Evaluation of Real-Time Satellite Precipitation Data for Global Drought Monitoring

TIAN ZHOU AND BART NIJSSEN

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

GEORGE J. HUFFMAN

NASA Goddard Space Flight Center, Greenbelt, Maryland

DENNIS P. LETTENMAIER

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

(Manuscript received 29 July 2013, in final form 7 February 2014)

ABSTRACT

The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) near-real-time (RT) data are considered less accurate than the TMPA research quality (RP) data because of the simplified data processing algorithm and the lack of gauge adjustments. However, for near-real-time hydrological applications, such as drought nowcasting, the RT data must play a key role given latency considerations and consistency is essential with products like RP, which have a long-term climatology. The authors used a bivariate test to examine the consistency between the monthly RT and RP precipitation estimates for 12 yr (2000–12) and found that, for over 75% of land cells globally, RT and RP were statistically consistent at 0.05 significance level. The inconsistent grid cells are spatially clustered in western North America, northern South America, central Africa, and most of Australia. The authors also show that RT generally increases with time relative to RP in northern South America and western Australia, while in western North America and eastern Australia, RT decreases relative to RP. In other areas such as the eastern part of North America, Eurasia, and southern part of the South America, the RT data are statistically consistent with the RP data and are appropriate for global- or macroscale hydrological applications.

1. Introduction

Precipitation is a key component of the water cycle and is the most important determinant of land surface hydrologic extremes (e.g., droughts), which result in billions of dollars of economic losses (Smith and Katz 2013) and human suffering each year. However, especially in the developing world, ground-based precipitation records are insufficient to support modern hydrologic prediction methods (Su et al. 2008; Yong et al. 2010; Bitew and Gebremichael 2011). Satellite-based remote sensing of precipitation offers an alternative source of precipitation information for hydrologic prediction that can resolve the space–time resolution deficiencies of in situ networks (Sapiano and Arkin 2009). The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) provides 3-hourly, 0.25° × 0.25°, near-global precipitation estimates (50°N–50°S) based on data from multiple passive microwave and infrared satellite sensors. The TMPA real-time product (RT) is available within about 9 h after the time of observation. There is also a post-real-time research-quality TMPA product (RP), available about 2 months following the month of the observation, which includes an adjustment to available in situ gauges and certain other corrections (Huffman et al. 2007).

Although the RT data are considered less accurate than the RP data (a result of the gauge adjustment in RP and simplified data processing in RT), the RT data are most appropriate for real-time hydrological applications, such as drought monitoring (e.g., Nijssen et al. 2014) and flood prediction (e.g., Wu et al. 2012), because of data latency considerations. The latest version of TMPA is version 7 (V7), which has recently been updated

Corresponding author address: Dr. Dennis P. Lettenmaier, Department of Civil and Environmental Engineering, University of Washington, 202D Wilson Ceramic Laboratory, Box 352700, Seattle, WA 98195-2700.
E-mail: dennisl@uw.edu

DOI: 10.1175/JHM-D-13-0128.1

© 2014 American Meteorological Society

and reprocessed back to March 2000 (Huffman and Bolvin 2013). This period of consistently processed precipitation record is a potential basis for estimation of the empirical precipitation distributions that are now commonly used in drought characterizations (e.g., Sheffield et al. 2004; Andreadis et al. 2005).

In this paper, we evaluate the TMPA V7 RT dataset for (statistical) consistency in accumulated precipitation amounts with the RP product over the period 2000–12 for the entire global (land) domain. We use the RP product as the benchmark for our comparisons because, on a monthly basis, the RP data essentially reproduce the characteristics of the gridded station data used in the adjustment process (Huffman et al. 2007). In the RP product, the satellite precipitation estimates (processed similar to the RT) are used only to partition the monthly totals to 3-hourly values. Insofar as we focus here only on the statistical stationarity of RT relative to RP at a monthly time scale, our results are of particular interest for applications that are sensitive to accumulated precipitation amounts, such as large-scale drought monitoring systems. For flood analysis and prediction, statistical properties of the short-term (3 hourly) data are more important than for drought, and our discussion here is not aimed at such applications.

2. Data and methods

The V7 RT and RP 3-hourly products were extracted for the period from 1 March 2000 (the earliest available V7 RT) to 29 February 2012 (144 months) from the National Aeronautics and Space Administration (NASA) Goddard Earth Sciences (GES) Data and Information Services Center (DISC) FTP sites. Both products have a spatial resolution of 0.25°, and we aggregated them to monthly totals.

We applied the Maronna and Yohai (1978) bivariate test as adapted by Potter (1981) for hydrological use. It has been adopted by the hydrology community to detect the statistical homogeneity of meteorological and hydrological time series (e.g., Lettenmaier et al. 1994; Plummer et al. 1995; Štěpánek et al. 2009; Jones 2012). It essentially provides a statistical hypothesis test framework for double mass curve analysis, which is used in hydrology and applied climatology (e.g., Dingman 2008) to detect systematic drifts of one time series relative to another. A constant slope of the double mass curve indicates consistency between two datasets, while changes in the slope indicate the departure of one dataset from the other. The departure could be caused by changes in data collection methods or the algorithms used in data processing (Searcy and Hardison 1960). An underlying assumption of the bivariate test is that the two time series to be compared are each temporally independent (but may be, and usually are, cross correlated) with a bivariate normal distribution. The test arguably is appropriate to long-term monthly precipitation time series, which either are or can be transformed to be approximately normally distributed (Lettenmaier et al. 1994) and usually are only modestly serially correlated (Yevjevich 1967).

The null hypothesis for the test is that the two time series $\{x_i\}$ and $\{y_i\}$ come from the same bivariate normal distribution. The test statistic T_0 is described as

$$X_i = \frac{1}{i}\sum_{j=1}^{i} x_j, \quad Y_i = \frac{1}{i}\sum_{j=1}^{i} y_j,$$

$$\overline{X} = X_n, \quad \overline{Y} = Y_n,$$

$$S_x = \sum_{j=1}^{n} (x_j - \overline{X})^2, \quad S_y = \sum_{j=1}^{n} (y_j - \overline{Y})^2,$$

$$S_{xy} = \sum_{j=1}^{n} (x_j - \overline{X})(y_j - \overline{Y}),$$

$$F_i = S_x - \frac{(X_i - \overline{X})^2 ni}{n-1}, \quad i < n,$$

$$D_i = \frac{[S_x(\overline{Y} - Y_i) - S_{xy}(\overline{X} - X_i)]n}{[(n-1)F_i]},$$

$$T_i = \frac{[i(n-1)D_i^2 F_i]}{S_x S_y - S_{xy}^2}, \quad \text{and} \tag{1}$$

$$T_0 = \max\{T_i\}. \tag{2}$$

For each time step i ($i = 1, 2, \ldots, n$, where $n = 144$), T_i is calculated, which represents the adjusted cumulative departure from one time series to the other. The test statistic T_0 is then determined as the maximum value of T_i over n time steps. The time i_0 corresponds to the time step when T_i equals T_0 and is an estimate of the time at which one time series begins to shift relative to the other. Critical values of the test statistic T_c for given significance levels α can be computed via Monte Carlo simulation for $n = 144$. We determined that T_c is 9.6 at $\alpha = 0.05$ and 12.8 at $\alpha = 0.01$. We also further examined the slope breaks of the double mass curves in those grid cells for which rejections of the null hypothesis occurred. This analysis identified grid cells where RT increased or decreased relative to RP.

3. Results

Figures 1a and 1b show results of the bivariate test applied to the V7 RT and RP products. The null

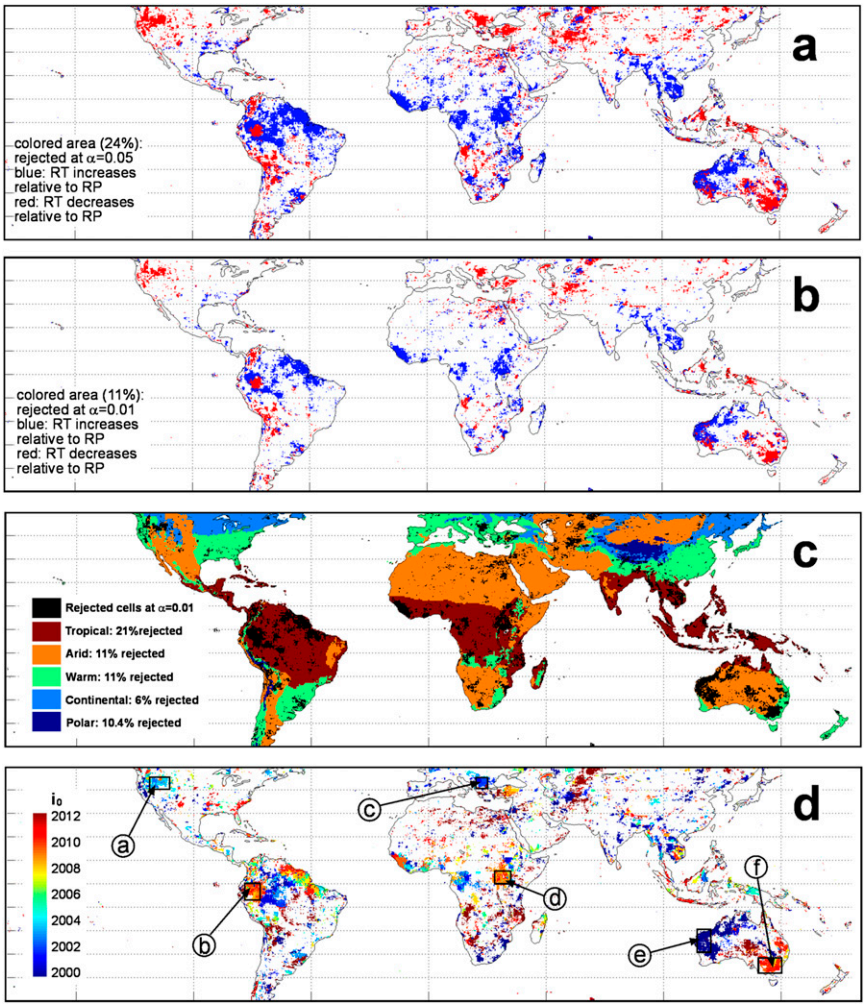

FIG. 1. Areas that are rejected by the bivariate test at (a) $\alpha = 0.05$ and (b) $\alpha = 0.01$. Blue indicates RT increases relative to RP, and red indicates RT decreases relative to RP. (c) Percentages of rejected grid cells in five Köppen climate regions. (d) The i_0 (time of the max test statistic in the bivariate test) distribution for rejections at $\alpha = 0.05$. Six rectangular regions selected for additional analyses are shown (see Fig. 2). They are identified (from west to east) as: the northwestern United States – a, northern South America – b, eastern Europe – c, central Africa – d, western Australia – e, and southeastern Australia – f (see Table 1).

hypothesis, that there is no systematic difference, was rejected for about 24% of the grid cells at the 0.05 significance level (Fig. 1a). The figure also shows the general trend of the double mass curve, with areas in which RT increases relative to RP shown in blue and areas in which RT decreases relative to RP in red. In general, Eurasia, northern Africa, most of North America, and the eastern and southern part of South America show a high degree of consistency between RT and RP. Spatially coherent areas of inconsistency include the northwestern part of the contiguous United States, the northern part of South America, central Africa, and most of Australia. There are also inconsistent grid cells in southern Africa and eastern Europe, but they are less spatially coherent than those noted above. In addition, the values of the test statistic T_0 (maximum value of T_i) are quite close to the 0.05 critical level in these areas. Therefore, it is useful to focus on the areas in which the null hypothesis was rejected at a higher significance level. Figure 1b shows the areas where the null hypothesis was rejected at $\alpha = 0.01$ (about 11% of grid cells).

The percentage of rejected cells changes in different climate regions. Figure 1c shows the cells rejected at $\alpha = 0.01$ in five Köppen climate regions. The tropical rain forest region has the highest rejection percentage

TABLE 1. Summary of six selected rectangular areas. See Fig. 1d for area locations.

	Northwestern United States (a)	Northern South America (b)	Eastern Europe (c)	Central Africa (d)	Western Australia (e)	Southeastern Australia (f)
Lat	40°–46°N	5°S–0°	40°–45°N	0°–5°N	30°–20°S	38°–32°S
Lon	120°–110°W	78°–70°W	20°–25°E	30°–36°E	114°–120°E	141°–150°E
Pixels rejected at $\alpha = 0.05$ (%)	72	94	75	87	82	78
Pixels rejected at $\alpha = 0.01$ (%)	53	83	52	72	77	62
Coordinates of the pixel with max T_0	43.5°N, 119°W	4°S, 73.5°W	44.5°N, 23°E	0.5°N, 32.5°E	24°S, 115°E	36.5°S, 148.5°E
T_0 of area mean bivariate test	15.62	34.93	18.39	46.02	53.56	22.13
Year of i_0 in area mean bivariate test	2003	2009	2002	2009	2000	2010

(21%), almost twice the global average, and 6% of grid cells were rejected in polar regions, which is the lowest rejection rate globally. The other three regions (arid, warm, and continental) have about 11% rejection, which is the same as the global average.

We then examined i_0, the time when a change or inconsistency began to develop between RT and RP. Figure 1d shows i_0 for all cells in which T_0 exceeds the $\alpha = 0.05$ threshold. It is clear that i_0 has a strong spatial coherence in these rejected areas, indicating that RT and RP have experienced a relative shift at about the same time (and hence, there may be a common cause within the spatially coherent areas).

To analyze the relationship between the test statistics and the RT and RP time series, we examined six rectangular areas in which most of the cells (over 70%) were rejected in the bivariate test at $\alpha = 0.05$ (Fig. 1d; Table 1). For each of these areas, we show the area-averaged RT

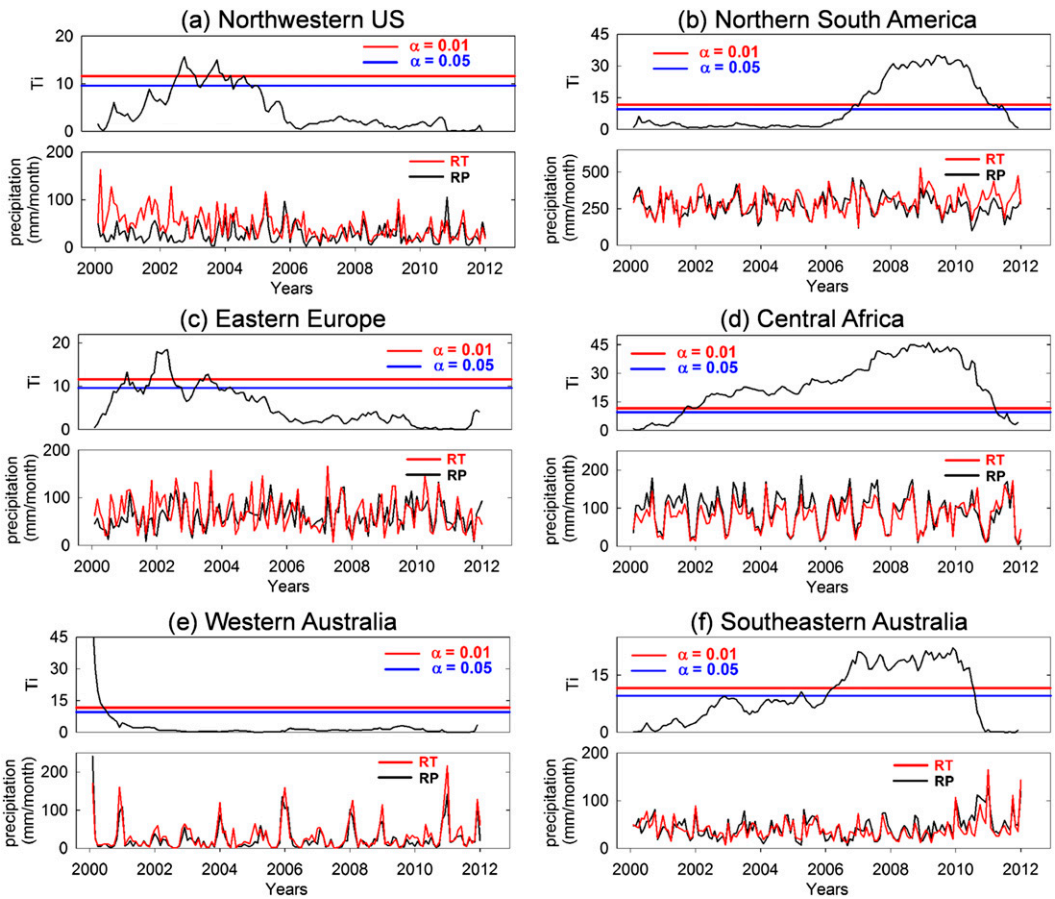

FIG. 2. Time series of the test statistic T_i (cumulative departure between RT and RP) and area-averaged RT and RP estimates for regions a, b, c, d, e, and f in Fig. 1d.

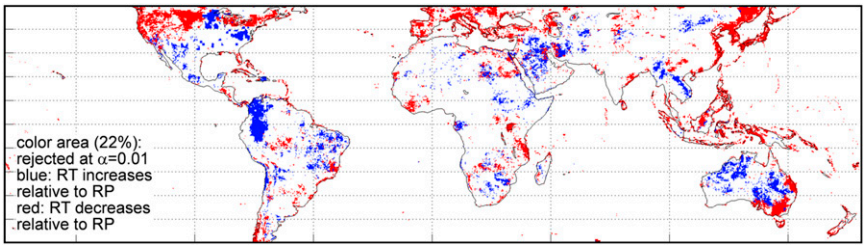

FIG. 3. Areas that are rejected in the bivariate test between TMPA V6 RT and RP for $\alpha = 0.01$. Blue indicates that RT increases relative to RP; red indicates that RT decreases relative to RP.

and RP time series as well as the bivariate test results for the two area-averaged time series (Fig. 2). For the northwestern United States (Fig. 2a), T_i exceeds the T_c threshold ($\alpha = 0.01$) in 2003 and 2004. The nature of the inconsistency is clear from the time series, with cumulative RT decreasing relative to RP over the period of analysis. Note that the test does not evaluate the bias of RT relative to RP, but merely whether there is a change in this bias over time. We note that this is a mostly mountainous region, so one is tempted to conclude that the divergence of RT relative to RP is related to terrain and/or mountain surface processes (such as snow). However, there are many other areas of the globe with topographically complex terrain where no such divergence is apparent, and the reasons for the divergence remain elusive. In northern South America (Fig. 2b), T_i reaches a peak in 2009. Time series plots indicate that the cumulative RT and RP series remain consistent until 2007, when RT increases relative to RP. In eastern Europe (Fig. 2c), the T_i pattern is similar to that in Fig. 2a, with a maximum around 2002 and relatively low values afterward. However, unlike the northwestern United States, there is no obvious increase of the RT time series relative to RP. Instead, RT and RP agree for most of the period except for a few extremes in RT series that occurred in 2002 and 2003. In central Africa (Fig. 2d), T_i exceeds the 0.01 critical value through 2002–11 with the peak at 2008. The continued exceeding of the T_c threshold indicates that the bias of RT relative to RP changes throughout this period. The time series suggests that the RT values are generally a little lower than RP from 2000 to 2008 and then become close thereafter. In western Australia (Fig. 2e), T_i only exceeded the 0.01 critical value in the first year and remains lower than the critical value afterward. The time series indicates that RP had a very large peak value during the first month compared to RT, after which the two records were similar. In southeastern Australia (Fig. 2f), the maximum T_i occurred in 2010, exceeding the 0.01 significance level. Even though the change in bias is difficult to detect visually, RT is smaller than RP for several events from 2008 to 2011, which is the apparent cause of the null hypothesis being rejected.

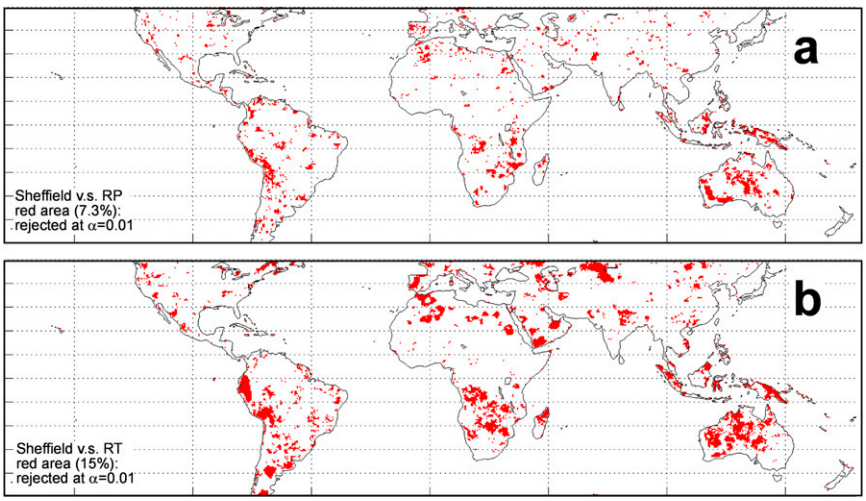

FIG. 4. Bivariate test from 2001 to 2010 between Sheffield et al. (2006) data and TMPA V7 (a) RP and (b) RT.

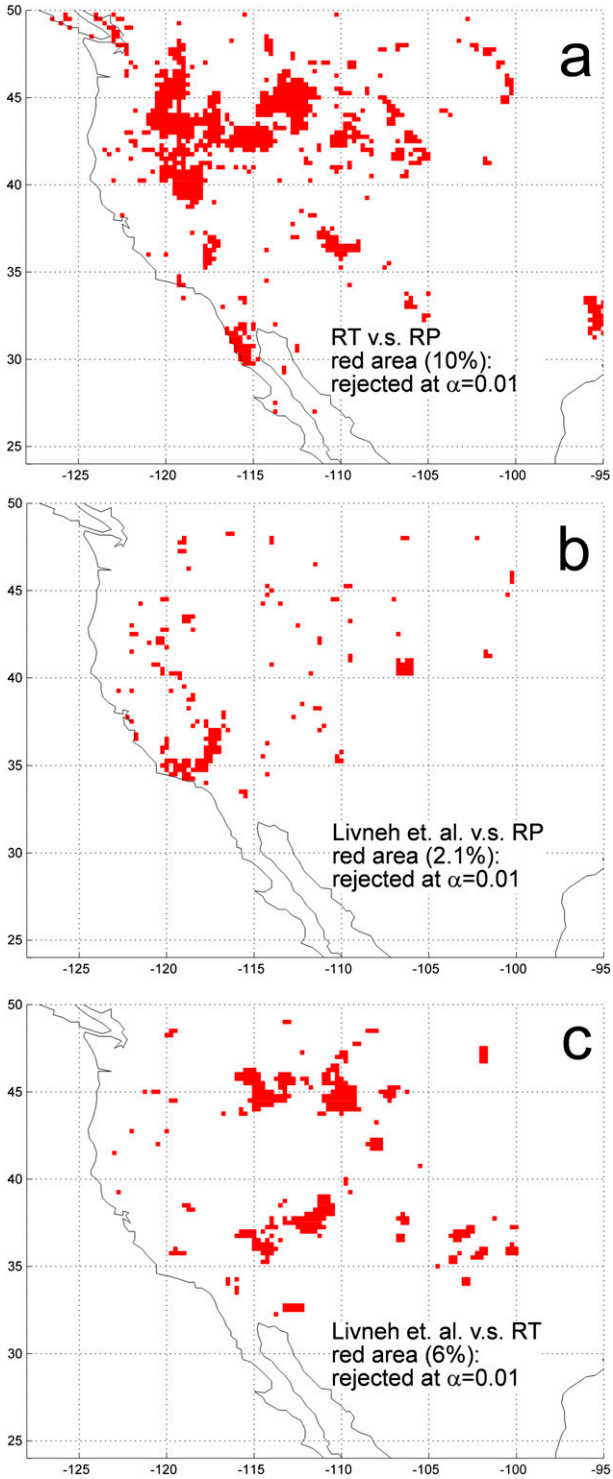

FIG. 5. Bivariate test in the western United States (from 100°W) from 2001 to 2011 between TMPA V7 (a) RT and RP, (b) Livneh et al. (2013) and TMPA V7 RP, and (c) Livneh et al. (2013) and TMPA V7 RT.

4. Discussion

Although a thorough analysis of the causes of the illustrated inconsistencies is not the focus of this paper, we conducted a series of diagnostic analyses to explore possible reasons for the observed inconsistencies. First, we applied the bivariate test to the TMPA V6 RT and RP data (V6 is the predecessor to V7) from 2004 to 2010. We found that even though the V6 test period (7 years) is 40% shorter than the V7 test period (12 years) and hence the test is considerably less powerful, the inconsistent area between V6 RT and RP at $\alpha = 0.01$ (22%) is about double that between V7 RT and RP (11%; Fig. 3). This suggests that inconsistencies also exist in the previous version of TMPA products, which is expected given the inhomogeneous V6 RT processing and the shift in V6 RP gauge source in May 2005 (Huffman et al. 2010). Our tests also suggest that the data consistency has been improved in V7.

Second, we applied the bivariate test to the TMPA products and a ¼° global gauge-based precipitation dataset (Sheffield et al. 2006) from 2001 to 2010 (note that the dataset has been extended by the authors subsequent to the original paper) to determine whether the benchmark dataset (RP) itself might be responsible for the observed differences. The test results (Fig. 4) show that only 7.3% of the grid cells were inconsistent between the Sheffield et al. data and RP at $\alpha = 0.01$. However, more than twice as many grid cells (15%) were rejected in the Sheffield et al. versus RT test, indicating that the RP dataset is much more consistent with the gauge observations than is RT.

We further investigated the western United States (25°–50°N, west of 100°W) by applying the bivariate test to a ¼° gauge-based gridded precipitation dataset [aggregated from the 1/16° dataset for the conterminous United States of Livneh et al. (2013)] and the TMPA RT and RP datasets from 2001 to 2011. This region is one of the largest contiguous, inconsistent areas in our test and generally has higher-quality gauge-based observations than outside the United States. We found that the percentage of the western United States for which the Livneh et al. data and RP are inconsistent at $\alpha = 0.01$ was only about 2%, much smaller than for RT versus RP over this region (about 10%; Fig. 5). Both the global and western U.S. tests suggest that the inconsistencies between RT and RP are real and are not an artifact of drift in RP, which effectively serves as a benchmark in our analyses.

Third, we examined the consistency between RT and RP for different seasons. RT and RP estimates in summer [June–September (JJAS) in the Northern Hemisphere and December–March (DJFM) in the Southern Hemisphere] and in winter (DJFM in the

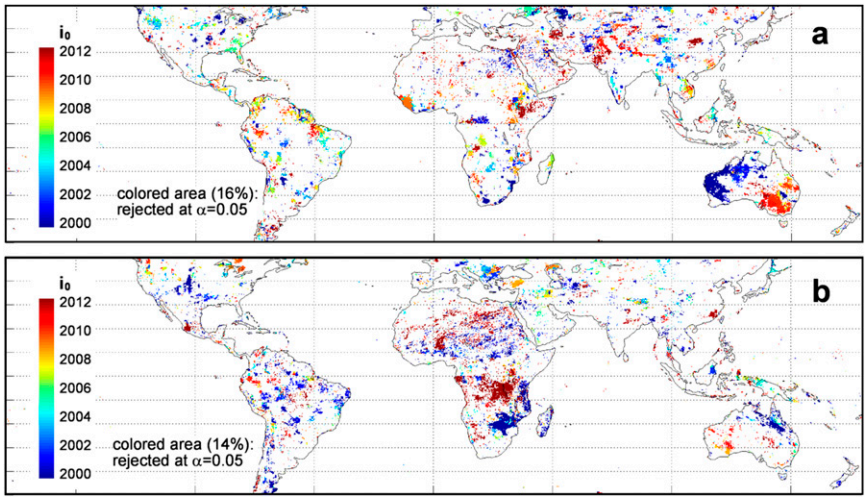

FIG. 6. Distribution of the time of the max test statistic (i_0) for grid cells rejected at $\alpha = 0.05$ in (a) summer (JJAS for the Northern Hemisphere, DJFM for the Southern Hemisphere) and (b) winter (DJFM for the Northern Hemisphere, JJAS for the Southern Hemisphere).

Northern Hemisphere and JJAS in the Southern Hemisphere) were selected for two additional bivariate tests. Although the sample size for each test (4 months × 12 years = 48) was relatively small in a statistical sense, and hence the test power is lower than for the entire series, the results do give some hint of the effects of seasonality on data consistency. The results (Fig. 6) indicate that the total number of inconsistent grid cells in summer and winter are similar—about 16% in summer and 14% in winter at $\alpha = 0.05$. However, the spatial distribution of the rejected grid cells is very different. In Southern Hemisphere summer, the largest inconsistent area is concentrated in Australia, but in winter most of the inconsistent areas are located in mid- and southern Africa. These findings imply that the consistency between RT and RP has strongly seasonal variations in different locations.

Fourth, we examined the relationship between the number of precipitation gauges used in RP and the consistency between RT and RP. The gauge observations used in the RP product come from the Global Precipitation Climatology Centre (GPCC; Huffman et al. 2010). For the period 1998–2010, RP used the GPCC Full Data Reanalysis (V6) product at 1° spatial resolution, which superimposes the observed monthly anomalies on the month's climatology from the period 1951–2000 (Becker et al. 2013; Huffman et al. 2010). After 2010, RP used the GPCC Monitoring Product (V4), which is processed in a manner consistent with the GPCC Full analysis but generally has fewer observations (Adler et al. 2003). Furthermore, the number of gauges included in the GPCC V6 product decreased from more than 30 000 to less than 10 000 over the period 2001–10. To investigate the effect of changes in the number of gauges included in RP, we selected two months (May 2001 and December 2010) that have the largest (32 018) and smallest (9493) gauge numbers in the underlying GPCC V6 product. The number of GPCC gauges used in RP in these two months was noted for each 3 × 3 gridcell window in the domain, centered on the grid cell of interest. The consistency for each grid cell was represented by T_0 calculated from the bivariate test. The percentiles of T_0 are plotted in Fig. 7 as a function of the numbers of gauges. Figure 7 suggests that there is essentially no relationship between (the median of) T_0 and the number of gauges. Generally, declining values of the 100th percentile with increasing number of stations are a reflection of the larger range of T_0 for small numbers of stations and appear to be related to occasionally anomalous behavior of RP when the number of stations in the constraining GPCC product is small.

Further examination of the shape of the T_i distribution as well as the RT and RP time series in the inconsistent grid cells allows us to group the inconsistencies qualitatively into two categories: 1) those largely caused by under- or overestimation of several peak values, with minimal change between RT and RP most of the time (type I, e.g., areas c, e, and f), and 2) those largely caused by systematic drift, with the difference between RT and RP changing over the test period (type II, e.g., areas a, b, and d). To differentiate the two types of inconsistencies, we simply eliminated the k pairs of data that had the largest absolute differences between RT and RP in each rejected grid cell and then performed the bivariate test

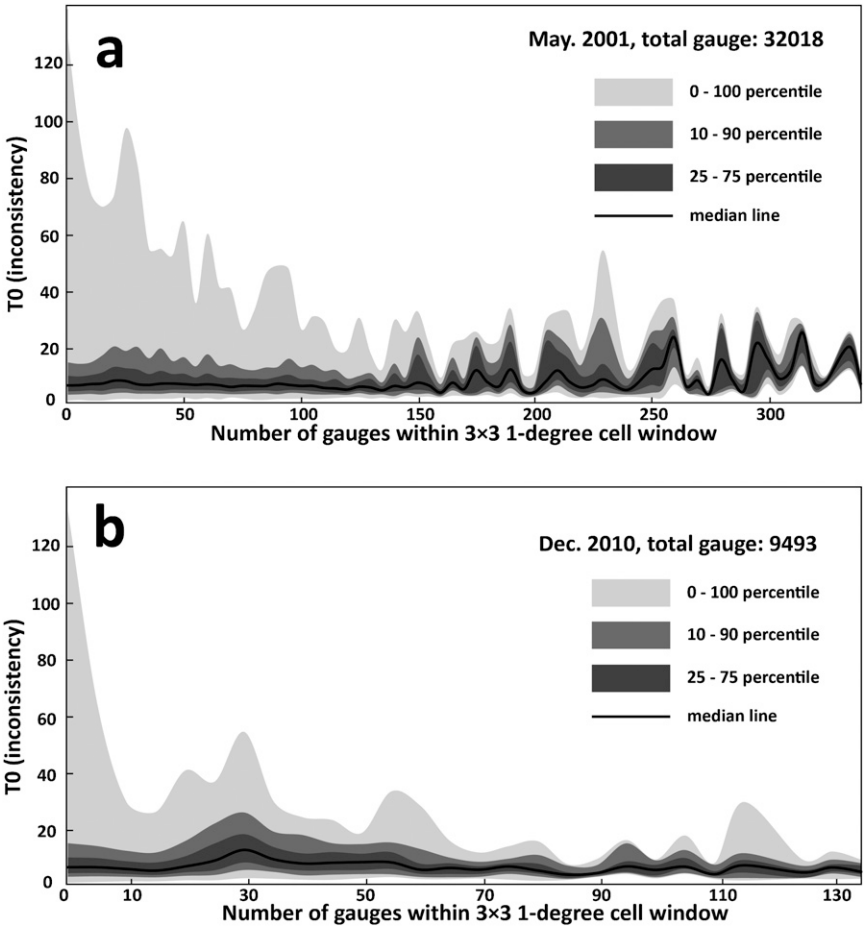

FIG. 7. Percentile of T_0 (higher means more inconsistent) against the number of gauges within 3×3 1° cell window for (a) May 2001 and (b) December 2010.

again on the reduced dataset ($k = 15$ in this case, or about 10% of the sample size). Grid cells that passed the second round of testing were considered to have type I inconsistencies, while the others were considered to have type II inconsistencies (Fig. 8). Overall, we found on this basis that at $\alpha = 0.05$, about 40% of the rejected grid cells had type I inconsistencies and the remaining 60% had type II inconsistencies.

Type I inconsistencies may be associated with the poor quality of measurements and estimates during extreme precipitation events, documented, for example, by Katsanos et al. (2004) in the eastern Mediterranean and Su et al. (2008) in central South America. Both studies found that previous versions of TMPA RT products tended to overestimate precipitation for large precipitation thresholds. Another reason for the first type inconsistency could be calibration issues in the TMPA-RT algorithm at and past the edge of the TRMM satellite coverage, as suggested by Villarini (2010) in Rome, Italy [for a previous version (V6) of RP data].

Type II inconsistencies may be induced by changes to the mix of satellite instruments that are used in the construction of the precipitation products as well as changes in the gauge datasets used for calibration in the construction of the RP product. These include satellite additions or termination during the test period, shifts in the timing of the satellite observations during the day, algorithmic changes in the weighting of the contribution of particular satellites, and changes in RT calibration performance in regions with strong gradients. One example of a change to the gauge dataset is a shift in December 2010 from the GPCC Full analysis to the GPCC Monitoring analysis, as noted above. These changes directly impact the RP product because the GPCC analysis has a dominant role at the monthly scale if "sufficient" gauge observations are available. In real-time hydrological applications, the first type of inconsistency may induce under- or overestimation of extreme events, while the second type of inconsistency can introduce systematic bias in hydrologic predictions.

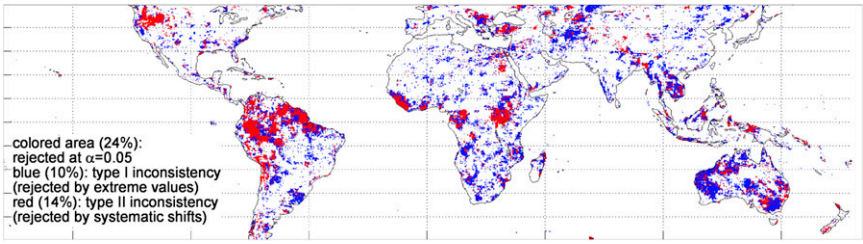

FIG. 8. Areas rejected by the bivariate test between TMPA V7 RT and RP at $\alpha = 0.05$. Blue indicates type I inconsistency, which mainly is caused by under- or overestimations of several peak values; red indicates type II inconsistency, which mainly is caused by systematic drifts.

In summary, the inconsistencies between TMPA V7 RT and RP appear to result from multiple causes. Although the specific causes for inconsistencies in various regions remain unclear, the bivariate test provides a powerful tool to examine algorithm deficiencies and identify inconsistent regions where satellite and/or gauge data availability changes.

5. Conclusions

Our analysis of the recently released TMPA V7 RT and RP datasets from 2001 to 2012 shows that there is no shift in the mean between the two datasets for over 75% of the global land area from 50°N to 50°S at a significance level of 0.05 and for nearly 90% at a significance level of 0.01. Our analysis also shows that cumulative RT increases relative to RP in northern South America and western Australia over time, while in western North America and eastern Australia, RT decreases relative to RP. In other areas such as eastern North America, Eurasia, and southern South America, the RT data are statistically consistent with the RP data and arguably can be considered to be a reliable precipitation input for global-scale or macroscale hydrological applications. In these areas, any existing bias is constant with time. Remarkable spatial coherence was observed in the date(s) at which the test statistic T_i reached its maximum value. Further explorations and analyses will be required to understand the underlying causes of the inconsistency between RT and RP.

Acknowledgments. This study was supported by NASA Grant NNX10AG87G to the University of Washington. The authors thank David Bolvin for thoughtful comments, Huilin Gao for data sharing, and Faisal Hossain for his suggestions on an earlier draft of the manuscript.

REFERENCES

Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.,* **4,** 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier, 2005: Twentieth-century drought in the conterminous United States. *J. Hydrometeor.,* **6,** 985–1001, doi:10.1175/JHM450.1.

Becker, A., P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider, and M. Ziese, 2013: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth Syst. Sci. Data,* **5,** 71–99, doi:10.5194/essd-5-71-2013.

Bitew, M. M., and M. Gebremichael, 2011: Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. *Water Resour. Res.,* **47,** W06526, doi:10.1029/2010WR009917.

Dingman, S. L., 2008: *Physical Hydrology.* 2nd ed. Waveland Press, 656 pp.

Huffman, G. J., and D. T. Bolvin, 2013: TRMM and other data precipitation data set documentation. Internal document, NASA, 40 pp. [Available online at ftp://meso-a.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf.]

——, and Coauthors, 2007: 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.

——, R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM Multi-Satellite Precipitation Analysis (TMPA). *Satellite Rainfall Applications for Surface Hydrology,* M. Gebremichael and F. Hossain, Eds., Springer, 3–22.

Jones, R. N., 2012: Detecting and attributing nonlinear anthropogenic regional warming in southeastern Australia. *J. Geophys. Res.,* **117,** D04105, doi:10.1029/2011JD016328.

Katsanos, D., K. Lagouvardos, V. Kotroni, and G. Huffmann, 2004: Statistical evaluation of MPA-RT high-resolution precipitation estimates from satellite platforms over the central and eastern Mediterranean. *Geophys. Res. Lett.,* **31,** L06116, doi:10.1029/2003GL019142.

Lettenmaier, D. P., E. F. Wood, and J. R. Wallis, 1994: Hydro-climatological trends in the continental United States, 1948–88. *J. Climate,* **7,** 586–607, doi:10.1175/1520-0442(1994)007<0586:HCTITC>2.0.CO;2.

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Updates and extensions. *J. Climate,* **26,** 9384–9392, doi:10.1175/JCLI-D-12-00508.1.

Maronna, R., and V. J. Yohai, 1978: A bivariate test for the detection of a systematic change in mean. *J. Amer. Stat. Assoc.,* **73,** 640–645, doi:10.1080/01621459.1978.10480070.

Nijssen, B., S. Shukla, C. Lin, H. Gao, T. Zhou, J. Sheffield, E. F. Wood, and D. P. Lettenmaier, 2014: A prototype global drought information system based on multiple land surface models. *J. Hydrometeor.,* doi:10.1175/JHM-D-13-090.1, in press.

Plummer, N., Z. Lin, and S. Torok, 1995: Trends in the diurnal temperature range over Australia since 1951. *Atmos. Res.,* **37,** 79–86, doi:10.1016/0169-8095(94)00070-T.

Potter, K. W., 1981: Illustration of a new test for detecting a shift in mean in precipitation series. *Mon. Wea. Rev.,* **109,** 2040–2045, doi:10.1175/1520-0493(1981)109<2040:IOANTF>2.0.CO;2.

Sapiano, M., and P. Arkin, 2009: 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.

Searcy, J., and C. Hardison, 1960: Double-mass curves. USGS Water Supply Paper 1541-B, 66 pp. [Available online at http://pubs.usgs.gov/wsp/1541b/report.pdf.]

Sheffield, J., G. Goteti, F. Wen, and E. F. Wood, 2004: A simulated soil moisture based drought analysis for the United States. *J. Geophys. Res.,* **109,** D24108, doi:10.1029/2004JD005182.

——, ——, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate,* **19,** 3088–3111, doi:10.1175/JCLI3790.1.

Smith, A. B., and R. W. Katz, 2013: US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Nat. Hazards,* **67,** 387–410.

Štěpánek, P., P. Zahradníček, and P. Skalák, 2009: Data quality control and homogenization of air temperature and precipitation series in the area of the Czech Republic in the period 1961–2007. *Adv. Sci. Res.,* **3,** 23–26, doi:10.5194/asr-3-23-2009.

Su, F., Y. Hong, and D. P. Lettenmaier, 2008: 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.

Villarini, G., 2010: Evaluation of the research-version TMPA rainfall estimate at its finest spatial and temporal scales over the Rome metropolitan area. *J. Appl. Meteor. Climatol.,* **49,** 2591–2602, doi:10.1175/2010JAMC2462.1.

Wu, H., R. F. Adler, Y. Hong, Y. Tian, and F. Policelli, 2012: Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. *J. Hydrometeor.,* **13,** 1268–1284, doi:10.1175/JHM-D-11-087.1.

Yevjevich, V., 1967: An objective approach to definitions and investigations of continental hydrologic droughts. Hydrology Paper 23, Colorado State University, Fort Collins, CO, 18 pp.

Yong, B., L. Ren, Y. Hong, J. Wang, J. J. Gourley, S. Jiang, X. Chen, and W. Wang, 2010: Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. *Water Resour. Res.,* **46,** W07542, doi:10.1029/2009WR008965.