in capturing** localized, convective, and orographic events due to the

Olivier Prat 3/1/2015 10:48 PM
Deleted: to capture

coarser resolution. Furthermore, those differences can be more important at the seasonal scale and for selected basins in particular over the Western part of CONUS (Pacific Northwest, Rocky Mountains) due to the difficulty **in retrieving** 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% +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 skills 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

Olivier Prat 3/1/2015 10:48 PM
Deleted: to retrieve

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 **the** Eastern US, results illustrate the challenge of retrieving extreme precipitation (top 1% percentile) from remote sensing.

Olivier Prat 3/2/2015 3:13 AM

Deleted: A similar study is currently underway for the different satellite QPE products available (3B42, 3B42RT, 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.

Acknowledgement

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. Bull. Am. Meteorol. Soc., in press, 2015.

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

Ciach, G. J., Habib, E., and Krajewski, W. F.: Zero-covariance hypothesis in the error variance separation method of radar rainfall verification. Adv. Water Resour., 26, 573-580, 2003.

Ciach, G. J., and Krajewski, W. F.: On the estimation of rainfall error variance. Adv. Water Resour., 2, 585-595, 1999.

Ciach, G. J., Krajewski, W. F., and Villarini, G.: Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. J. Hydrometeorol. 8, 1325-1347, 2007.

Chen, S., Hong, Y., Gourley, J. J., Huffman, G. J., Tian, Y., Cao, Q., Yong, B., Kirstetter, P.-E., Hu, J., Hardy, J., Li, Z., Khan, S. I., and Xue, X.: Evaluation of the successive V6 and V7 TRMM multisatellite precipitation analysis over the Continental United States, Water Resour. Res., 49, 8174–8186, doi:10.1002/2012WR012795, 2013.

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. Meteor., 33, 140-158, 1994.

Dinku, T., Ruiz, F., Connor, S. J., and Ceccato, P.: Validation and Intercomparison of Satellite Rainfall Estimates over Colombia, J. Appl. Meteorol. Climatol., 49, 1004-1014, doi:10.1175/2009JAMC2260.1, 2010.

Dinku, T., Ceccato, P., and Connor, S. J.: Challenges of satellite rainfall estimation over mountainous and arid parts of east Africa, Int. J. Remote Sens., 30, 5965-5979, doi:10.1080/01431161.2010.499381, 2011.

Goodison, B. E., Louie, P. Y. T., and Yang, D.: WMO solid precipitation measurement intercomparison. WMO Instruments and Observing Methods Rep., WMO/TD-872, 212 pp., 1998.

Groisman, P. Y., Peck, E. L., and Quayle, R. G.: Intercomparison of Recording and Standard Nonrecording U.S. Gauges , J. Atmos. Ocean. Tech., 16, 602-609, 1999.

Gutowski Jr., W. J., Decker, S. G., Donavon, R. A., Pan, Z., Arritt, R. W, and Takle E. S.: Temporal–Spatial Scales of Observed and Simulated Precipitation in Central U.S. Climate. J. Climate, 16, 3841-3847, 2003.

Habib, E., Ciach, G. J, and Krajewski, W. F.: A method for filtering out raingauge representativeness errors from the verification distributions of radar and raingauge rainfall. Adv. Water Resour., 27, 967-980, 2004.

Hou, D., Charles, M., Luo, Y., Toth, Z., Zhu, Y., Krzysztofowicz, R., Lin, Y., Xie, P., Seo, D.-J., Pena, M., and Cui, B.: Climatology-Calibrated Precipitation Analysis

at Fine Scales: Statistical Adjustment of Stage IV toward CPC Gauge-Based Analysis. J. Hydrometeorol., 15, 2542-2557, 2014.

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.

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., 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 **28 Jan** 201**5**), 2013.

Joyce, R., Janowiak, J., Arkin, P., and Xie, X.: 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.: Quantitative precipitation estimation from earth observation satellites. Rainfall: State of the Science, Geophys. Monogr. Ser. http://dx.doi.org/10.1029/2009GM000920, online first, 2010.

Olivier Prat 2/27/2015 11:06 AM
Deleted: 7

Olivier Prat 2/27/2015 11:06 AM
Deleted: Sep

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.

Leeper, R. D., Rennie, J., and Palecki, M. A.: Observational Perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) Network: Temperature and Precipitation Comparison, J. Atmos. Ocean. Tech., in press, 2015.

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

McMillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: rainfall, river discharge, and water quality, Hydrol. Process., 26, 4078-4111, 2012.

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. Oceanic Technol., 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., Prat, O. P., Seo, D.-J., and Habib, E.: Assessment and implications of NCEP Stage IV quantitative precipitation estimates. Wea. Forecast. **Under revision**, 2015.

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

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.

Ochoa, A., Pineda, L., Crespo, P., and Willems, P.: Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific–Andean region of Ecuador and Peru. Hydrol. Earth Syst. Sci., 18, 3179–3193, 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, 2010a.

Prat, O. P., and Barros, A. P.: Ground observations to characterize the spatial gradients and vertical structure of orographic precipitation - Experiments in the inner region of the Great Smoky Mountains, J. Hydrol., 391, 143-158, doi: 10.1016/j.jhydrol.2010.07.013, 2010b.

Olivier Prat 2/27/2015 11:53 AM
Deleted: Submitted.

Olivier Prat 2/27/2015 11:52 AM
Deleted: 4

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., 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., and Nanjundiah, R.: Diurnal-scale signatures of monsoon rainfall over the Indian region from TRMM satellite observations. J. Geophys. Res., 115, D02103. http://dx.doi.org/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.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer,, A., Rudolf, B., and Ziese, M.: GPCC Monitoring Product: Near Real-Time Monthly Land-Surface Precipitation from Rain-Gauges based on SYNOP and CLIMAT data. DOI: 10.5676/DWD_GPCC/MP_M_V4_100; http://dx.doi.org/10.5676/DWD_GPCC/MP_M_V4_100, 2011.

Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre, DWD, 12 pp., available at ftp://ftp-anon.dwd.de/pub/data/gpcc/PDF/GPCC_intro_products_2008.pdf, last access 28 January 2015, 2010.

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. Bull. Am. Meteorol. Soc., 81, 2035-2046, 2000.

Sevruk, B., Ondras, M., and Chvila, B.: The WMO precipitation measurement intercomparison, Atmos. Res., 192, 376-380, 2009.

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. Bull. Am. Meteorol. Soc., 92, 1321-1338, 2011.

LIST OF TABLES

LIST OF FIGURES

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

Figure 3: Scatterplots (black) and Quantile-Quantile (red) plots for annual precipitation derived from PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT when compared to GHCN-D 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 then for rows a), b), and d).

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.

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.

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).

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).

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 Missouri Basin River (MB).

Figure 9: a) Number of rainy days (RR > 0 mm/day), b) Wet Millimeter Days (WMMD), c) Extreme Precipitation Days greater than 2 in day^{-1} (RR > 50.8 mm/day: EPD2), and d) Extreme Precipitation Days greater than 4 in day^{-1} (RR > 101.6 mm/day: EPD4) for Stage IV and TMPA 3B42.

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).

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, 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).

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 **pixel** [YY: red], and successively at the rain gauge only [YN: blue], or at the radar **pixel** only [NY: green] over CONUS (circle) and for the 12 RFCs (bars). Data are for the period 2002-2012.

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.

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

Table 2: Average rain rate (mm/day), 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. **The asterisk indicates that the datasets are statistically different at the 5% significance level with respect to surface observations.**

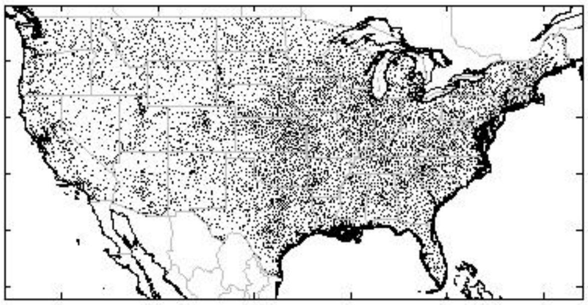
ID	GHCN -D	PRISM	Stage IV	TMPA	
				3B42	3B42RT
CONUS	2.47	2.42 [-2.2|0.98|0.98]*	2.31 [-6.4|0.93|0.93]*	2.52 [+2.1|1.00|0.83]*	2.62 [+6.1|0.99|0.36]*
ABRFC	2.13	2.06 [-3.0|0.97|0.99]	**2.14 [+0.7|1.00|0.96]**	**2.27 [+6.9|1.06|0.96]***	2.40 [+12.5|1.12|0.94]*
CBRFC	0.89	0.90 [+0.7|0.99|0.87]	0.77 [-13.0|0.89|0.86]*	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]*

Table 3: Average rain rate (mm/day) 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. **The asterisk indicates that the datasets are statistically different at the 5% significance level with respect to surface observations.**

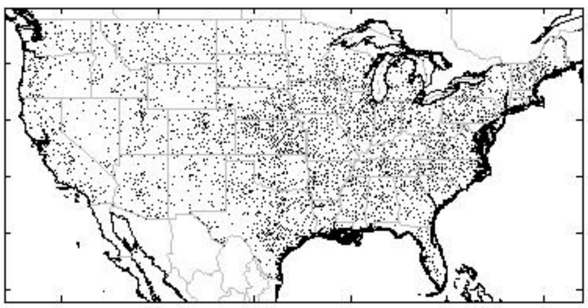
Season	ID	GHCN-D	PRISM	Stage IV	TMPA	
					3B42	3B42RT
DJF	CONUS	2.09	2.03 [-2.8]	1.89 [-9.5]*	2.09 [+0.1]	2.31 [+10.7]*
	ABRFC	1.17	1.14 [-2.4]	**1.15 [-1.5]**	1.28 [+9.2]	1.31 [+11.6]
	CBRFC	1.01	1.01 [-0.2]	0.86 [-14.8]*	0.83 [-17.4]*	1.32 [+30.4]*
	CNRFC	3.09	3.06 [-1.2]	2.72 [-12.2]	***1.92 [-38.0]******	***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]*

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).

a)



b)



c)

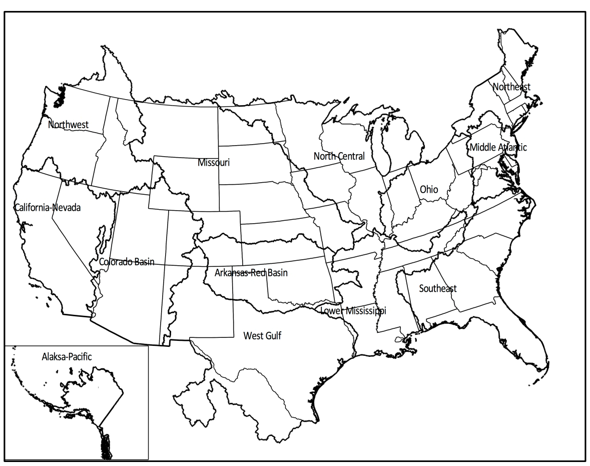


Figure 2: Annual average precipitation derived from: a) PRISM, b) Stage IV, c) TMPA 3B42, and d) TMPA 3B42RT for the period 2002-2012.

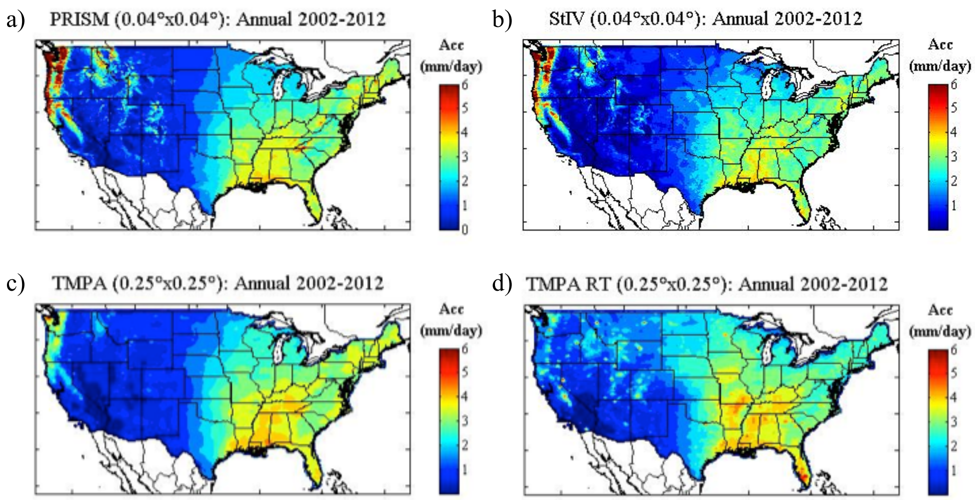


Figure 3: Scatterplots (black) and Quantile-Quantile (red) plots for annual precipitation derived from PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT when compared to GHCN-Daily network for: a) CONUS, b) Lower Mississippi River Basin (LM), c) Missouri Basin River (MB), and d) Northwest (NW) for the period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

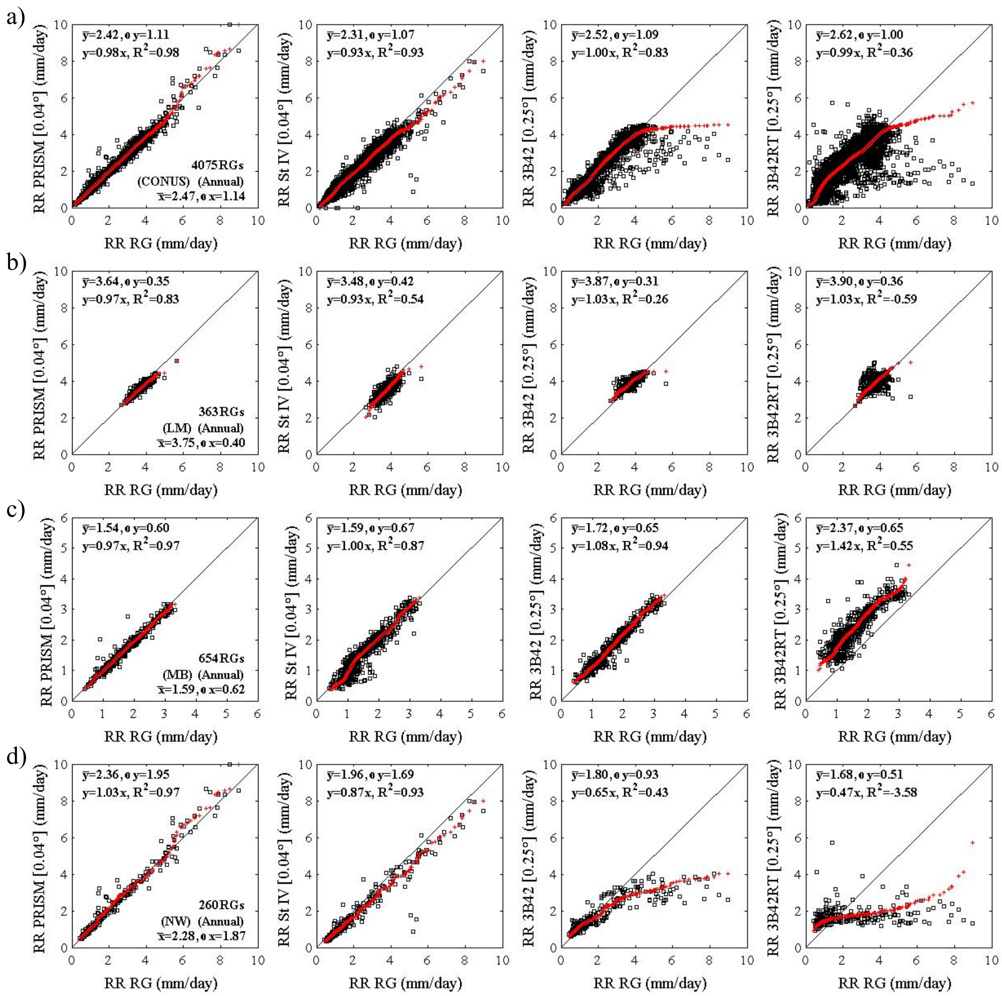


Figure 4: Average annual precipitation derived from GHCN-D, PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for the different RFCs. Differences (%) with respect to GHCN-D surface measurements.

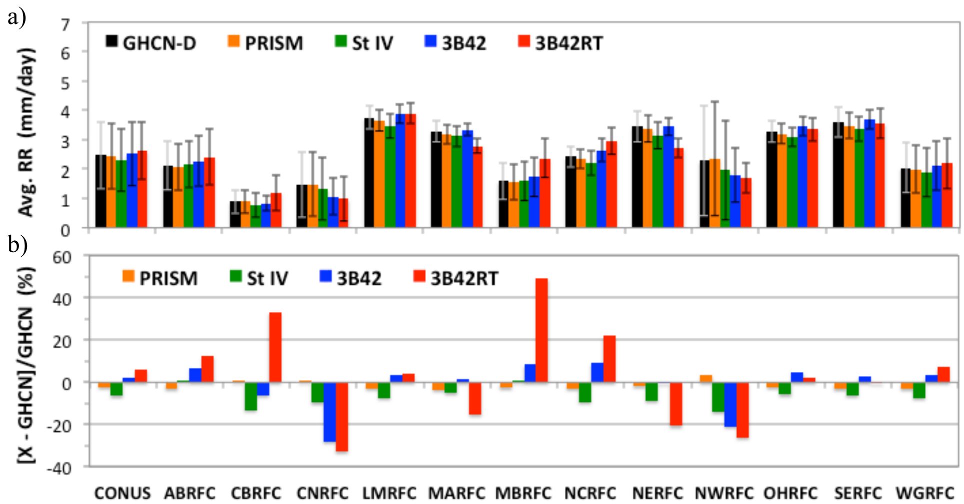


Figure 5: Winter (DJF: Left Column) and summer (JJA: Right Column) precipitation derived from: a) PRISM, b) Stage IV, c) TMPA 3B42, and d) TMPA 3B42RT for the period 2002-2012.

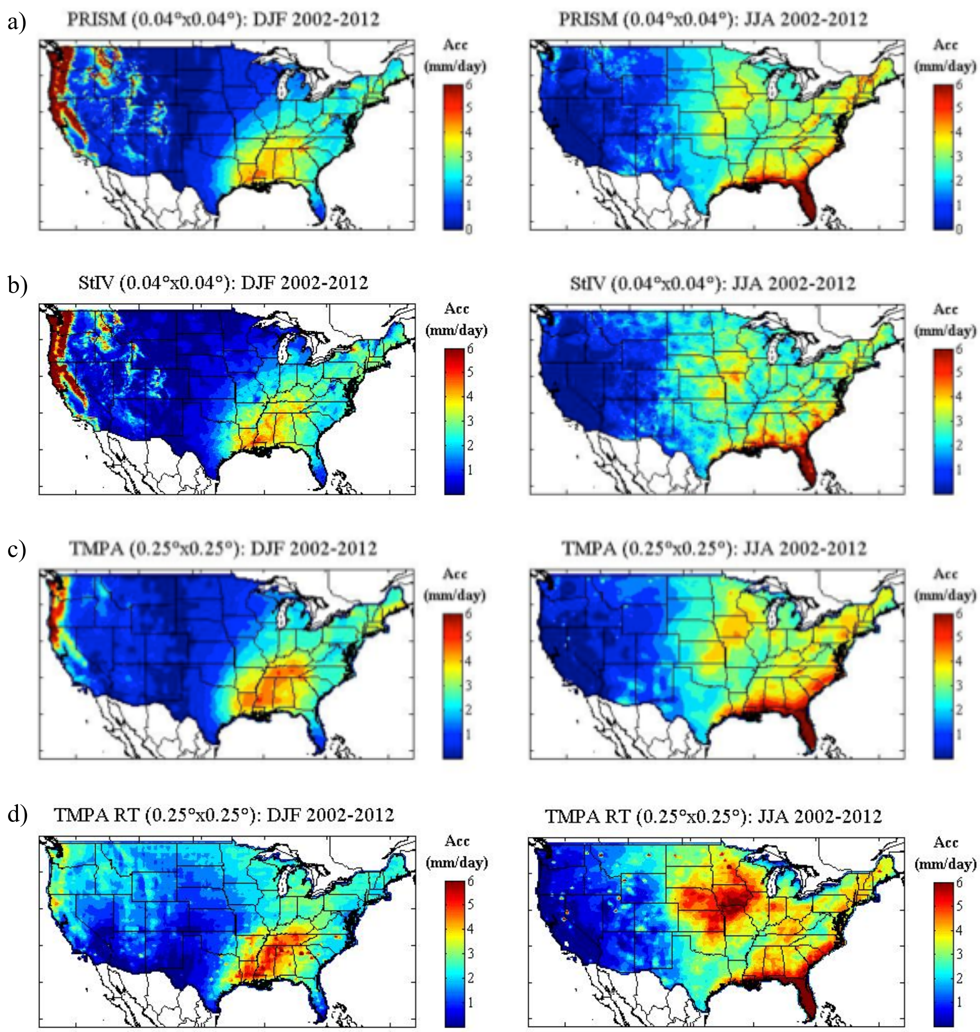


Figure 6: Seasonal rain-rate derived from GHCN-D, PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for each RFC and for: a) Winter (DJF), and b) Summer (JJA). Differences between GHCN-D and PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for: c) Winter (DJF), and d) Summer (JJA).

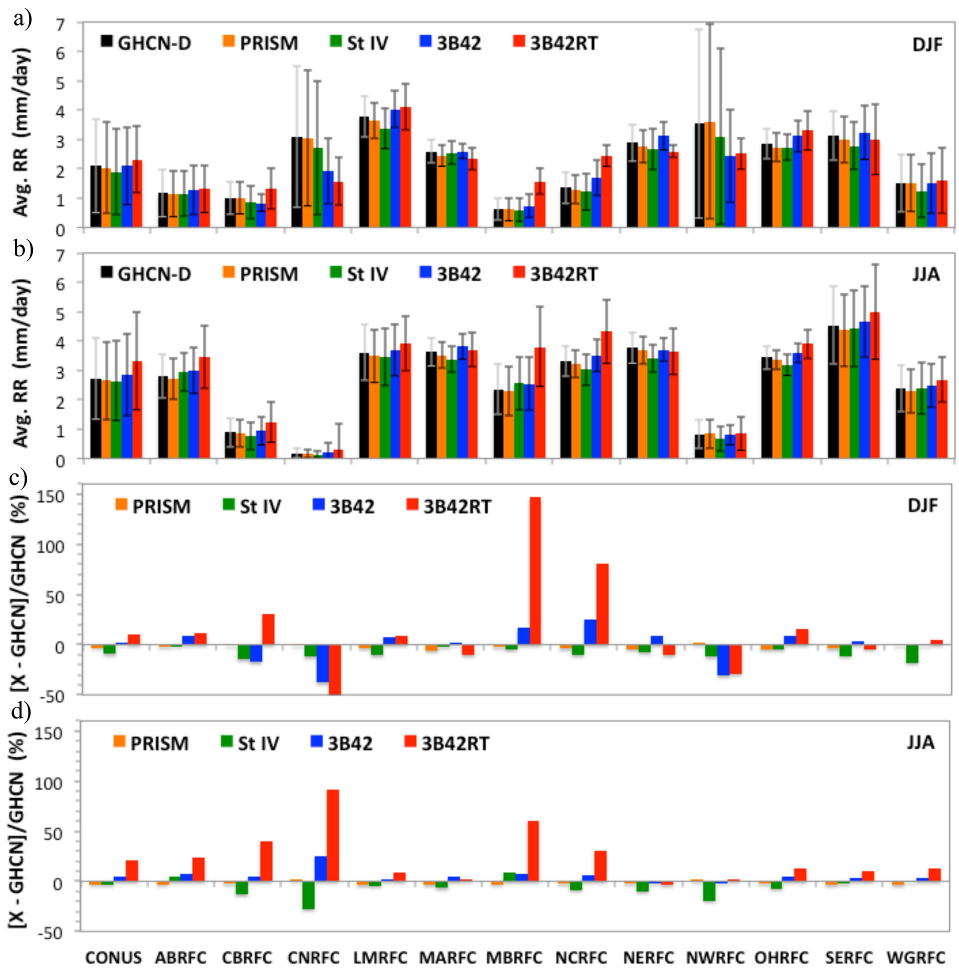


Figure 7: Scatterplots (black) and Q-Q plots (red) for the seasonal rain-rate for PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for: a) Winter (DJF), and b) Summer (JJA) over Northwest (NW).

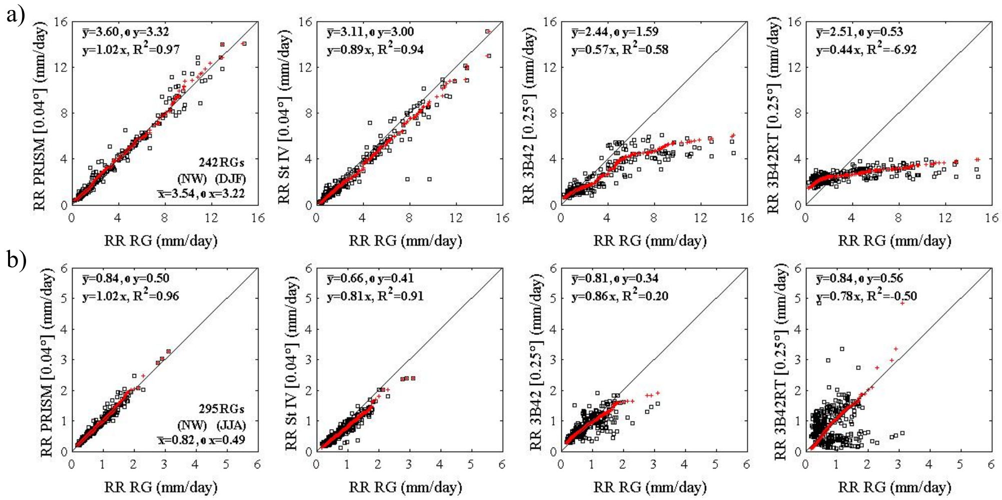


Figure 8: Scatterplots (black) and Q-Q plots (red) for the seasonal rain-rate for PRISM, Stage IV, TMPA 3B42, and TMPA 3B42RT for: a) Winter (DJF), and b) Summer (JJA) over the Missouri Basin River (MB).

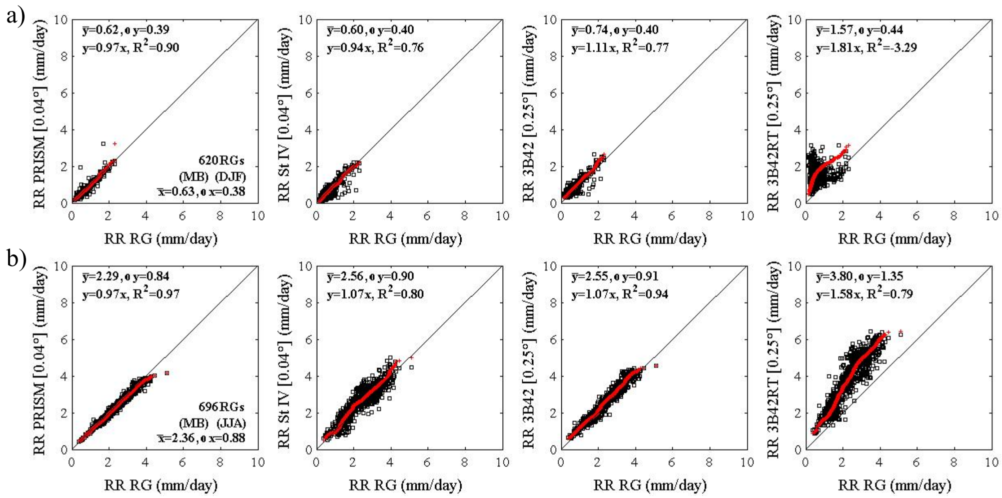


Figure 9: a) Number of rainy days (RR > 0 mm/day), b) Wet Millimeter Days (RR > 17.8 mm/day: WMMD), c) Precipitation Days with accumulation greater than 2 in day^{-1} (RR > 50.8 mm/day: EPD2), and d) Precipitation Days with accumulation greater than 4 in day^{-1} (RR > 101.6 mm/day: EPD4) for GHCN-D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

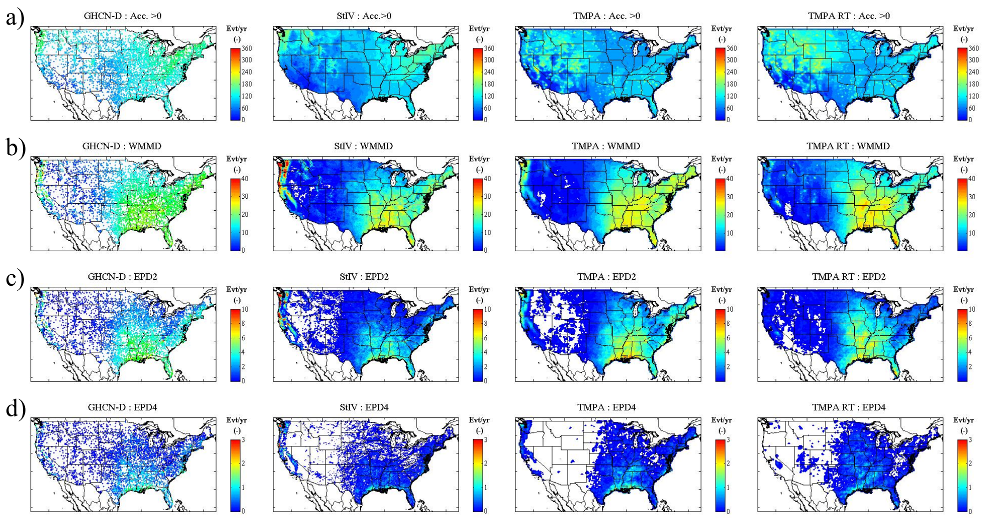


Figure 10: Proportion (%) of stations (GHCN-D) and corresponding pixel (Stage IV, TMPA 3B42, TMPA 3B42RT) experiencing WMMD, EPD2, and EPD4 over CONUS (central figure) and for the 12 River Forecast Centers (border figures).

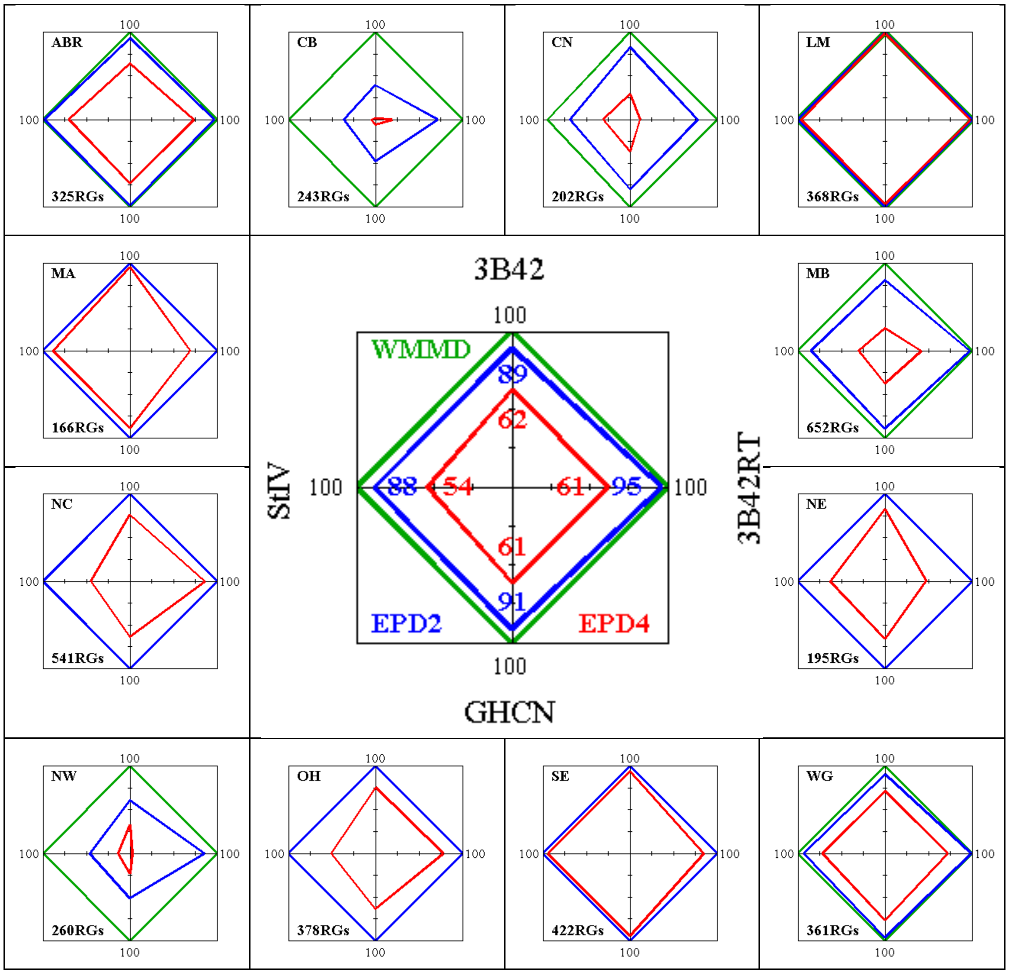


Figure 11: Average number of: a) Rainy days, b) Wet Millimeter Days (WMMD) i.e. precipitation days with accumulation greater than 17.8 mm/day, c) Precipitation Days with accumulation greater than 2 in day^{-1} (EPD2), and d) Precipitation Days with accumulation greater than 4 in day^{-1} (EPD4) for GHCN-D, 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).

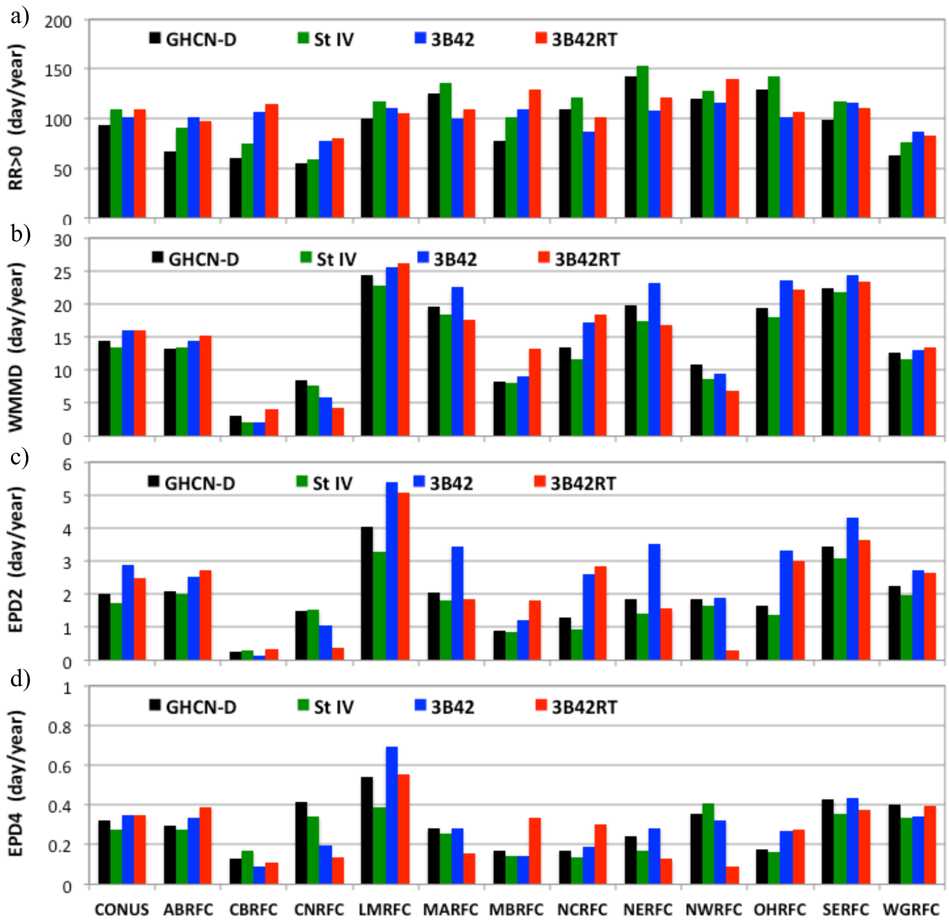


Figure 12: Contingency as a function of the daily threshold selected: a) RR>0, b) RR>WMMD, c) RR>EPD2, and d) RR>EPD4 for rainfall observed simultaneously at the rain gauge and at the radar **pixel** [YY: red], and successively at the rain gauge only [YN: blue], or at the radar **pixel** only [NY: green]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

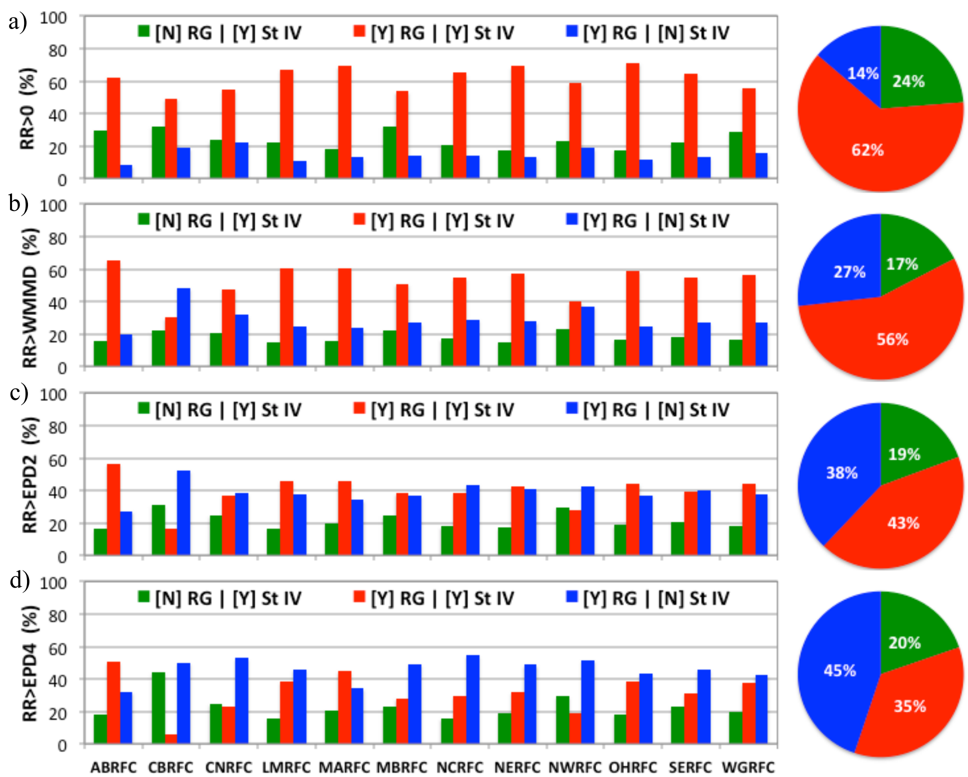
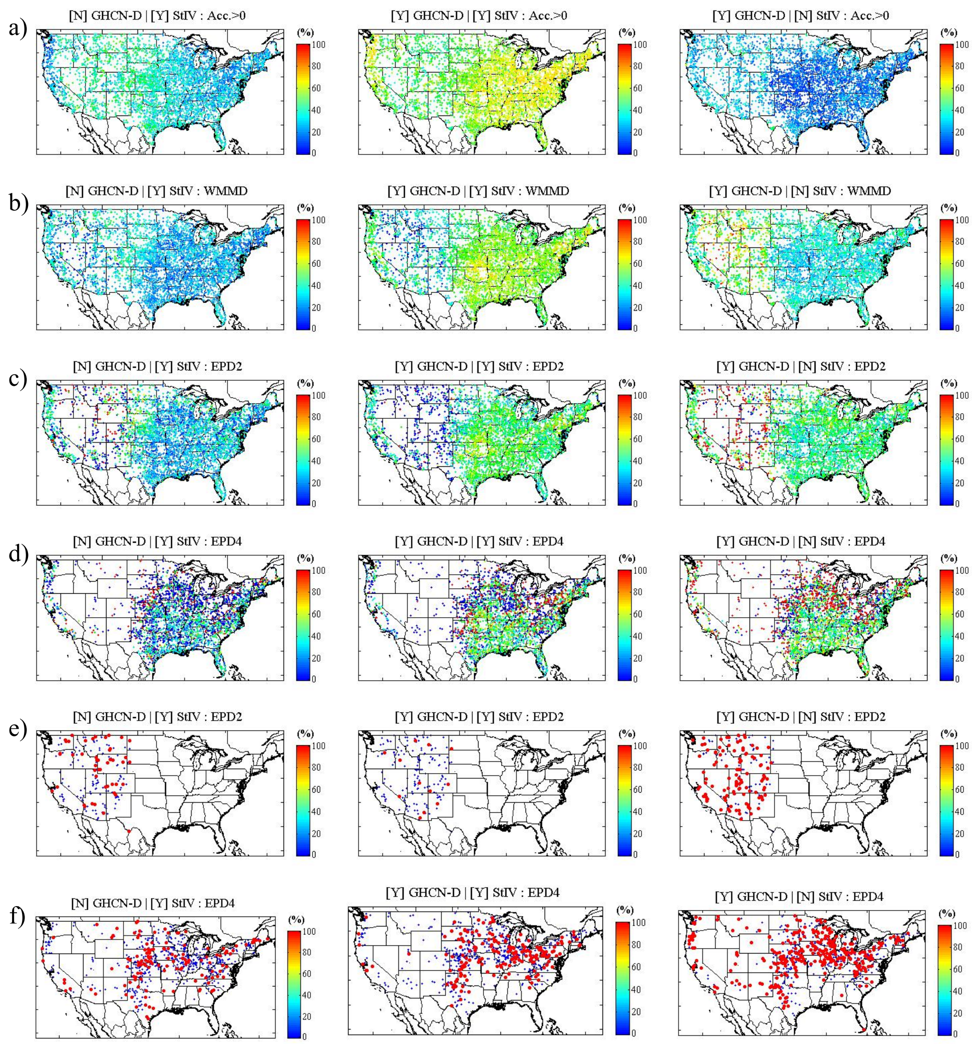


Figure 13: Contingency analysis at the rain gauge site with respect to the daily rainfall accumulation: a) RR > 0, b) RR > WMMD, c) RR > EPD2, and d) RR > EPD4 for rain observed at the radar pixel only (first column), simultaneously at the rain gauge and radar (second column), and at the rain gauge only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



Page 20: [1] Deleted **Olivier Prat** **3/1/15 1:13 PM**

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), the fact that estimates for each individual RFC are processed independently and uses different algorithms that are ultimately combined into a CONUS mosaic (Nelson et al. 2014).

Page 20: [2] Deleted **Olivier Prat** **3/1/15 10:05 AM**

Figure 8 displays the scatterplots over the Missouri Basin River (MB) for winter (Fig. 8a) and summer (Fig. 8b). Once again,

Page 20: [3] Deleted **Olivier Prat** **3/1/15 10:18 AM**

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%).

Page 20: [4] Deleted **Olivier Prat** **3/1/15 10:42 AM**

During winter, the near real time 3B42RT exhibits a strong overestimation (+147%) with a rain rate of 1.57 mm/day to be compared with the 0.64 mm/day measured by GHCN-D (Table 3).

Page 20: [5] Deleted	Olivier Prat	3/1/15 11:09 AM

For summer, 3B42RT displays a severe overestimation (+61%) when compared to surface observations.

Page 20: [6] Deleted	Olivier Prat	3/1/15 11:56 AM

These strong differences are due to uncertainties in PMW retrieval in presence of strong sub-cloud evaporation that could be mistaken for precipitation. Regardless of the season, the