Manuscript prepared for Hydrol. Earth Syst. Sci.
with version 2014/07/29 7.12 Copernicus papers of the LaTeX class copernicus.cls.
Date: 8 July 2015

Characterization of Precipitation Product Errors across the United States using Multiplicative Triple Collocation

Seyed Hamed Alemohammad[1], Kaighin A. McColl[1], Alexandra G. Konings[1], Dara Entekhabi[1], and Ad Stoffelen[2]

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Koninklijk Nederlands Meteorologisch Instituut (KNMI), R&D Satellite Observations, De Bilt, the Netherlands

Correspondence to: Seyed Hamed Alemohammad (hamed_al@mit.edu)

Abstract.

Validation of precipitation estimates from various products is a challenging problem, since the true precipitation is unknown. However, with the increased availability of precipitation estimates from a wide range of instruments (satellite, ground-based radar, and gauge), it is now possible to apply the Triple Collocation (TC) technique to characterize the uncertainties in each of the products. Classical TC takes advantage of three collocated data products of the same variable and estimates the mean squared error of each, without requiring knowledge of the truth. In this study, triplets among NEXRAD-IV, TRMM ~~3B42, GPCP~~ 3B42RT, GPCP 1DD and GPI products are used to quantify the associated spatial error characteristics across a central part of the continental US. Data are aggregated to biweekly accumulations from January 2002 through April 2014 across a 2° × 2° spatial grid. This is the first study of its kind to explore precipitation estimation errors using TC across the United States (US). A multiplicative (logarithmic) error model is incorporated in the original TC formulation to relate the precipitation estimates to the unknown truth. For precipitation application, this is more realistic than the additive error model used in the original TC derivations, which is generally appropriate for existing applications such as in the case of wind vector components and soil moisture comparisons. This study provides error estimates of the precipitation products that can be incorporated into hydrological and meteorological models, especially those used in data assimilation. Physical interpretations of the error fields (related to topography, climate, etc) are explored. The methodology presented in this study could be used to quantify the uncertainties associated with precipitation estimates from each of the constellation of GPM satellites. Such quantification is prerequisite to optimally merging these estimates.

1 Introduction

Precipitation is one of the main drivers of the water cycle; therefore, accurate precipitation estimates are necessary for studying land-atmosphere interactions as well as linkages between the water, energy and carbon cycles. Surface precipitation is also a principal driver of hydrologic models with a wide range of applications. A wide suite of instruments (in-situ and remote sensing) monitor precipitation incident at the Earth surface. Specifically, there has been a great effort during the last two decades to use microwave radar and radiometer instruments on board low-earth orbit satellites to accurately estimate precipitation over large areas. These estimates when combined with infrared based cloud top temperature observations from geostationary satellites provide high spatial and temporal resolution precipitation estimates that are appropriate for hydrological and climatological studies.

However, precipitation estimation is inevitably subject to error. The errors are caused by different factors depending on the measurement instrument. For gauge measurements, the sparse distribution of gauges, environmental conditions such as wind and evaporation, and topography contribute to the errors. For ground-based radars, beam blockages in mountainous regions, the empirical backscatter-rain rate relationship (and the simplifications ~~embededded~~ embedded in their functional form) and clutter are among the sources of error. Lastly, for satellite retrievals (both radiometer and radar), assumptions about the surface emissivity, neglecting evaporation below clouds, and empirical relationships are the driving factors of error.

The new Global Precipitation Measurement (GPM) mission aims to integrate precipitation estimates from a constellation of satellites to provide high spatial and temporal resolution estimates of precipitation over the Earth (Hou et al., 2013). However, successful data integration requires that the errors in each estimate are known. Since the truth is not known, only indirect methods are generally developed to estimate errors.

Several studies investigate and model the uncertainties in remotely-sensed precipitation estimates; however, they all depend on assuming the ground-based (gauge and/or radar) observations or models representing the zero-error precipitation (~~Krajewski et al. (2000); McCollum et al. (2002); Ebert et al. (2007); Su et al. (2008); Sapia~~
among others).

Triple Collocation (TC) provides a platform for quantifying the Root-Mean-Square-Error (RMSE) in three or more products that estimate the same geophysical variable. Developed by Stoffelen (1998), TC takes advantage of at least three spatially and temporally collocated measurements of the variable of interest to solve a system of equations and estimate the error variances of each of the measurements. To make this system of equations determined, some assumptions are built into the technique including zero error cross covariance between different products and zero covariance between errors and truth.

While TC has been used extensively to estimate errors in soil moisture products (Miralles et al., 2010; Dorigo et al., 2010; Parinussa et al., 2011; Anderson et al., 2012; Draper et al., 2013), it has also been successfully applied to other geophysical variables such as ocean wind speed and wave height (Stoffelen, 1998; Janssen et al., 2007; Portabella and Stoffelen, 2009), leaf area index (LAI) (Fang et al., 2012), fraction of absorbed photosynthetically active radiation (FAPAR) (D'Odorico et al., 2014), sea-ice thickness (Scott et al., 2014), atmospheric columnar integrated water vapor (Cimini et al., 2012; Thao et al., 2014), sea surface salinity (Ratheesh et al., 2013), and land water storage (van Dijk et al., 2014).

Roebeling et al. (2012) for the first time apply the TC technique to precipitation products and estimate errors for three precipitation products across Europe. The results show that a gridded gauge product and satellite retrievals (microwave) have TC errors less than 1.0 mm day^{-1} while the European weather radar estimates have errors up to 18 mm day^{-1} in some mountainous regions.

New variants of TC are introduced with wider applications in recent years. McColl et al. (2014) introduce the Extended TC (ETC) that can be used to easily estimate the correlation coefficient between each of the triplets and the unknown truth as well as their RMSEs. ETC is mathematically equivalent to the original TC; however, the ease of calculating the correlation coefficients in ETC provides a different perspective on the performance of each product.

Su et al. (2014) introduce an implementation of instrument variables to reduce the minimum number of products necessary for TC analysis to two. In this framework, the lagged version of one of the two products is used as the third product to conduct the TC analysis (lagged-TC). If the lagged product is sampled at time intervals shorter than the temporal correlation length of the variable of interest, this approach can provide RMSE estimates of two collocated products.

In this study, we estimate the spatial RMSE between triplets of precipitation products across a central part of the US. Unlike Roebeling et al. (2012), we introduce a new logarithmic (multiplicative) error model that is more realistic for precipitation estimates. Moreover, the ETC approach is used in this study to estimate the correlation coefficients for each of the products.

Yilmaz and Crow (2014) present an extensive evaluation of the TC assumptions when applied to soil moisture products. We take a similar approach here, and use rain gauge data to validate the error estimates from TC analysis in a subset of pixels of the study domain. These pixels (located in the state of Oklahoma) have a dense network of rain gauges with a high quality data processing system that enables us to do this evaluation. The results of this evaluation provide a better understanding of the errors in precipitation products estimated by TC.

This paper is organized as following: Section 2 introduces the multiplicative TC analysis. Section 3 reviews the products used in this study. Section 4 presents the results of TC error estimates. Section 5 evaluates the assumptions of TC analysis using gauge data and Section 6 discusses the results and conclusions.

2 Triple Collocation Formulation

In this section, we review the TC formulation and introduce the multiplicative error model. In the multiplicative error model for precipitation, the true precipitation is related to the estimation as:

$$\mathbf{R_i} = a_i \mathbf{T}^{\beta_i} e^{\epsilon_\mathbf{i}} \tag{1}$$

in which $\mathbf{R_i}$ is the precipitation intensity estimate from product i, $\mathbf{T}$ is the true precipitation intensity, a_i is the multiplicative error, β_i is the deformation error and $\epsilon_\mathbf{i}$ is the residual (random) error. The multiplicative error model is used in several studies to investigate the errors associated with precipitation estimates (Hossain and Anagnostou, 2006; Ciach et al., 2007; Villarini et al., 2009; Tian et al., 2013). It is generally concluded that the multiplicative model is more appropriate for quantifying errors in precipitation estimates. Moreover, Tian et al. (2013) present a comparison between the linear and multiplicative error models applied to daily precipitation estimates across the US. They show that the multiplicative model has a better prediction skill and it is applicable to the variable and wide range of daily precipitation values. We also evaluated the joint probability density functions (PDF) of pairs of products to check their spread across different values of precipitation. Results show that PDFs generated from the multiplicative model have better spread compared to the additive model. Therefore, we concluded that for biweekly data it is better to assume the multiplicative model.

In this study, we use the multiplicative model to relate the precipitation estimates to the true value; however, without having the truth or making any assumptions about the distribution of the error, we estimate the RMSE of each estimate. Taking the logarithm of (1), results in:

$$ln(\mathbf{R_i}) = \alpha_i + \beta_i ln(\mathbf{T}) + \epsilon_\mathbf{i} \tag{2}$$

in which, $\alpha_i = ln(a_i)$ is the offset. Defining $\mathbf{r_i} = ln(\mathbf{R_i})$ and $\mathbf{t} = ln(\mathbf{T})$ the equation is simplified to:

$$\mathbf{r_i} = \alpha_i + \beta_i \mathbf{t} + \epsilon_\mathbf{i} \tag{3}$$

This linear equation makes it possible to apply TC to the precipitation data assuming a multiplicative error model. Therefore, log-transformation of the precipitation estimates from all the products is performed in this study and then TC is applied. Assuming there are three collocated estimates of precipitation with zero mean residual errors ($E(\epsilon_\mathbf{i}) = 0$) that are uncorrelated with each other ($Cov(\epsilon_\mathbf{i}, \epsilon_\mathbf{j}) = 0$) and with the true precipitation ($Cov(\epsilon_\mathbf{i}, \mathbf{t}) = 0$), the RMSE of each product can be estimated using the following sets of equations (McColl et al., 2014):

$$\sigma^2_{\mathbf{r_1}} = C_{11} - \frac{C_{12}C_{13}}{C_{23}} \tag{4}$$

$$\sigma^2_{\mathbf{r_2}} = C_{22} - \frac{C_{12}C_{23}}{C_{13}} \tag{5}$$

$$\sigma^2_{\mathbf{r_3}} = C_{33} - \frac{C_{13}C_{23}}{C_{12}} \tag{6}$$

where C_{ij} is the (i,j)th element of the sample covariance matrix between the transformed triplets, and $\sigma_{\mathbf{r_i}}$ is the RMSE of the $\mathbf{r_i}$ product. Equations (4) - (6) estimate the mean-square-error of each product in logarithmic scale. In Section 4, the results of these estimates along with RMSE estimates of $\mathbf{R_i}$ products are presented.

Based on the ETC introduced by McColl et al. (2014), the correlation coefficient between the truth and each of the triplets is:

$$\rho^2_{\mathbf{t},1} = \frac{C_{12}C_{13}}{C_{11}C_{23}} \tag{7}$$

$$\rho^2_{\mathbf{t},2} = \frac{C_{12}C_{23}}{C_{22}C_{13}} \tag{8}$$

$$\rho^2_{\mathbf{t},3} = \frac{C_{13}C_{23}}{C_{33}C_{12}} \tag{9}$$

where $\rho^2_{\mathbf{t},i}$ is the correlation coefficient between the truth and product i in the logarithmic scale (i.e. between $\mathbf{t}$ and $\mathbf{r_i}$). In defining the sign of the $\rho_{\mathbf{t},i}$, it is assumed that the measurements are positively correlated with the truth to overcome sign ambiguity.

3 Study Domain and Data Pre-Processing

Figure 1 shows the analysis domain and the spatial grid used in this study. The study domain ranges from 30° to $40^\circ N$ latitudes and 110° to 80° W longitudes. This region is selected to maximize the overlapping spatial coverage between the data sets that are used here. Major water-bodies (Great Lakes and the Gulf of Mexico) and strong terrain (i.e. Rocky Mountains) are excluded.

Precipitation estimates from ~~four~~ five products NEXRAD-IV, TRMM 3B42RT, TRMM 3B42, GPI and GPCP 1DD are evaluated. NEXRAD-IV is the national mosaicked precipitation estimates from the National Weather Service ground-based WSR-88D radar network ~~(Fulton et al., 1998)~~ (National Center for Environmental Predic

This product is based on merged gauge and radar estimates from 12 river forecast centers across the Continental United States (CONUS) that are mosaicked to a 4km grid over CONUS. The product is available through the National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL; Lin and Mitchell (2005)). Using nearest neighbor sampling, we map this product to a $0.05^\circ \times 0.05^\circ$ latitude-longitude grid. The original NEXRAD-IV (hereafter called NEXRAD) product is hourly accumulation in mm and is available from Jan. 2002 to present.

TRMM ~~3B42~~ 3B42RT is a multi-satellite precipitation estimate from the Tropical Rainfall Measuring Mission (TRMM) together with other low Earth-orbit microwave instruments ~~(Huffman et al., 2007)~~ (Tropical Rainfall Measu

The precipitation estimates from several microwave instruments are calibrated against the merged radar and radiometer precipitation product from TRMM, and then merged to produce a near-global 3-hr precipitation product. The pixels with no microwave instrument observations are filled with measurements from IR instruments on board geostationary satellites that are calibrated using Passive Microwave (PMW) measurements. The TRMM 3B42RT is the real-time version of the product that does not have a gauge correction; however, the TRMM 3B42 ~~(hereafter called TRMM)~~ is a gauge corrected product meaning that the monthly accumulation of estimates in each pixel are calibrated against GPCC gauge product to have similar monthly magnitudes. ~~This product is~~ These two products are available on a $0.25^\circ \times 0.25^\circ$ latitude-longitude grid from Jan. 1998 to present. We use the current V7 of ~~this product~~them.

The GOES Precipitation Index (GPI) is a rainfall retrieval algorithm that only uses cloud-top temperatures from IR-based observations of geostationary satellites to estimate rain rate ~~(Arkin and Meisner, 1987; ?)~~ (Arkin and Meisne The main advantage of this product is that it only uses observations from geostationary satellites that are frequently available across the globe. However, the physics of the precipitation process is not considered in this retrieval algorithm. Therefore, the estimates are only useful in the tropics and warm-season extra-tropics in which most of the precipitation originates from deep convective cloud systems. This product contains daily precipitation rates on a $1^\circ \times 1^\circ$ spatial grid from Oct. 1996 to present.

The Global Precipitation Climatology Project (GPCP) is globally merged daily precipitation rate at $1^\circ \times 1^\circ$ spatial resolution from Oct. 1996 to the present ~~(Huffman et al., 2001)~~ (Huffman et al., 2001, 2013) . This is a merged estimate of precipitation from low earth orbit ~~Passive Microwave (PMW)~~ PMW instruments, the GOES IR-based observations, and surface rain gauge measurements. The merging approach utilizes the higher accuracy of the PMW observations to calibrate the more frequent GOES observations. In this study, V1.2 of the One-Degree Daily (1DD) product of GPCP is used.

The NEXRAD~~and TRMM~~ , TRMM 3B42 and TRMM 3B42RT data are upscaled to a $1^\circ \times 1^\circ$ spatial grid to be consistent with the spatial resolution of the GPI and GPCP 1DD data.

The time domain for this error estimation study is from Jan. 2002 until Apr. 2014. All the data products have complete record within this time window which is more than one decade. Moreover, to generate temporally uncorrelated samples that do not have zero precipitation, the data from each product is temporally aggregated to biweekly values. ~~A large number of zero values~~ Precipitation is a bounded variable and can only take values greater and equal to zero. If the precipitation estimate at a specific time and space is equal to zero; then, the error in that estimate can be from a limited set of numbers (basically any number greater than zero). Therefore, the error is dependent on the measurement (or equivalently the truth). As a result, if we have zero value in the precipitation measurement for all the triplets, the error of each of them is dependent on the measurement; and therefore, on each other. This dependence would violate the assumption that all errors are independent and identically distributed. The error dependence decreases as the measurement value moves away from zero. Among the aggregated data, there are a few percentage of samples that have zero biweekly precipitation accumulation which are removed from the analysis. The percentage of samples with zero value is less than 2% in most of the region other than 8 pixels in the southwest of the region (the driest part of the domain) that have up to 8% of the samples equal to zero. In the accumulation algorithm, any biweekly data with missing hourly or daily measurements is treated as a missing value.

This data aggregation reduces the number of samples across the temporal domain of this study. TC analysis needs enough samples to be able to provide an accurate estimation of the error. Therefore, we combine the estimates from four neighboring 1° × 1° pixels to form data points for the 2° × 2° grids shown in Figure 1. This means measurements taken over each of the four 1° × 1° pixels inside the 2° × 2° pixel are each treated to be measurements over the 2° × 2° pixel, increasing the total number of samples for each 2° × 2° pixel. Under the assumption that the estimated rainfall is statistically homogeneous over each 2° × 2° pixel, we can trade off space and time in this way to increase the number of samples.

In the main analyses of the paper, the four products NEXRAD, TRMM 3B42RT, GPI and GPCP 1DD are used. The TRMM 3B42 is used in Section 5 to show the impact of gauge correction on the estimated error characteristics. Figure 2 shows the climatology of precipitation derived from each of the four products. There is a good agreement between NEXRAD, TRMM ~~and GPCP~~ 3B42RT and GPCP 1DD estimates; however, GPI has a different climatological pattern across the domain. This difference is not unexpected. GPI's retrieval algorithm is very simple and only considers the cloud top temperature; therefore, it is less accurate compared to the other three products that are either based on ground-based radar or have microwave estimates of precipitation combined with IR-based observations.

4 Results of TC Analysis

In this section, we apply the multiplicative TC technique to the precipitation products introduced in Section 3 and present the estimated RMSE and correlation coefficients for each of the products. The four products are grouped to two triplets; Group 1 includes NEXRAD, TRMM 3B42RT and GPI products, and Group 2 includes NEXRAD, TRMM ~~and GPCP~~3B42RT and GPCP 1DD. Similar results were obtained from other triplet combinations (not shown here).

Figures 3 and 4 show the RMSE of each $\mathbf{r_i}$ product in groups 1 and 2 respectively. These figures also show the number of data points (biweekly precipitation measurements) that are used in each pixel to do the TC estimate. Generally there are more than 1000 data points in each pixel. The sharp decline in the number of data points in the pixels in the south west of the study domain is due to the NEXRAD product, which had one of its radar systems repeatedly inactive during 2002 and 2003.

The RMSE reported in these figures is based on a bootstrap analysis. We run 1000 bootstrap simulations (i.e. sampling with replacement from the original data time series) and estimate the RMSE using Equations (4) - (6). The mean of these 1000 RMSE estimates are reported in Figures 3 and 4. Additionally, the standard deviation of these bootstrap estimates is reported in the supplementary materials Figure S1. The standard deviations of RMSE from the bootstrap simulations are one order of magnitude smaller than the RMSE estimate itself and the results are consistent between the two groups. GPI has a more uniform pattern for standard deviation of RMSE compared to NEXRAD, TRMM ~~and GPCP~~ 3B42RT and GPCP 1DD that have the east-west pattern. The standard deviation plots provide a range of confidence on the RMSE estimates from TC analysis. Since the standard deviations are an order of magnitude smaller than the RMSE itself, the mean RMSE from the bootstrap simulations is a reasonable estimate of the RMSE.

The first observation and control check from Figures 3 and 4 is that the RMSE estimates of precipitation from NEXRAD and TRMM 3B42RT in both of the groups are very similar. This shows that the TC analysis is robust and the results are not in general dependent on the choice of triplets. Moreover, TRMM 3B42RT product has a lower RMSE in most of the region.

The RMSE estimates shown in Figures 3 and 4 are in logarithmic scale which are informative and useful if someone is assimilating the products in the logarithmic scale (equivalently using the $\mathbf{r_i}$ products). However, the RMSE estimates of $\mathbf{R_i}$ products in units of precipitation intensity (mm/day in this case) provide another perspective and might be simpler to interpret. Denoting $\mu_{\mathbf{R_i}}$ as the mean of $\mathbf{R_i}$, expansion of (2) using Taylor series results in:

$$ln(\mathbf{R_i}) \approx ln(\mu_{\mathbf{R_i}}) + (\mathbf{R_i} - \mu_{\mathbf{R_i}})\frac{1}{\mu_{\mathbf{R_i}}} \tag{10}$$

Therefore,

$$Var[\mathbf{r_i}] = (\frac{1}{\mu_{\mathbf{R_i}}})^2 Var[(\mathbf{R_i} - \mu_{\mathbf{R_i}})] \tag{11}$$

$$\sigma^2_{\mathbf{r_i}} = (\frac{1}{\mu_{\mathbf{R_i}}})^2 \sigma^2_{\mathbf{R_i}} \tag{12}$$

$$\sigma_{\mathbf{R_i}} = \mu_{\mathbf{R_i}} \sigma_{\mathbf{r_i}} \tag{13}$$

Equation (13) is used to report the RMSE of each of the precipitation product errors after carrying out the TC analysis on the log-transformed products. Figures 5 and 6 show the RMSE of precipitation products in each group in units of mm/day. The standard deviation of these RMSE estimate are also presented in Figure S2 of supplementary materials.

There is again consistency between the results of NEXRAD and TRMM 3B42RT in both groups. ~~Similar to Figure 3 and 4, the~~ The RMSE of the TRMM 3B42RT product in both of the triplets and in majority of the pixels is small compared to the other two products~~and~~ , and it is also relatively small compared to the mean precipitation from climatology maps in Figure 2. NEXRAD has relatively higher RMSE compared to TRMM 3B42RT, but is considerably smaller than GPCP 1DD or GPI.

Comparing the pattern of RMSE in NEXRAD, TRMM ~~and GPCP~~ 3B42RT and GPCP 1DD with the climatology maps (Figure 2), it is clear that the RMSE in each product increases east to west similar to the climatology. This means that in regions with higher mean precipitation rate, the RMSE is higher. This is consistent with other studies that have found that the mean error of precipitation estimates is proportional to the mean precipitation (Tian et al. (2013); Gebregiorgis and Hossain (2014); Tang et al. (2014); Alemohammad et al. (2014), among others).

A recent study by Prat and Nelson (2014) investigates the error of several precipitation products (ground-based radar and microwave instruments) over CONUS by assuming the gauge data as truth.

They mainly characterize the bias in precipitation estimates and evaluate detection of precipitation events at different intensity thresholds and time scales. However, their results show a similar pattern in the error estimates; higher estimation errors for higher mean precipitation.

Figure 7 shows the estimated correlation coefficients between the underlying truth and each precipitation product in the logarithmic scale. Similar to Figures S1 and S2 each column is showing the results of one of the triplet groups. Estimates of ρ^2 for TRMM 3B42RT and NEXRAD products from the two groups are very similar and again shows the robustness of results from the TC technique. Among the products analyzed here, the TRMM 3B42RT product has the highest correlation coefficient with the truth in ~~almost all~~ majority of the pixels~~. NEXRAD also has high correlation with the truth but there is~~ , and NEXRAD is ranked second after TRMM 3B42RT. There is also a pattern that pixels toward the east of the region have higher correlation coefficients ~~in the NEXRAD product. GPCP~~ compared to the west of the region. GPCP 1DD has less correlation with the truth, and it has a similar east-west pattern. GPI exhibits very low correlation coefficients (~ 0.1) toward the west of the region.

The combined and quantitative analyses of the RMSE estimate and the correlation coefficients show that the TRMM 3B42RT product has the best performance among the four products considered here. The RMSE ~~and correlation coefficient for TRMM have little~~ for TRMM 3B42RT has relatively less variations across the domain. This means that the TRMM 3B42RT product has better performance in diverse climatic and geographical conditions. However, the correlation coefficients in TRMM 3B42RT decrease in the west side of the domain. This region is the coldest and snowiest part of the domain and it is covered with snow during the winter. The accuracy of microwave-based precipitation retrievals, which are the input measurements to the TRMM 3B42RT product, are affected by the snow on the ground. Some of the retrieval algorithms for these instruments cannot appropriately distinguish the snow on the ground from the falling precipitation. This phenomenon can contribute to the low correlation coefficient between the TRMM 3B42RT and the truth in the west part of the domain.

The NEXRAD product has a distinct error pattern. Both the RMSE and correlation coefficient of the NEXRAD estimates are small toward the west of the domain. However, comparing the error estimates from NEXRAD with the climatology values reveals that the errors are sometimes on the same order as the climatology toward the west of the domain. This is also revealed by the correlation coefficient values, which have a smaller value in the west side of the domain for NEXRAD. This pattern is consistent with the NEXRAD coverage maps provided by Maddox et al. (2002) that shows the effect of terrain on radar beam blockage in mountainous regions of CONUS. Beam blockage is one of the sources of error in ground-based radar estimates of precipitation in mountainous regions.

The GPI and GPCP 1DD products have, in general, lower quality than TRMM 3B42RT and NEXRAD. They have higher RMSE and lower correlation coefficients with the truth. They both have the east-west pattern in the correlation coefficient; however, the GPI product has a sharper gradient and is poorly correlated with the truth toward the west of the study domain. Precipitation events in this region are mostly driven by frontal systems that generate clouds not necessarily well-correlated to precipitation; therefore, the GPI estimates that are solely based on cloud-top temperature are not well correlated with the truth. GPCP 1DD also uses IR-based observations of the clouds, but those are merged with microwave observations from low earth orbit satellites that are more accurate. Therefore, the resulting correlation coefficients are generally higher, especially in the west side of the study domain. If the analysis was limited to the RMSE estimates, GPI might be considered to be performing uniformly well across the entire domain. But with the correlation coefficients we can clearly see the change in quality of GPI estimates across the domain.

5 Gauge Analysis

In this section, we will review the assumptions that are embedded in TC estimates of RMSE and evaluate them using in-situ gauge data. Gauge data are used a proxy for truth. As mentioned in Section 2, TC assumes zero correlation between errors of the triplets (zero error cross-covariance assumption) and between the errors and the truth (error orthogonality assumption). However, this assumption can be violated in many applications if the retrieval algorithms have similar error structures. Yilmaz and Crow (2014) investigated the assumptions of TC and introduced a decomposition of RMSE derived from TC as following:

$$\sigma^2_{TC_1} = \sigma^2_{TRE_1} + \sigma^2_{LS_1} + \sigma^2_{OE_1} + \sigma^2_{XCE_1} \tag{14}$$

In this equation, $\sigma^2_{TC_1}$ is the error variance of product 1 that is estimated by TC, and $\sigma^2_{TRE_1}$ is the true error variance of product 1 that TC is aiming to estimate. $\sigma^2_{LS_1}$ is the leaked portion of σ^2_T (the variance of the true data), $\sigma^2_{OE_1}$ represents the bias term due to the violation of error orthogonality assumption and $\sigma^2_{XCE_1}$ is the bias term due to the violation of zero error cross-covariance assumption between different products. Note, $\sigma^2_{XCE_1}$ is affected by non-zero error cross covariance between any pair of the products, and it is not only between product 1 and the gauge. Using similar notations as in Section 2, these four elements are defined as:

$$\sigma^2_{TRE_1} = \overline{\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_1}} \tag{15}$$

$$\sigma^2_{LS_1} = (\beta_1 - c_{3|1}\beta_3)(\beta_1 - c_{2|1}\beta_2)\sigma^2_{\mathbf{t}} \tag{16}$$

$$\sigma^2_{OE_1} = (\beta_1 - c_{3|1}\beta_3)(\overline{\mathbf{t}\boldsymbol{\epsilon_1}} - c_{2|1}\overline{\mathbf{t}\boldsymbol{\epsilon_2}}) + (\beta_1 - c_{2|1}\beta_2)(\overline{\mathbf{t}\boldsymbol{\epsilon_1}} - c_{3|1}\overline{\mathbf{t}\boldsymbol{\epsilon_3}}) \tag{17}$$

$$\sigma^2_{XCE_1} = -c_{2|1}\overline{\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_2}} - c_{3|1}\overline{\boldsymbol{\epsilon_1}\boldsymbol{\epsilon_3}} + c_{3|1}c_{2|1}\overline{\boldsymbol{\epsilon_2}\boldsymbol{\epsilon_3}} \tag{18}$$

in which $c_{i|j}$ is the scaling factor of product i assuming product j as the reference and overbar refers to temporal averaging. Equations (15) - (18) indicate the error decomposition for product 1 in the triplet. Similar equations can be derived for other products. Derivations of equations for these decomposition terms using the multiplicative error model is presented in Appendix A. For a detailed explanation on how to estimate different variables in these equations, the reader is referred to Section 2.c of Yilmaz and Crow (2014).

For this evaluation analysis, we need accurate ground based observations in order to avoid errors due to differences in the spatial coverage between the gauges and the other products. The six pixels shown in Figure 1 are selected for this evaluation since they have a dense network of rain gauges. These pixels are located in the state of Oklahoma and the gauge data are retrieved from the Oklahoma Mesonet network. This network provides quality controlled daily precipitation estimates across the state of Oklahoma from an automatic and spatially dense set of rain gauges. We have located the gauges in each of the pixels; each pixel at every time contains at least 12 gauges and some of the pixels have up to 39 monitoring gauges. The daily data from the gauges in each pixel are averaged to estimate the true rain of the pixel and are then accumulated to biweekly values.

It is understood that gauge data also have errors including representativeness error (they are point measurements unlike the other products that provide an average value over each pixel); however, as it is shown in Yilmaz and Crow (2014) (Appendix A) the representativeness error in the gauge measurements causes a positive bias in the TC-based RMSE estimates while the ~~cross correlation error~~ cross-correlation between the errors of different products in each triplet causes a negative bias. Therefore, it is reasonable to assume gauge data as an unbiased estimate of truth. Moreover, in this study the average of estimates from several gauges is used as the unbiased estimate of the truth. The representativeness error of the gauge estimates is basically interpreted as part of the total error variance in the gauge product. However, since the gauge estimates are unbiased estimates of the truth, it can be used a proxy to decompose the error variance estimates from TC technique.

Figure 8 shows the results of error decomposition for the RMSE of the NEXRAD product. These estimates are based on another bootstrap simulation with 1000 samples, with corresponding one standard deviation confidence intervals. This figure shows that the bias caused by the leaked signal and error orthogonality assumption is almost zero in all of the cases. However, the zero error ~~cross-covariance~~ cross covariance assumption is causing significant underestimation in the RMSE estimated by TC. Therefore, the NEXRAD RMSE estimate from TC is a lower bound for the error. Figures S3 - S5 in the supplementary material show similar decomposition of the RMSE in TRMM ~~, GPCP~~ 3B42RT, GPCP 1DD and GPI products across these pixels. These figures also confirm that the violation of the zero cross covariance error leads to underestimation of the true RMSE by TC analysis. The noticeable difference between Figures 8, S3, S4 and S5 is that in Figure S5 that shows the error decomposition of GPI product the contribution of error cross covariance to the total TC estimate is small, and in 4 of the pixels is almost zero. This is consistent with the fact that GPI has a completely different retrieval algorithm and is only based on could top temperature measurements. Therefore, it has less correlation with other products. These results are consistent with the findings in Yilmaz and Crow (2014). Moreover, this analysis shows that similar to the soil moisture data it is appropriate to assume that the errors of precipitation products are not correlated with the truth.

~~The estimates in Figure 8 are based on another bootstrap simulation with 1000 samples, with corresponding one standard deviation confidence intervals.~~ Here, we compare the ranking of the products based on the TC-derived errors and the ones based on the gauge analysis (σ_{TRE}). The goal of this comparison is to show how much the violation of zero error cross covariance impacts the RMSE estimates. In all of the 6 pixels that we conducted the gauge-based analysis, the TRMM 3B42RT and NEXRAD products are ranked 1st and 2nd for the lowest error based on the RMSE from TC, respectively. Moreover, based on the rankings in the gauge-based TC analysis (σ_{TRE} in Figures 8, S3, S4 and S5) in 5 out of the 6 pixels, TRMM 3B42RT has the lowest error, and in 4 out of the 6 NEXRAD has the best error after TRMM 3B42RT. However, GPCP 1DD and GPI rankings are only preserved on 3 out of the 6 pixels. Therefore, in general, we can make the conclusion that the relative rankings for the products with the lowest error remains almost the same but it is hard to make any conclusion about the ranking of the other products. Nevertheless, this is based on only 6 pixels out of the 75 pixels across the whole domain. Therefore, it is not possible to extend this conclusion to the whole study. We can conclude from this comparison that the cross-correlation error can impact the performance ranking of the precipitation products, but the relative impact needs further analysis.

To further evaluate the impact of error cross covariance, we replace the TRMM 3B42RT product with the TRMM 3B42 product and estimate the RMSEs in each triplet again. As it was mentioned in Section 3, the TRMM 3B42 product has a monthly gauge correction in its estimation algorithm. Our hypothesis is that this gauge dependence increases the error cross covariance between different products and will lead to lower RMSE estimates in NEXRAD, TRMM 3B42 and GPCP 1DD (these three have gauge correction in their algorithms) compared to the initial estimate using TRMM 3B42RT. We conducted this analysis and the resulting RMSE estimates from two triplets (NEXRAD, TRMM 3B42, GPI) and (NEXRAD, TRMM 3B42, GPCP 1DD) are presented in Figures S6 and S7. Comparing Figures 5 and 6 with S6 and S7, it is evident that the TC-derived error estimates of NEXRAD, TRMM 3B42 and GPCP 1DD are smaller when using the non real-time version of the TRMM 3B42 product. This further confirms that violation of the zero error cross covariance causes a negative bias in the TC-based RMSE estimates.

6 Conclusions

This study presents, for the first time, error estimates of four precipitation products across a central part of the continental US using Triple Collocation (TC). A multiplicative error model is introduced to TC analysis that is a more realistic error model for precipitation. Furthermore, an extended version of TC is used with which not only the standard deviation of random errors in each product, but the correlation coefficient of each product with respect to an underlying truth are estimated. The results show that the TRMM 3B42RT product is performing relatively better than the other three products.

TRMM 3B42RT has the lowest RMSE across the domain, and the highest correlation coefficient with the underlying truth. Meanwhile, NEXRAD performs relatively poorly in the west side of the study domain that is probably caused by the terrain beam blockage. The performance of the GPCP 1DD and GPI product were lower than that of TRMM 3B42RT and NEXRAD. GPI has significantly lower performance in the west side of the study domain that is likely caused by the simple retrieval algorithm used in this product. Meanwhile, GPI has a reasonably good correlation with underlying the truth in the east side of the domain.

In the second part of the paper, an evaluation of the assumptions built into TC is carried out using surface gauge data as proxy for the truth across selective pixels. These pixels have a dense coverage of in-situ gauges. The results of this evaluation reveal that the TC error estimates underestimate the true error in different products due to a violation of the assumption of zero error cross covariance. Moreover, replacing the TRMM 3B42RT with TRMM 3B42 revealed that the gauge correction in the TRMM 3B42 violates the zero error cross covariance assumption and leads to smaller RMSE estimates. However, the result of RMSE estimates from TC have a lot of potential to be incorporated into data assimilation and data merging algorithms.

Triple Collocation analysis has a lot of potential to be applied to various precipitation products at a wide range of spatial and temporal resolutions. This will provide a better understanding of the true error patterns in different products. Error quantification of precipitation products is a necessity if one aims to merge precipitation estimates from several instruments/models. However, care should be taken in choosing triplets that have zero or small error cross covariance. Otherwise, the error variances will be underestimated.

The multiplicative error model used in this study is shown to be an appropriate choice relative to the additive model. However, it would be beneficial to investigate more complex models that can take into account any higher order dependence of the estimate on the truth. A modification to this study would be to include a gauge-only precipitation product. This would reduce the error cross covariance between the products, since the gauge measurement system is different from the remote-sensing instruments. Although gauge estimates have representativeness error, this error will be part of the total error in the gauge product resulting in higher RMSE values of gauge product. Furthermore, conducting TC analysis on precipitation data with different temporal resolution will provide valuable insight on the performance of different products at different temporal scales. However, this should be carried out with care, as precipitation errors at certain temporal resolutions are highly correlated and are not appropriate for TC analysis.

Appendix A: Error Decomposition

In this section, we derive Equations 15 - 18 starting with the multiplicative error model in logarithmic scale:

$$\mathbf{r_i} = \alpha_i + \beta_i \mathbf{t} + \epsilon_\mathbf{i} \tag{A1}$$

Without loss of generality, we assume $\mathbf{r_i}$ and $\mathbf{t}$ be the anomalies from a climatological mean; then, the model is simplified to:

$$\mathbf{r_i} = \beta_i \mathbf{t} + \epsilon_\mathbf{i} \tag{A2}$$

Choosing product r_1 as the reference, the scaling factors are defined as:

$$c_{2|1} = \frac{\overline{\mathbf{r_1 r_3}}}{\overline{\mathbf{r_2 r_3}}} \tag{A3}$$

$$c_{3|1} = \frac{\overline{\mathbf{r_1 r_2}}}{\overline{\mathbf{r_3 r_2}}} \tag{A4}$$

Therefore, the rescaled data sets are defined as: $\mathbf{r_2}^* = c_{2|1}\mathbf{r_2}$ and $\mathbf{r_3}^* = c_{3|1}\mathbf{r_3}$. Then, TC-based error variance of product 1 is defined as:

$$\sigma^2_{TC_1} = \overline{(\mathbf{r}_1 - \mathbf{r}_3^*)(\mathbf{r}_1 - \mathbf{r}_2^*)} \tag{A5}$$

Inserting r_2^*, r_3^* and (A2) into (A5):

$$\sigma^2_{TC_1} = \overline{[(\beta_1 - c_{3|1}\beta_3)\mathbf{t} + (\epsilon_\mathbf{1} - c_{3|1}\epsilon_\mathbf{3})][(\beta_1 - c_{2|1}\beta_2)\mathbf{t} + (\epsilon_\mathbf{1} - c_{2|1}\epsilon_\mathbf{2})]} \tag{A6}$$

$$\begin{aligned}\sigma^2_{TC_1} &= (\beta_1 - c_{3|1}\beta_3)(\beta_1 - c_{2|1}\beta_2)\sigma^2_\mathbf{t} \\ &+ (\beta_1 - c_{3|1}\beta_3)(\overline{\mathbf{t}\epsilon_\mathbf{1}} - c_{2|1}\overline{\mathbf{t}\epsilon_\mathbf{2}}) + (\beta_1 - c_{2|1}\beta_2)(\overline{\mathbf{t}\epsilon_\mathbf{1}} - c_{3|1}\overline{\mathbf{t}\epsilon_\mathbf{3}}) \\ &+ (\overline{\epsilon_\mathbf{1}\epsilon_\mathbf{1}} - c_{2|1}\overline{\epsilon_\mathbf{1}\epsilon_\mathbf{2}} - c_{3|1}\overline{\epsilon_\mathbf{1}\epsilon_\mathbf{3}} + c_{3|1}c_{2|1}\overline{\epsilon_\mathbf{2}\epsilon_\mathbf{3}})\end{aligned} \tag{A7}$$

Rewriting (A7) as:

$$\sigma^2_{TC_1} = \sigma^2_{TRE_1} + \sigma^2_{LS_1} + \sigma^2_{OE_1} + \sigma^2_{XCE_1} \tag{A8}$$

where:

$$\sigma^2_{TRE_1} = \overline{\epsilon_\mathbf{1}\epsilon_\mathbf{1}} \tag{A9}$$

$$\sigma^2_{LS_1} = (\beta_1 - c_{3|1}\beta_3)(\beta_1 - c_{2|1}\beta_2)\sigma^2_\mathbf{t} \tag{A10}$$

$$\sigma^2_{OE_1} = (\beta_1 - c_{3|1}\beta_3)(\overline{\mathbf{t}\epsilon_\mathbf{1}} - c_{2|1}\overline{\mathbf{t}\epsilon_\mathbf{2}}) + (\beta_1 - c_{2|1}\beta_2)(\overline{\mathbf{t}\epsilon_\mathbf{1}} - c_{3|1}\overline{\mathbf{t}\epsilon_\mathbf{3}}) \tag{A11}$$

$$\sigma^2_{XCE_1} = -c_{2|1}\overline{\epsilon_\mathbf{1}\epsilon_\mathbf{2}} - c_{3|1}\overline{\epsilon_\mathbf{1}\epsilon_\mathbf{3}} + c_{3|1}c_{2|1}\overline{\epsilon_\mathbf{2}\epsilon_\mathbf{3}} \tag{A12}$$

Equations (A9) - (A12) are the same as (15) - (18) that are used to decompose the RMSE estimates of TC analysis.

Acknowledgements. The authors wish to thank Dr. Wade Crow and another anonymous reviewer for their constructive feedback that led to improvements in this paper. The authors also thank all the producers and distributors of the data used in this study. The TRMM 3B42 and TRMM 3B42RT data used in this study were acquired as part of the ~~NASA's~~ NASA Earth-Sun System Division and archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). The GPCP 1DD data were provided by the NASA/Goddard Space Flight Center's Mesoscale Atmospheric Processes Laboratory, which develops and computes the 1DD as a contribution to the GEWEX Global Precipitation Climatology Project. The GPI data are produced by science investigators, Drs. Phillip Arkin and John Janowiak of the Climate Analysis Center, NOAA, Washington, D.C., and distributed by the Distributed Active Archive Center (Code 610.2) at the Goddard Space Flight Center, Greenbelt, MD, 20771. The Oklahoma Mesonet data are provided courtesy of the Oklahoma Mesonet, a cooperative venture between Oklahoma State University and The University of Oklahoma and supported by the taxpayers of Oklahoma.

References

Alemohammad, S. H., Entekhabi, D., and McLaughlin, D. B.: Evaluation of Long-Term SSM/I-Based Precipitation Records over Land, J. Hydrometeor, 15, 2012–2029, doi:10.1175/JHM-D-13-0171.1, 2014.

Alemohammad, S. H., McLaughlin, D. B., and Entekhabi, D.: Quantifying Precipitation Uncertainty for Land Data Assimilation Applications, Mon. Wea. Rev., pp. –, doi:10.1175/MWR-D-14-00337.1, http://dx.doi.org/10.1175/MWR-D-14-00337.1, 2015.

Anagnostou, E., Maggioni, V., Nikolopoulos, E., Meskele, T., Hossain, F., and Papadopoulos, A.: Benchmarking High-Resolution Global Satellite Rainfall Products to Radar and Rain-Gauge Rainfall Estimates, Geoscience and Remote Sensing, IEEE Transactions on, 48, 1667–1683, doi:10.1109/TGRS.2009.2034736, 2010.

Anderson, W. B., Zaitchik, B. F., Hain, C. R., Anderson, M. C., Yilmaz, M. T., Mecikalski, J., and Schultz, L.: Towards an integrated soil moisture drought monitor for East Africa, Hydrology and Earth System Sciences, 16, 2893–2913, doi:10.5194/hess-16-2893-2012, 2012.

Arkin, P. and Janowiak, J.: Global Precipitation Index (GPI), Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC).

Arkin, P. A. and Meisner, B. N.: The Relationship between Large-Scale Convective Rainfall and Cold Cloud over the Western Hemisphere during 1982-84, Mon. Wea. Rev., 115, 51–74, doi:10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2, 1987.

Chen, S., Kirstetter, P. E., Hong, Y., Gourley, J. J., Tian, Y. D., Qi, Y. C., Cao, Q., Zhang, J., Howard, K., Hu, J. J., and Xue, X. W.: Evaluation of Spatial Errors of Precipitation Rates and Types from TRMM Spaceborne Radar over the Southern CONUS, J. Hydrometeor, 14, 1884–1896, doi:10.1175/JHM-D-13-027.1, 2013.

Ciach, G. J., Krajewski, W. F., and Villarini, G.: Product-Error-Driven Uncertainty Model for Probabilistic Quantitative Precipitation Estimation with NEXRAD Data, J. Hydrometeor, 8, 1325–1347, doi:10.1175/2007JHM814.1, 2007.

Cimini, D., Pierdicca, N., Pichelli, E., Ferretti, R., Mattioli, V., Bonafoni, S., Montopoli, M., and Perissin, D.: On the accuracy of integrated water vapor observations and the potential for mitigating electromagnetic path delay error in InSAR, Atmospheric Measurement Techniques, 5, 1015–1030, doi:10.5194/amt-5-1015-2012, 2012.

D'Odorico, P., Gonsamo, A., Pinty, B., Gobron, N., Coops, N., Mendez, E., and Schaepman, M. E.: Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe, Remote Sensing of Environment, 142, 141–154, doi:10.1016/j.rse.2013.12.005, 2014.

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture datasets, Hydrology and Earth System Sciences, 14, 2605–2616, doi:10.5194/hess-14-2605-2010, 2010.

Draper, C., Reichle, R., de Jeu, R., Naeimi, V., Parinussa, R., and Wagner, W.: Estimating root mean square errors in remotely sensed soil moisture over continental scale domains, Remote Sensing of Environment, 137, 288 – 298, doi:http://dx.doi.org/10.1016/j.rse.2013.06.013, 2013.

Ebert, E. E., Janowiak, J. E., and Kidd, C.: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models, Bull. Amer. Meteor. Soc., 88, 47–64, doi:10.1175/BAMS-88-1-47, 2007.

Fang, H., Wei, S., Jiang, C., and Scipal, K.: Theoretical uncertainty analysis of global MODIS, CYCLOPES, and {GLOBCARBON} {LAI} products using a triple collocation method, Remote Sensing of Environment, 124, 610 – 621, doi:http://dx.doi.org/10.1016/j.rse.2012.06.013, 2012.

Fulton, R. A., Breidenbach, J. P., Seo, D.-J., Miller, D. A., and O'Bannon, T.: The WSR-88D Rainfall Algorithm, Wea. Forecasting, 13, 377–395, doi:10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2, 1998.

Gebregiorgis, A. and Hossain, F.: Estimation of Satellite Rainfall Error Variance Using Readily Available Geophysical Features, Geoscience and Remote Sensing, IEEE Transactions on, 52, 288–304, doi:10.1109/TGRS.2013.2238636, 2014.

Gebregiorgis, A. S. and Hossain, F.: How well can we estimate error variance of satellite precipitation data around the world?, Atmospheric Research, 154, 39 – 59, doi:http://dx.doi.org/10.1016/j.atmosres.2014.11.005, 2015.

Habib, E., Haile, A. T., Tian, Y., and Joyce, R. J.: Evaluation of the High-Resolution CMORPH Satellite Rainfall Product Using Dense Rain Gauge Observations and Radar-Based Estimates, J. Hydrometeor, 13, 1784–1798, doi:10.1175/JHM-D-12-017.1, 2012.

Hossain, F. and Anagnostou, E. N.: A Two-Dimensional Satellite Rainfall Error Model, IEEE Transactions of Geoscience and Remote Sensing, 44, 1511–1522, doi:10.1109/TGRS.2005.863866, 2006.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The Global Precipitation Measurement Mission, Bull. Amer. Meteor. Soc., 95, 701–722, doi:10.1175/BAMS-D-13-00164.1, 2013.

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations, J. Hydrometeor, 2, 36–50, doi:10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2, 2001.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, J. Hydrometeor, 8, 38–55, doi:10.1175/JHM560.1, 2007.

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: GPCP 1-Degree Daily Combination (Version 1.2), Goddard Space Flight Center's Mesoscale Atmospheric Processes Laboratory, http://precip.gsfc.nasa.gov/gpcp_daily_comb.html, 2013.

Janssen, P. A. E. M., Abdalla, S., Hersbach, H., and Bidlot, J.-R.: Error Estimation of Buoy, Satellite, and Model Wave Height Data, J. Atmos. Oceanic Technol., 24, 1665–1677, doi:10.1175/JTECH2069.1, 2007.

Kirstetter, P.-E., Hong, Y., Gourley, J. J., Schwaller, M., Petersen, W., and Zhang, J.: Comparison of TRMM 2A25 Products, Version 6 and Version 7, with NOAA/NSSL Ground Radar-Based National Mosaic QPE, J. Hydrometeor, 14, 661–669, doi:10.1175/JHM-D-12-030.1, 2012.

Kirstetter, P.-E., Viltard, N., and Gosset, M.: An error model for instantaneous satellite rainfall estimates: evaluation of BRAIN-TMI over West Africa, Quarterly Journal of the Royal Meteorological Society, 139, 894–911, doi:10.1002/qj.1964, 2013.

Krajewski, W. F., Ciach, G. J., McCollum, J. R., and Bacotiu, C.: Initial Validation of the Global Precipitation Climatology Project Monthly Rainfall over the United States, J. Appl. Meteor., 39, 1071–1086, doi:10.1175/1520-0450(2000)039<1071:IVOTGP>2.0.CO;2, 2000.

Lin, Y. and Mitchell, K. E.: The NCEP Stage II/IV hourly precipitation analyses: development and applications, in: 19th Conf. on Hydrology, American Meteorological Society, San Diego, CA, p. Paper 1.2, http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/refs/stage2-4.19hydro.pdf, 2005.

Maddox, R. A., Zhang, J., Gourley, J. J., and Howard, K. W.: Weather Radar Coverage over the Contiguous United States, Wea. Forecasting, 17, 927–934, doi:10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2, 2002.

Maggioni, V., Sapiano, M. R. P., Adler, R. F., Tian, Y., and Huffman, G. J.: An Error Model for Uncertainty Quantification in High-Time-Resolution Precipitation Products, J. Hydrometeor, 15, 1274–1292, doi:10.1175/JHM-D-13-0112.1, 2014.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, Geophysical Research Letters, 41, 6229–6236, doi:10.1002/2014GL061322, 2014.

McCollum, J. R., Krajewski, W. F., Ferraro, R. R., and Ba, M. B.: Evaluation of Biases of Satellite Rainfall Estimation Algorithms over the Continental United States, J. Appl. Meteor., 41, 1065–1080, doi:10.1175/1520-0450(2002)041<1065:EOBOSR>2.0.CO;2, 2002.

Miralles, D. G., Crow, W. T., and Cosh, M. H.: Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations, J. Hydrometeor, 11, 1423–1429, doi:10.1175/2010JHM1285.1, 2010.

National Center for Environmental Prediction: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data, National Center for Atmospheric Research, Earth Observing Laboratory, http://data.eol.ucar.edu/codiac/dss/id=21.093, 2005.

Parinussa, R. M., Holmes, T. R. H., Yilmaz, M. T., and Crow, W. T.: The impact of land surface temperature on soil moisture anomaly detection from passive microwave observations, Hydrology and Earth System Sciences, 15, 3135–3151, doi:10.5194/hess-15-3135-2011, 2011.

Portabella, M. and Stoffelen, A.: On Scatterometer Ocean Stress, J. Atmos. Oceanic Technol., 26, 368–382, doi:10.1175/2008JTECHO578.1, 2009.

Prat, O. P. and Nelson, B. R.: Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge datasets (2002–2012), Hydrology and Earth System Sciences Discussions, 11, 11 489–11 531, doi:10.5194/hessd-11-11489-2014, 2014.

Ratheesh, S., Mankad, B., Basu, S., Kumar, R., and Sharma, R.: Assessment of Satellite-Derived Sea Surface Salinity in the Indian Ocean, Geoscience and Remote Sensing Letters, IEEE, 10, 428–431, doi:10.1109/LGRS.2012.2207943, 2013.

Roebeling, R. A., Wolters, E. L. A., Meirink, J. F., and Leijnse, H.: Triple Collocation of Summer Precipitation Retrievals from SEVIRI over Europe with Gridded Rain Gauge and Weather Radar Data, J. Hydrometeor, 13, 1552–1566, doi:10.1175/JHM-D-11-089.1, 2012.

Salio, P., Hobouchian, M. P., Skabar, Y. G., and Vila, D.: Evaluation of High-Resolution Satellite Precipitation Estimates over Southern South America using a Dense Rain Gauge Network, Atmospheric Research, doi:10.1016/j.atmosres.2014.11.017, 2014.

Sapiano, M. R. P. and Arkin, P. A.: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data, J. Hydrometeor, 10, 149–166, doi:10.1175/2008JHM1052.1, 2009.

Scott, K., Buehner, M., and Carrieres, T.: An Assessment of Sea-Ice Thickness Along the Labrador Coast From AMSR-E and MODIS Data for Operational Data Assimilation, Geoscience and Remote Sensing, IEEE Transactions on, 52, 2726–2737, doi:10.1109/TGRS.2013.2265091, 2014.

Seyyedi, H.: Performance Assessment of Satellite Rainfall Products for Hydrologic Modeling, Ph.D. thesis, University of Connecticut, http://digitalcommons.uconn.edu/dissertations/335, 2014.

Stampoulis, D. and Anagnostou, E. N.: Evaluation of Global Satellite Rainfall Products over Continental Europe, J. Hydrometeor, 13, 588–603, doi:10.1175/JHM-D-11-086.1, 2012.

Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, Journal of Geophysical Research: Oceans, 103, 7755–7766, doi:10.1029/97JC03180, 1998.

Su, C.-H., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: Applications to soil moisture monitoring, Journal of Geophysical Research: Atmospheres, 119, 6419–6439, doi:10.1002/2013JD021043, 2014.

Su, F., Hong, Y., and Lettenmaier, D. P.: Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin, J. Hydrometeor, 9, 622–640, doi:10.1175/2007JHM944.1, 2008.

Tang, L., Tian, Y., and Lin, X.: Validation of precipitation retrievals over land from satellite-based passive microwave sensors, Journal of Geophysical Research: Atmospheres, 119, 4546–4567, doi:10.1002/2013JD020933, 2014.

Thao, S., Eymard, L., Obligis, E., and Picard, B.: Trend and Variability of the Atmospheric Water Vapor: A Mean Sea Level Issue, J. Atmos. Oceanic Technol., 31, 1881–1901, doi:10.1175/JTECH-D-13-00157.1, 2014.

Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., Hsu, K.-l., Turk, F. J., Garcia, M., and Zeng, J.: Component analysis of errors in satellite-based precipitation estimates, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2009JD011949, 2009.

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., and Wu, H.: Modeling errors in daily precipitation measurements: Additive or multiplicative?, Geophysical Research Letters, 40, 2060–2065, doi:10.1002/grl.50320, 2013.

Tropical Rainfall Measurement Mission Project (TRMM): TRMM 3-Hourly 0.25 deg. TRMM and Others Rainfall Estimate Data, version 7, Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC), http://disc.sci.gsfc.nasa.gov/datacollection/TRMM_3B42_V7.html.

van Dijk, A. I. J. M., Renzullo, L. J., Wada, Y., and Tregoning, P.: A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble, Hydrology and Earth System Sciences, 18, 2955–2973, doi:10.5194/hess-18-2955-2014, 2014.

Vila, D. A., de Goncalves, L. G. G., Toll, D. L., and Rozante, J. R.: Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America, J. Hydrometeor, 10, 533–543, doi:10.1175/2008JHM1048.1, 2009.

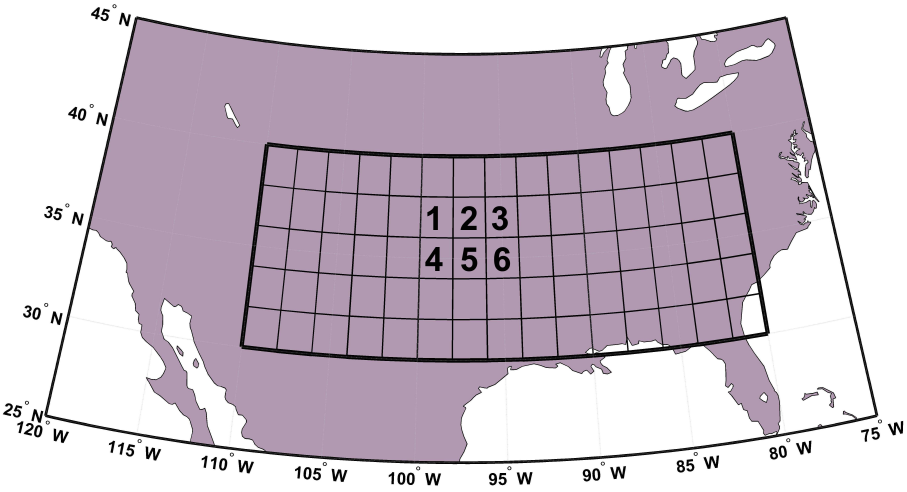
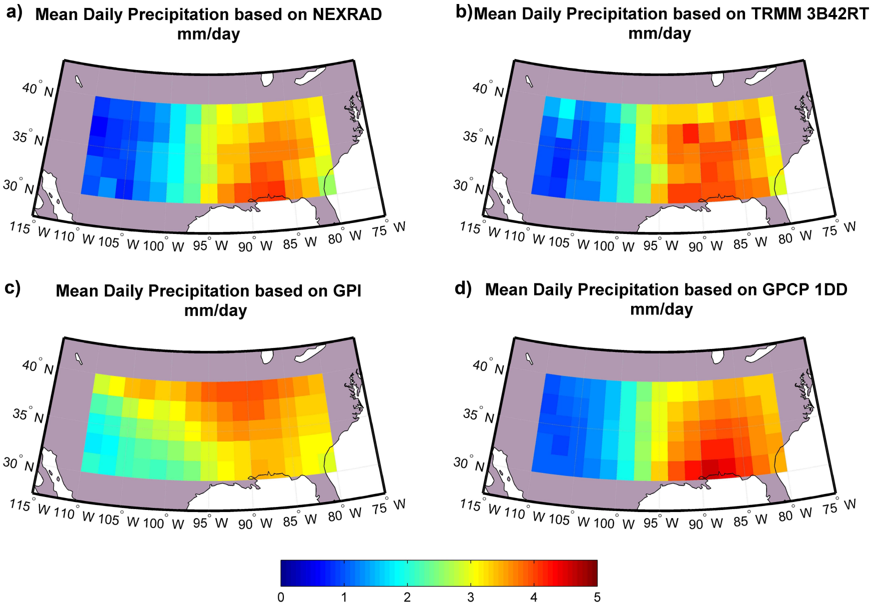


Figure 1. Study Domain. The six numbered pixels are used in Section 5 for evaluation of TC assumptions in estimating RMSE.


a) Mean Daily Precipitation based on NEXRAD mm/day

b) Mean Daily Precipitation based on TRMM 3B42RT mm/day

c) Mean Daily Precipitation based on GPI mm/day

d) Mean Daily Precipitation based on GPCP 1DD mm/day

Figure 2. Climatology of precipitation across the study domain from each of the products.

Villarini, G., Krajewski, W. F., Ciach, G. J., and Zimmerman, D. L.: Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates, Water Resources Research, 45, doi:10.1029/2008WR006946, 2009.

Yilmaz, M. T. and Crow, W. T.: Evaluation of Assumptions in Soil Moisture Triple Collocation Analysis, J. Hydrometeor, 15, 1293–1302, doi:10.1175/JHM-D-13-0158.1, 2014.

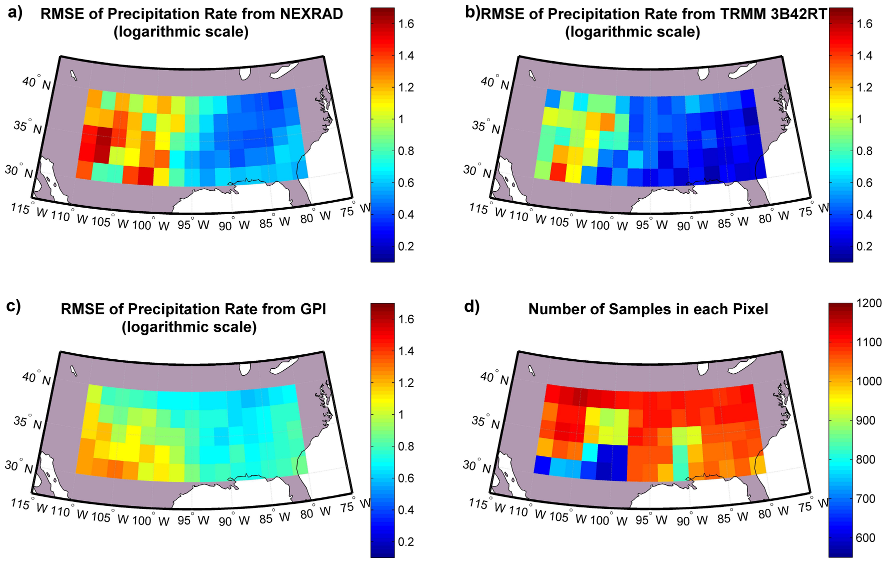


Figure 3. RMSE of the precipitation rate in logarithmic scale estimated from TC using triplets in group 1; a) NEXRAD, b) TRMM 3B42RT, c) GPI. Panel d) shows the number of data points (biweekly measurements) in each pixel that are used for error estimation in TC analysis.

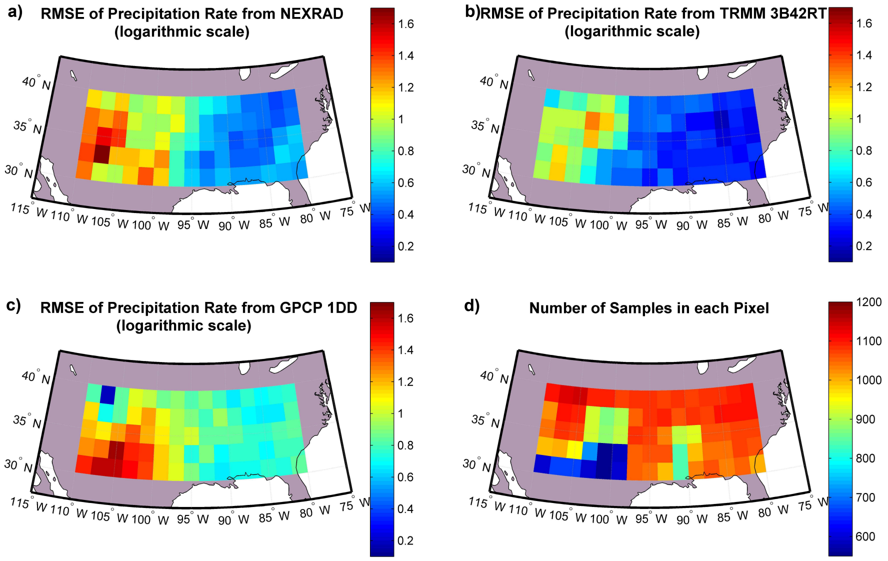


Figure 4. RMSE of the precipitation rate in logarithmic scale estimated from TC using triplets in group 2; a) NEXRAD, b) TRMM 3B42RT, c) ~~GPI~~GPCP 1DD . Panel d) shows the number of data points (biweekly measurements) in each pixel that are used for error estimation in TC analysis.

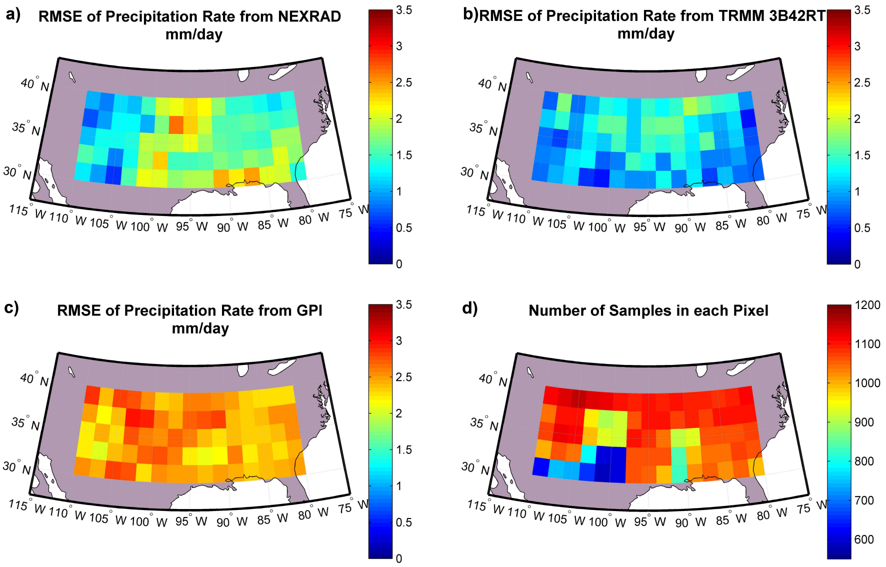


Figure 5. RMSE of the precipitation rate estimated from TC using triplets in group 1; a) NEXRAD, b) TRMM 3B42RT, c) GPI. Panel d) shows the number of data points (biweekly measurements) in each pixel that are used for error estimation in TC analysis.

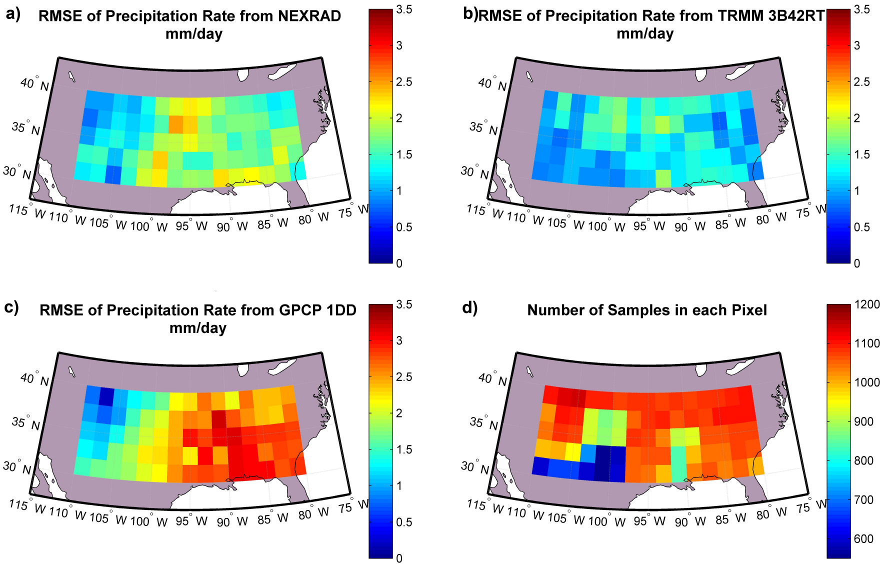


Figure 6. RMSE of the precipitation rate estimated from TC using triplets in group 2; a) NEXRAD, b) TRMM 3B42RT, c) GPCP 1DD. Panel d) shows the number of ~~datapoints~~ data points (biweekly measurements) in each pixel that are used for error estimation in TC analysis.

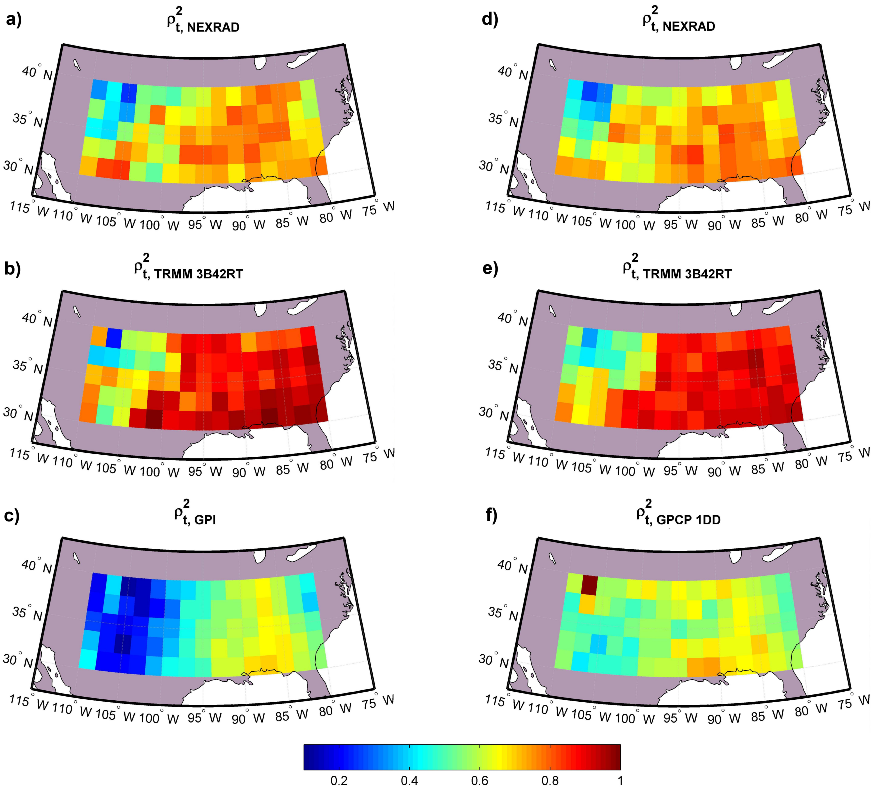


Figure 7. Correlation coefficient between the truth and each precipitation product. The left column shows the results for triplets in group 1, and the right column shows the results for triplets in group 2.

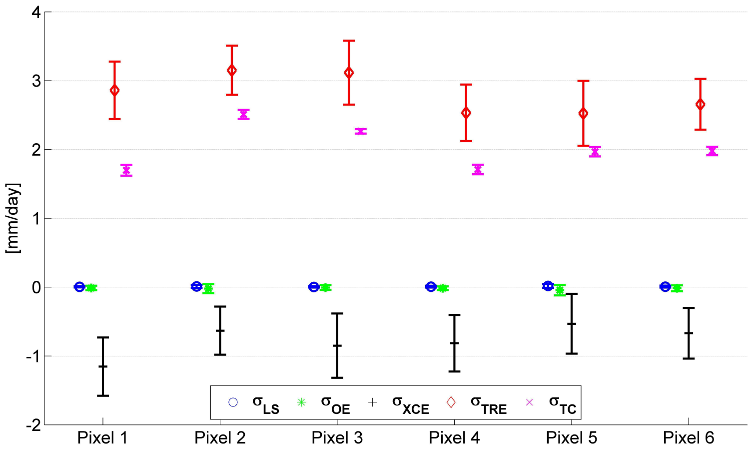


Figure 8. Decomposition of TC-based estimate of RMSE in the NEXRAD product across the six pixels shown in Figure 1. Error bars show one standard deviation of the estimates from a bootstrap run with 100 samples.

Hydrol. Earth Syst. Sci. Discuss., 12, C1978–C1981, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C1978/2015/

Hydrology and Earth System Sciences
Discussions
Open Access

Interactive comment on "Characterization of precipitation product errors across the US using multiplicative Triple Collocation" *by* S. H. Alemohammad et al.

S. H. Alemohammad et al.

hamed_al@mit.edu

Received and published: 6 June 2015

We thank referee #1 for their positive and insightful comments. Here, we respond to the general and specific comments included in their review:

General Comments:

- The authors provide a creative and original study of the errors of several "standard" precipitation data sets using the Triple Collocation approach. Critically, this allows them to use the radar analyses without having to assume that they are exact. It also raises the interesting question of what the result would be if the

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

gauges used in Section 5 were entered as yet another dataset in the Triple Collocation study (obviously, only for the subset of boxes that have gauge data). How close would they turn out to be to the unknown true precipitation?

Response/Action: *We evaluated all possible triplets of gauge with two other products among the four products we have in this study. The results show that the gauge is in most of the cases the product with lowest RMSE, and in some others the second one with RMSEs close to the first product. This result is from the six pixels that we have dense gauge measurements; therefore, it should be interpreted carefully and may not be transferable to the whole domain. This is addressed in the supplementary material, Figures S3-S5, which was submitted together with the initial submission.*

- The fundamental assumption is that the errors are multiplicative. The background literature tends to advocate this approach for short-interval data – daily or sub-daily. By the time you get to monthly averages the precipitation itself (not the logarithm) tends to be settling toward Gaussian, indicating additive error, although this depends on how frequent the precipitation is. The biweekly interval is in between; is there any way to assess how correct a multiplicative model is?

Response/Action: *Here, we distinguish between the distribution of the precipitation data and the additive/multiplicative error model. In our study, we do not assume any distribution for either the data or the errors. We only make the assumption that the error model be multiplicative. This can be evaluated using the joint PDFs of each pair of the products. If the joint PDFs have a constant spread across different values of precipitation; then, the error model is an appropriate one. We made this evaluation, and the PDFs resulting from the multiplicative model have better spread compared to the additive model. So, we concluded that for biweekly data it is better to assume the multiplicative model. We added this discussion on the model selection to the final submission to better justify the choice of multiplicative model.*

- Finally, the English is extremely clean; if there were an annual award for such excellence, you would deserve it. Overall, a very strong manuscript that just needs some tune-up on the way to acceptance.

Specific Comments:

1. Abstract: It would strengthen the Abstract to be more specific about the details of the comparison: $2^\circ \times 2^\circ$ grid boxes for a specific part of CONUS (not just "across the U.S."), using biweekly accumulations for the period January 2002 through April 2014.

   Response/Action: *We incorporated this in the final submission and revise the abstract accordingly.*

2. Dataset citation: The various datasets used are not cited and acknowledged in a consistent fashion, but should be. However, I would suggest that one of the newly emerging best practices in publication is to provide a reference-list citation for the data sets used, just as is done for journal articles. See the AMS policy http://www2.ametsoc.org/ams/index.cfm/publications/authors/journal-and-bamsauthors/journal-and-bams-authors-guide/data-archiving-and-citation/ for a discussion and examples. I would urge the authors to adopt this approach to give proper credit and guide the interested reader to the appropriate archives.

   Response/Action: *We thank the reviewer for the informative reference on data citation. We list appropriate citation to the datasets in the final version.*

3. Dataset names: Shortening "TRMM 3B42" to "TRMM" is ambiguous, since there are many TRMM products, while "3B42" is specific. The same comment applies to the GPCP 1DD, for which "GPCP" is ambiguous, while "1DD" is not.

   Response/Action: *We replaced the appropriate abbreviations in the final version.*

4. P.14,L.2-3: It would seem that the insightful statement is that the cloud systems are driven by frontal systems. GPI reacts to clouds, and fronts generate clouds that are not necessarily well-correlated to precipitation.

   Response/Action: *We acknowledge the revised statement, and corrected it in the final version.*

   Technical Corrections

5. P.8,L.22: The IR in 3B42 is calibrated by microwave before use in the product.

   Response/Action: *We point out this calibration in the revised manuscript.*

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 2527, 2015.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Hydrol. Earth Syst. Sci. Discuss., 12, C1982–C1989, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C1982/2015/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Characterization of precipitation product errors across the US using multiplicative Triple Collocation" *by* S. H. Alemohammad et al.

S. H. Alemohammad et al.

hamed_al@mit.edu

Received and published: 6 June 2015

We thank Dr. Crow for his constructive and insightful comments. Here, we respond to the general and specific comments included in the review:

General Comments:

- The paper describes the application of a modified triple collocation approach to the problem of evaluating large-scale precipitation data sets. The proposed modification allows for the more appropriate treatment of precipitation errors as multiplicative in nature. Issues surrounding the potential impact of error cross-correlation are examined via the decomposition of triple collocation error results over heavily-gauged reference sites. Overall, this is a high-quality paper on a topic of significant interest. The application of a log-transform to deal with multiplicative precipitation errors is a very nice methodological extension and clearly superior to the existing approach (of naively assuming that all errors are additive in nature in order to shoehorn them into a TC framework).

I also appreciated the effort to explicitly examine the role of error cross-correlation on TC-derived error estimates (in Section 5). However, the one important thing I felt was missing was a re-examination of results in Section 4 based on the (non-trivial) impacts of cross-correlated errors (isolated in Section 5). For example, a key result in Section 4 is the relative lack of accuracy for the GPI precipitation product. However, in Section 5 (and the supplementary materials) we also see that GPI is relatively more independent (i.e. contains less error cross-correlation) than the other precipitation data sets. Given that TC will penalize GPI for this lack of dependence . . . does this mean that the analysis in Section 4 is truly even-handed? Is GPI being unfairly penalized due to being truly independent from the other products – as opposed to it being FAIRLY penalized for its weak relationship with "true" precipitation? So basically, I'd like a little bit of guidance about how the conclusions presented in Section 4 should be re-examined given the cross-correlation issues presented in Section 5. Should the reader really trust that relative rankings presented in Section 4?

I understand that this is a generic problem with any TC analysis; however, I think there are a couple of things that the authors could do to better address this point. First, they could examine whether or not the relative rankings that they derive using TC (at the 6 reference pixel sites examined in Section 5) accurately reflect the rankings they achieve when comparing all the products against the high-quality rain gauge observations acquired at each sites. If TC can successfully replicate the gauge-based rank correlation analysis at these 6 sites – that would be good evidence that the spatially-distributed TC results in Section 4 are robust in a relative sense (despite the known bias issues associated with the neglect of error cross-correlation).

Response/Action: *This is a very interesting comparison and we appreciate it. In all of the 6 pixels that we conducted the gauge-based analysis, the TRMM 3B42 and NEXRAD products are ranked 1st and 2nd for the lowest error based on the RMSE from TC, respectively. Looking at the rankings in the gauge-based TC analysis (*σ_{TRE} *in Figures 8, S3, S4 and S5 in the original submission) in 5 out of the 6 pixels, TRMM 3B42 has the lowest error, and in 4 out of the 6 NEXRAD has the best error after TRMM 3B42. Therefore, in general, we can make the conclusion, the relative rankings for the products with the lowest error remains almost the same. However, GPCP 1DD and GPI rankings are only preserved on 3 out of the 6 pixels. This makes it hard to make the conclusion that these rankings are preserved for all the cases. Nevertheless, this is based on only 6 pixels out of the 75 pixels across the whole domain. So it is not possible to extend this conclusion to the whole study. We can conclude from this comparison that the cross-correlation error can impact the performance ranking of the precipitation products, but the relative impact needs further analysis. We included this comparison in the final submission with detailed explanations.*

- Another step that could be taken would be replace the TMPA 3B42 dataset with its "real-time" (RT) equivalent (TMPA 3B42RT) which is not gauge-corrected. This transition would make the "TRMM" precipitation product relatively more independent from the NEXRAD and GPCP datasets (which also have a gauge-correction component). Therefore, this transition towards greater error independence should lead to an increase in TC-derived error for the NEXRAD and GPCP products and a decrease in error for GPI (when considered as part of triplet that includes GPI) . The size of this increase (or decrease) could be used of an indication of how serious the cross-correlation problem is across the entire study domain. Does it – for example – significantly close the gap between GPI and NEXRAD TC results over the eastern part of the study area?

Response/Action: *We implemented this analysis, and included the detailed results in the final paper. Here, we present two figures showing the result of applying TC to the two sets of triplets while replacing TRMM 3B42 with TRMM 3B42RT (Figures 1 and 2). These figures indicate that using the real-time product of TRMM 3B42 will increase the TC-derived errors in NEXRAD and TRMM products and decrease the errors in GPI. This is a direct result of the non-zero cross-correlation. However, the general pattern of the error and the relative values between the the three products remains the same. The TRMM 3B42RT is still doing a better job compared to NEXRAD and GPI; then, NEXRAD is ranked 2nd in performance and GPI is the last one. From this example, we can conclude that the cross-correlation problem is a not a major issue in applying TC analysis to precipitation products in this case; however, it definitely impacts the absolute value of RMSEs derived from TC.*

- Therefore, prior to publication, I would strongly recommend that the authors address this issue in some manner. At the very least, add 2-3 sentences describing the consequences of the analysis in Section 5 on earlier results in Section 4.

Minor Suggestions:

1. Page 2536, Line 9-11: Clarify what exactly is meant by “homogeneous”? You mean homogeneous in a statistical climate sense . . . correct?

   Response/Action: *Exactly, this refers to the spatial homogeneity of the precipitation in a statistical climate sense. The statement will be revised in the final paper to clarify this.*

2. Page 2535, Line 20-22: I don’t see how zeros would violate the assumption of error independence . . . however I can see how they would cause fatal problems

in a logtransform analysis. The authors might want to re-write this sentence (or improve its clarity). Also, what about extremely low rainfall values (right at the edge of numerical precision) . . . can they skew results conducted in log-transform space?

Response/Action: *The issue with zeros is both violation of the error independence and log-transformation. Precipitation can only take values greater and equal to zero; therefore, it is bounded from one side. And if the precipitation estimate at a specific time and space is equal to zero; then, the error in that estimate can be from a limited set of numbers (basically any number greater than zero). So, it is dependent on the measurement (or equivalently the truth). As a result, if we have zero value in the precipitation measurement for all the triplets, the error of each of them is dependent on the measurement; and therefore, on each other. This error dependence decreases as the measurement value moves away from zero, and it can be present for rainfall values close to numerical precision of the system. This is also a minor problem when using TC on soil moisture data that is a bounded variable but is it a major issue with daily and subdaily precipitation data which has a lot of zeros. Therefore, in our analysis we are using biweekly accumulation and excluding the few percentage of zero values to reduce the impact of this dependence. Moreover, it is not possible to represent a non-zero error together with zero precipitation as the truth in the multiplicative model of Eq. (1).*

3. Page 2541, Line 1-4: I had to read this sentence several times to follow it . . . I'd recommend re-writing to clarify its meaning (e.g. be a bit more specific . . . representativeness error in what? . . . and cross-correlation in errors between what and what?).

Response/Action: *We believe this statement refers to page 2542, Lines 1-4. Assuming this, we revised the sentence to the following in our final submission to clarify the points: "It is understood that gauge data also have errors including*

> *representativeness error (they are point measurements unlike the other products that provide an average value over each pixel); however, as it is shown in Yilmaz and Crow (2014) (Appendix) the representativeness error in the gauge measurements causes a positive bias in the TC-based RMSE estimates while the cross-correlation between the errors of different products in each triplet causes a negative bias."*

List of Figures

1. **Figure 1**: RMSE of the precipitation rate estimated from TC using triplets in group 1; a) NEXRAD, b) TRMM 3B42RT, c) GPI. Panel d) shows the number of data points (biweekly measurements) in each pixel that are used for error estimation in TC analysis.
2. **Figure 2**: RMSE of the precipitation rate estimated from TC using triplets in group 2; a) NEXRAD, b) TRMM 3B42RT, c) GPCP 1DD. Panel d) shows the number of data points (biweekly measurements) in each pixel that are used for error estimation in TC analysis.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 2527, 2015.

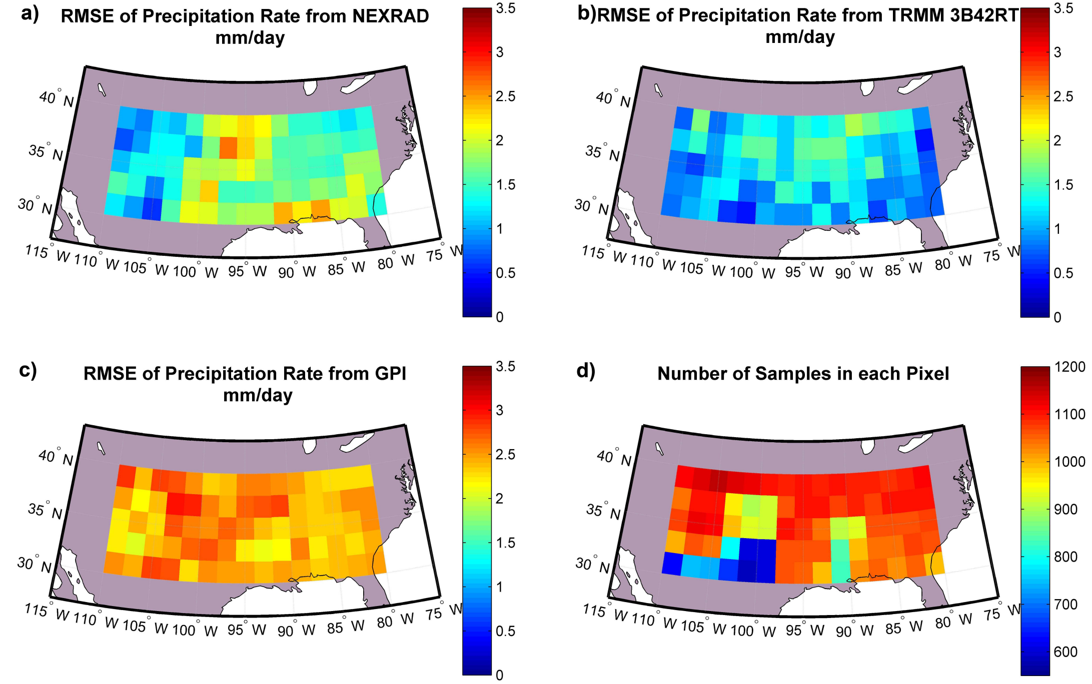


Fig. 1.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

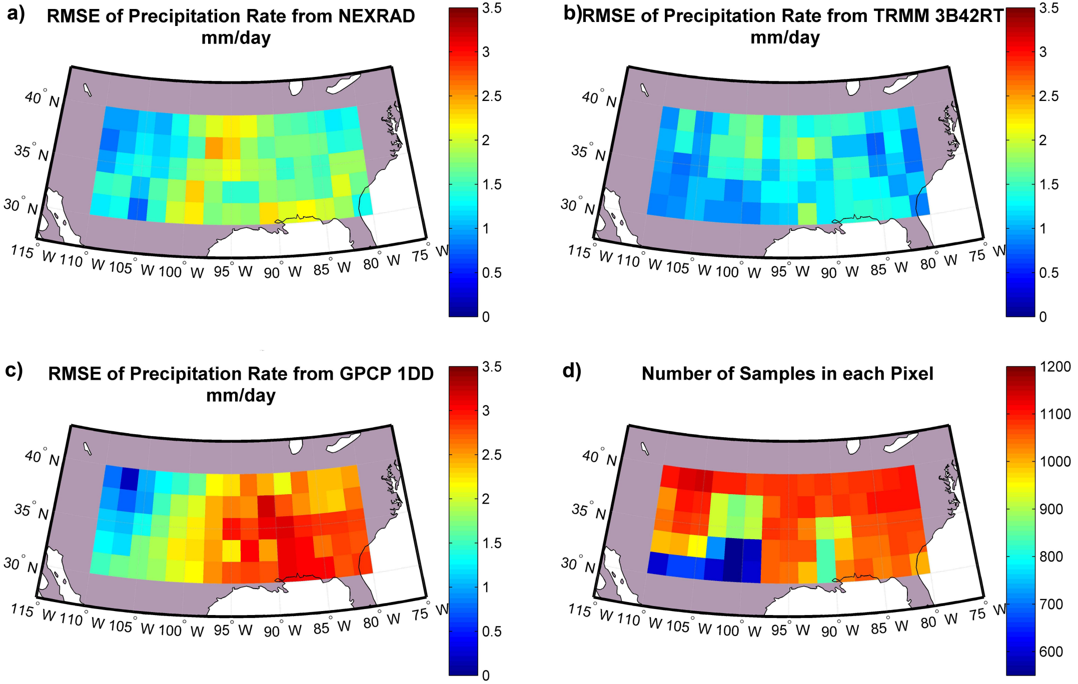


Fig. 2.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper