SPECIAL
State of the Science of Precipitation
COLLECTION

Real-Time Bias Reduction for Satellite-Based Precipitation Estimates

YUDONG TIAN

Earth System Science Interdisciplinary Center, University of Maryland, College Park, and Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

CHRISTA D. PETERS-LIDARD

Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

JOHN B. EYLANDER

Air Force Weather Agency, Air and Space Models Integration Branch, Offutt AFB, Nebraska

(Manuscript received 3 December 2009, in final form 24 May 2010)

ABSTRACT

A new approach to reduce biases in satellite-based estimates in real time is proposed and tested in this study. Currently satellite-based precipitation estimates exhibit considerable biases, and there have been many efforts to reduce these biases by merging surface gauge measurements with satellite-based estimates. Most of these efforts require timely availability of surface gauge measurements. The new proposed approach does not require gauge measurements in real time. Instead, the Bayesian logic is used to establish a statistical relationship between satellite estimates and gauge measurements from recent historical data. Then this relationship is applied to real-time satellite estimates when gauge data are not yet available. This new scheme is tested over the United States with six years of precipitation estimates from two real-time satellite products [i.e., the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) research product 3B42RT and the NOAA Climate Prediction Center (CPC) Morphing technique (CMORPH)] and a gauge analysis dataset [i.e., the CPC unified analysis]. The first 4-yr period was used as the training period to establish a satellite–gauge relationship, which was then applied to the last 2 yr as the correction period, during which gauge data were withheld for training but only used for evaluation. This approach showed that satellite biases were reduced by 70%–100% for the summers in the correction period. In addition, even when sparse networks with only 600 or 300 gauges were used during the training period, the biases were still reduced by 60%–80% and 47%–63%, respectively. The results also show a limitation in this approach as it tends to overadjust both light and strong events toward more intermediate rain rates.

1. Introduction

Precipitation estimates from satellite-based sensors have become indispensible in a wide range of hydrological applications, especially since the launch of the Tropical Rainfall Measuring Mission (TRMM). Their global coverage, timely availability and unprecedented spatial and temporal resolutions make them crucial in studies such as global water cycle, water resources, crop yield, droughts, floods, and landslides.

Despite their advantages, currently purely satellite-based precipitation estimates contain considerable errors. This is primarily a result of the inherently indirect nature of precipitation remote sensing, which mostly derives precipitation rates from infrared (IR) or microwave signatures of cloud or ice particles, and to the limited spatial and temporal sampling of the space-borne sensors. Most of the current data products take advantage of the availability of multiple IR and microwave sensors to optimally intercalibrate and merge the retrievals from these sensors, in an effort to reduce these errors. Nevertheless, compared with ground-based gauge or radar measurements, these products still have much room to improve (e.g., Sorooshian et al. 2000; Gottschalck et al. 2005; Yilmaz et al. 2005; Ebert et al. 2007; Tian

Corresponding author address: Yudong Tian, NASA Goddard Space Flight Center, Mail Code 614.3, Greenbelt, MD 20771.
E-mail: Yudong.Tian@nasa.gov

DOI: 10.1175/2010JHM1246.1

© 2010 American Meteorological Society

et al. 2007; Sapiano and Arkin 2009; Habib et al. 2009; Tian et al. 2010). For example, Gottschalck et al. (2005) showed that satellite-based estimates have a correlation of 0.5–0.8 with surface gauge measurements over the southeastern United States and less than 0.5 over other areas of the contiguous United States (CONUS). Ebert et al. (2007) showed that IR-based estimates underestimated summer precipitation by as much as 50% over the eastern CONUS, while they overestimated winter precipitation by 50%–100% throughout CONUS. Tian et al. (2010) found that among four different satellite-based precipitation datasets, the retrievals over the eastern United States have as much as a 32% overestimate in summer and a 48% underestimate in winter.

Currently the most practical approach to reduce these errors is to merge ground-based measurements from rain gauges or radar networks. A leading example is the TRMM Multisatellite Precipitation Analysis (TMPA) research product 3B42 version 6 (Huffman et al. 2007, 2009) produced at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center. After merging intercalibrated passive microwave (PMW) retrievals from multiple space-borne sensors, and filling PMW coverage gaps with IR-based estimates, 3B42 uses the monthly accumulation of global surface gauge measurements to rescale the satellite-based estimates in post–real time. This procedure results in estimates with substantially reduced biases, especially on the time scales of a month or longer. Smith et al. (2006) used the median of the long-term mean values from an ensemble of satellite-based products as the reference value, to estimate and reduce biases in satellite-based estimates, especially over the ocean when gauge data do not exist. This method works best when the errors in different satellite-based estimates are independent, but this condition is usually not satisfied. The Air Force Weather Agency (AFWA) has been producing a real-time global precipitation analysis, based on PMW and IR retrievals, superseded by gauge reports from the World Meteorological Organization (WMO)'s Global Telecommunication System (GTS). But the gauge reports available from GTS in real time are rather sparse, and this approach leaves some artifacts that result from the disparity between satellite-based estimates and isolated gauge reports (Tian et al. 2009). Recently Xiong et al. (2008) and Janowiak et al. (2009) proposed a procedure to correct the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) Morphing technique (CMORPH; Joyce et al. 2004) in real time. This procedure first performs bias correction for CMORPH with the (probability density function) PDF-matching technique against real-time global daily rain gauge data. Then it combines the corrected CMORPH with the gauge analysis itself with the optimal interpolation (OI) technique. Their test results over China showed substantial improvements in the merged CMORPH analysis. Over CONUS, Boushaki et al. (2009) used the real-time CPC daily gauge analysis (Higgins et al. 2000) to correct the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification Scheme (PERSIANN-CCS; Hong et al. 2004) data, and the corrected data show dramatic improvements over test areas in the southwestern United States.

A critical requirement in the existing gauge-correction schemes is the availability of gauge data as timely as that of the satellite-based estimates. Otherwise, gauge correction in real time is not possible; one has to resort to an approach similar to TMPA 3B42 with a monthly scale correction, at the price of a latency in the availability of such data products. However, the number of gauge reports available in real time over the globe is rather limited; a significant number of gauge reports, such as those within the WMO's GTS, are issued with various delays. In addition, many regions over the world, including the United States, are seeing many weather stations disappear recently (Stokstad 1999), leaving only historical data available. For example, there were over 18 000 daily gauge reports over United States before 2004, but there have been less than 10 000 since 2004 (Chen et al. 2008). In China, there had been ~700 hydrological stations before they ceased to operate in 1997 (Xie et al. 2007). Finally, collecting, processing, and quality-controlling real-time gauge reports are logistically tedious and complex. Therefore, it is highly desirable to explore a new approach to reduce the biases in the satellite-based estimates in real time, without depending on the timely availability of surface gauge observations.

In this study, we propose a scheme to explore this possibility. This new scheme is based on the observation that the error characteristics in satellite-based estimates are remarkably consistent. For example, over CONUS, estimates for summer show regularly positive biases (overestimates), while those for winter suffer negative biases (underestimates; Tian et al. 2009). In addition, most of the errors are caused by hit biases, meaning the satellite-based estimates have sufficient capability in detecting precipitation events, but are short in determining the correct rate rates of the events. Therefore, we developed a Bayesian approach to ''train'' an algorithm with the coincidental satellite and gauge data within a recent historical period. This algorithm essentially establishes a statistical relationship between coincidental gauge measurements and satellite estimates. Then we apply this ''learned'' relationship to real-time

satellite estimates, when gauge data are not available, to derive the mostly likely values of gauge measurements as the corrected satellite estimates.

This scheme was tested with nearly 6 yr of satellite-based estimates from 2 real-time products: TMPA 3B42RT and CMORPH (see section 3). We used the first 4 yr as the training period when historical gauge data are available. The last 2 yr were used as the correction period, during which the gauge data were withheld for correction but only used for validation. Our test results showed satellite biases were reduced by 70%–100% for the summers in the correction period. In addition, even when sparse networks simulated with only 600 or 300 gauges were used during the training period, the biases were still reduced by 60%–80% and 40%–60%, respectively. Thus, this method can potentially complement the existing real-time correction schemes.

Compared with other existing bias-correction schemes, this approach is theoretically more rigorous and generic, with minimal assumptions of the error characteristics in the data. We envision this scheme, if proven effective, could be used by data producers as an integral part of their processing algorithm, or by data consumers as a postprocessing step for their applications requiring low-bias precipitation estimates in real time. However, in practice, there are many factors affecting the effectiveness of this approach, which will be discussed in this paper. Because of these practical factors, this scheme can only yield incremental improvements, rather than an ultimate solution. Currently this scheme has been tested with real satellite-based data, but has not been employed to produce corrected data operationally.

The formulation of the Bayesian scheme is given in section 2. Data and methodology used in this study are described in section 3. Results are presented in section 4, followed by conclusions and a discussion in section 5.

2. Formulation

At a grid point, when both gauge measurements and satellite estimates are available, we denote their values as G_i and S_j, respectively. Then for the training period, we can compute the joint probability $P(G_i, S_j)$ from the existing data, and by the Bayesian theorem:

$$P(G_i|S_j) = \frac{P(G_i, S_j)}{P(S_j)}, \tag{1}$$

we can compute the conditional probability $P(G_i|S_j)$, which gives the likelihood of a gauge measurement G_i at the grid point when a satellite-based estimate has the value of S_j. Since all the terms on the right-hand side of Eq. (1) can be computed from the training data, we can calculate the most probable value of the gauge measurement G_{j0} given any S_j, which maximize $P(G_i|S_j)$. The end result from the training is a lookup table, which will be used during the correction period when only satellite estimates are available. We can then find G_{j0} from a given S_j in the lookup table, and we use G_{j0} to replace the corresponding S_j as the corrected satellite estimate value.

Theoretically, the most probable value of G_i, $(i = 1, 2, \ldots, N)$ should be the one that gives a unique and well-defined maximum of $P(G_i|S_j)$ for each S_j. In practice, however, the number of (G_i, S_j) sample pairs at an individual grid point is not always large enough to guarantee that, making the determination of maximum $P(G_i|S_j)$ unreliable. To increase the number of samples for more stable statistics, we encompassed the neighboring grid points within a limited range for the collection of the samples. The size of the neighborhood has to be small enough to keep the local precipitation regime and error characteristics uniform. In this study, we used a neighborhood within five grid boxes around the central box. In addition, we discretized the daily rain-rate values into logarithmic bins, to keep the number of strong events sufficiently large, and the events distribution closer to normal distribution. In this study, a constant size ratio of 1.16 between two adjacent bins was used. Finally, we used the average of all the G_i values for each particular S_j, weighted by $P(G_i|S_j)$, as an approximation of G_{j0}, instead of using the single value of G_i (which is often not unique), which gives the maximum of $P(G_i|S_j)$.

3. Data and methodology

Two real-time satellite-based estimates were used to test our correction scheme. One is the TMPA real-time product 3B42RT (Huffman et al. 2007, 2010), and the other is the CMORPH product (Joyce et al. 2004). Both products are purely satellite based; they have not been corrected in any way with ground-based gauge or radar measurements. TMPA 3B42RT derives its precipitation estimates primarily by merging the most recent PMW scans available from the array of sensors including the TRMM Microwave Imager (TMI), the Special Sensor Microwave Imager (SSM/I), the Advanced Microwave Sounding Unit-B (AMSU-B), and the most recent Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E). Then it uses IR-based retrievals from geostationary satellites to fill PMW coverage gaps. We note that 3B42RT data prior to February 2005 are considered obsolete because they were produced from an outdated algorithm, but for the purpose of this study we retained part of the outdated data for the training period, whose impact will be discussed later.

CMORPH's PMW-based retrievals are from almost the same set of sensors as 3B42RT, but it takes advantage of the high-resolution IR imagery to derive the advection of precipitation patches between PMW scans, and uses this information to obtain a smooth "morphing" of PMW rain patterns. Both the original 3B42RT and CMORPH used in this study have a 0.25° × 0.25° spatial resolution and a 3-h temporal resolution, and we used their daily accumulation to match that of the gauge data (below).

The gauge data we used to correct the satellite-based estimates are the NOAA CPC unified daily gauge analysis (CPC-UNI; Xie et al. 2007; Chen et al. 2008). This dataset employs an OI technique to reproject quality-controlled gauge reports over CONUS to a 0.25° × 0.25° grid. The OI-based interpolation has been shown to have higher correlation with individual gauge measurements than other techniques (Chen et al. 2008). There are about 6000–10 000 gauge reports daily in CPC-UNI for most of our study period.

For both the satellite-based and gauge-based data, we only train with and correct precipitation rates of 1 mm day^{-1} or higher, because estimates below 1 mm day^{-1} are much less reliable, and their contribution to the total precipitation amount is insignificant.

Our study period covers nearly 6 yr (5 yr and 8 months) from March 2003 to October 2008. The whole record is divided into two periods: the first period of nearly 4 yr (3 yr and 8 months) is designated as the training period, and the second period of 2 yr as the correction period. During the training period, both the gauge data and satellite-based estimates were used to establish the gauge–satellite statistical relationship. Then during the correction period, the gauge data were withheld for any training or correction. The correction was performed first, and then the corrected satellite data were evaluated against the withheld gauge data. Since the correction will work for the satellite data within the training period as well, we also went back to correct them for the first 4 yr.

During the training period, we further split each annual cycle into a warm season (from May to October) and a cold season (from November to April), resulting in four warm seasons and three cold seasons. The training was performed separately for the two different seasons. This is based on the observation that the error characteristics have a strong seasonal dependency (Ebert et al. 2007; Sapiano and Arkin 2009; Tian et al. 2009, 2010); the inconsistent performance between warm and cold seasons will degrade the training of the satellite data if lumped together. Ideally the annual cycle could be split into more segments, such as four seasons, if the training data are abundant. But with the current amount of training data available, the two-season separation is a compromise between the amount of training data and the resolution of the seasonality.

To investigate the impact of gauge density on this scheme, we use the gauge data to simulate gauge networks with three different densities:

1) A really dense network with gauge analysis data at every grid point over CONUS used. We denote this density as "100%."
2) A sparse network with 600 gauges over CONUS, whose locations were randomly picked from the CPC-UNI analysis. We denote this density as "10%," as the number of the simulated gauges is nearly 10% of the minimum number of gauges used in CPC-UNI analysis.
3) An even sparser network with only 300 gauges over CONUS, similarly picked from the CPC-UNI analysis data, and we denote this density as "5%."

The three gauge densities are illustrated in Figs. 1a–c. The correction scheme was tested with each of the three gauge densities, and the corresponding corrected data were all evaluated against the gauge analysis of the original density (100%). By testing with sparse networks, we can assess how the scheme would work in other areas over the globe where dense gauge deployments are rare.

4. Results

a. Bias correction

The correction scheme proved to be able to reduce the biases in satellite-based estimates substantially. Figure 1 shows comparisons of biases between the original, uncorrected 3B42RT and CMORPH, and their corrected counterparts produced with the three respective gauge densities (100%, 10%, and 5%) shown in the left column. During the last warm season in the correction period (May–October 2008), both of the original 3B42RT and CMORPH (Figs. 1d,h) show similar, large areas of overestimates over the central CONUS, a well-documented characteristics (Gottschalck et al. 2005; Tian et al. 2007, 2009). Corrections with 100% gauge data effectively removed these overestimates (Figs. 1e,i), considerably improving the overall estimates over CONUS. Some smaller areas of overestimates along the East Coast and over southern Florida, especially for 3B42RT, were also eliminated. But the narrow stripe of positive biases in southern California was largely not affected (Fig. 1e).

Corrections with 10% and 5% gauge densities also exhibit remarkable improvement, to a varying degree. With the gauge density of 10% (600 gauges), the positive

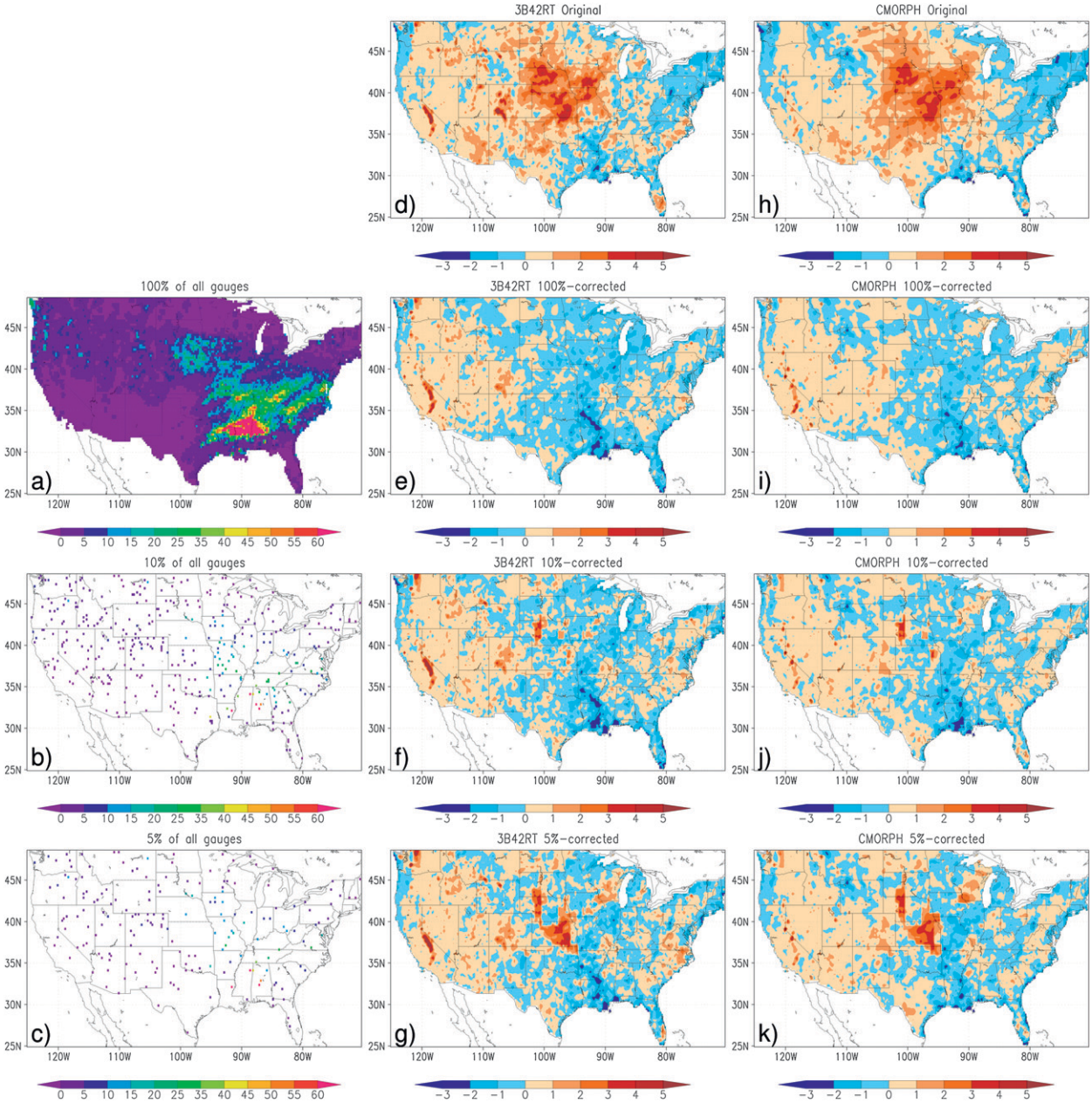

FIG. 1. (a)–(c) Three gauge densities and average biases (mm day^{-1}) before and after the correction for the last warm season (May–October 2008) for (d)–(g) 3B42RT and (h)–(k) CMORPH. The color bars for (a)–(c) are the daily rain rates reported by each gauge for 1 Mar 2003 as a sample day.

biases over central CONUS were largely removed, but not completely (Figs. 1f,j). Several small patches of biases still existed, such as the ones in western Nebraska and eastern Kansas. Further examination reveals that the gauge distribution over these areas is much sparser (Fig. 1b), and it is reasonable to expect the performance to degrade in such cases. With a gauge density of 5% (300 gauges), those areas saw even fewer gauges (Fig. 1c), and consequently the areas of less-corrected biases were larger (Figs. 1g,k). Obviously a too sparse gauge network does not offer much help to our correction scheme or to other schemes.

Bias reduction for cold seasons (not shown) is not significant. This is because for these seasons, both 3B42RT and CMORPH are dominated with underestimates over CONUS, especially over higher latitudes and mountainous areas (Tian et al. 2009). These underestimates are primarily caused by missed detection by satellite-based

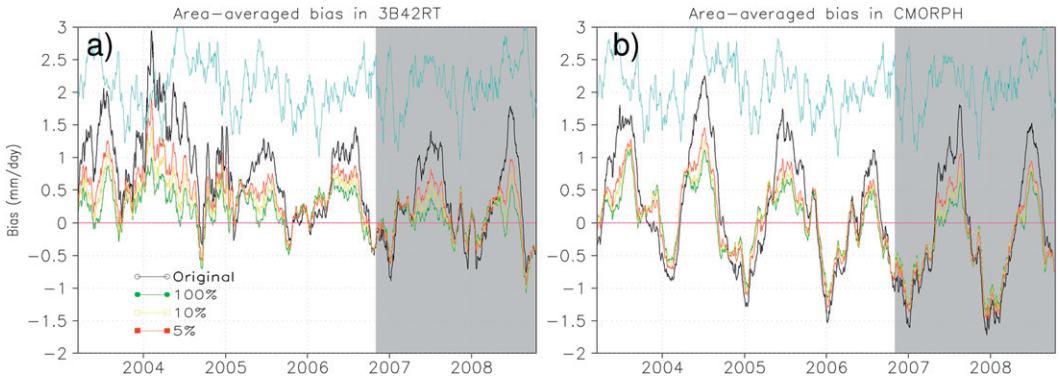

FIG. 2. Time series of biases before and after correction with different gauge densities for (a) 3B42RT and (b) CMORPH. The shaded area indicates the 2-yr correction period during which no gauge data were available for correcting the biases in satellite-based estimates. The light blue curve in the background is the area-averaged precipitation from gauge analysis, to help readers assess the relative amplitudes of the biases.

estimates over cold land surfaces such as snow and ice cover, whereas our current scheme can only correct satellite retrievals when they are detected events or false alarms.

Figure 2 shows the time series of the biases before and after the corrections, for the whole study period, including both the training (white area) and the correction period (shaded area). The biases in the original estimates (black curves) show a roughly regular seasonal dependence, especially for CMORPH (Fig. 2b): during warm seasons the satellite data tend to overestimate, while during cold seasons they tend to underestimate. For the reasons mentioned above, our scheme provides limited improvement to the satellite estimates during cold seasons, especially during the correction period. However, for the warm seasons, the reduction in biases is rather substantial. Specifically, both summers of 2007 and 2008 saw decreases of the overestimates in either 3B42RT (Fig. 2a) or CMORPH (Fig. 2b) by over 50% after the scheme was trained with 100% of the gauges (green curves). The reduction of the biases was gradually smaller with 10% and 5% gauge densities (yellow and red curves), but was still significant. A similar amplitude of reduction was achieved for the 4-yr training period.

The fact that our scheme did not completely eliminate the biases in either 3B42RT or CMORPH even in summer is primarily related to the inevitable inconsistencies in the error characteristics of the original data. For example, the first 2 yr of data from 3B42RT (produced from an obsolete algorithm) have distinctively different error behavior from other years. For most of the time during that period it shows overestimates, except several brief episodes during the winters. Such a different period will contribute misleading information to the scheme during the training period, reducing the accuracy of the correction.

Table 1 summarizes the seasonal-scale biases before and after the corrections for the two winters [December–February (DJF)] and summers [June–August (JJA)] in the correction period (i.e., the last two years). For either 3B42RT and CMORPH, the reductions in biases with 100% gauges are dramatic during the two summers: 3B42RT saw over 80% decrease in biases for the first summer and nearly 100% for the second, and CMORPH

TABLE 1. Average biases (mm day^{-1}) for the last two winters (DJF) and summers (JJA). The percentage numbers are the reductions in biases relative to the biases in the uncorrected data.

	3B42RT				CMORPH			
Season	Original	100	10	5	Original	100	10	5
DJF 2007	−0.23	−0.01	−0.05	−0.08	−1.15	−0.87	−0.90	−0.98
		97%	79%	67%		24%	21%	15%
DJF 2008	−0.40	−0.17	−0.21	−0.26	−1.45	−1.19	−1.22	−1.30
		58%	46%	35%		18%	15%	10%
JJA 2007	1.15	0.22	0.41	0.61	1.23	0.31	0.47	0.65
		81%	64%	47%		75%	62%	47%
JJA 2008	0.97	0.01	0.19	0.36	1.05	0.30	0.39	0.51
		99%	81%	63%		72%	63%	51%

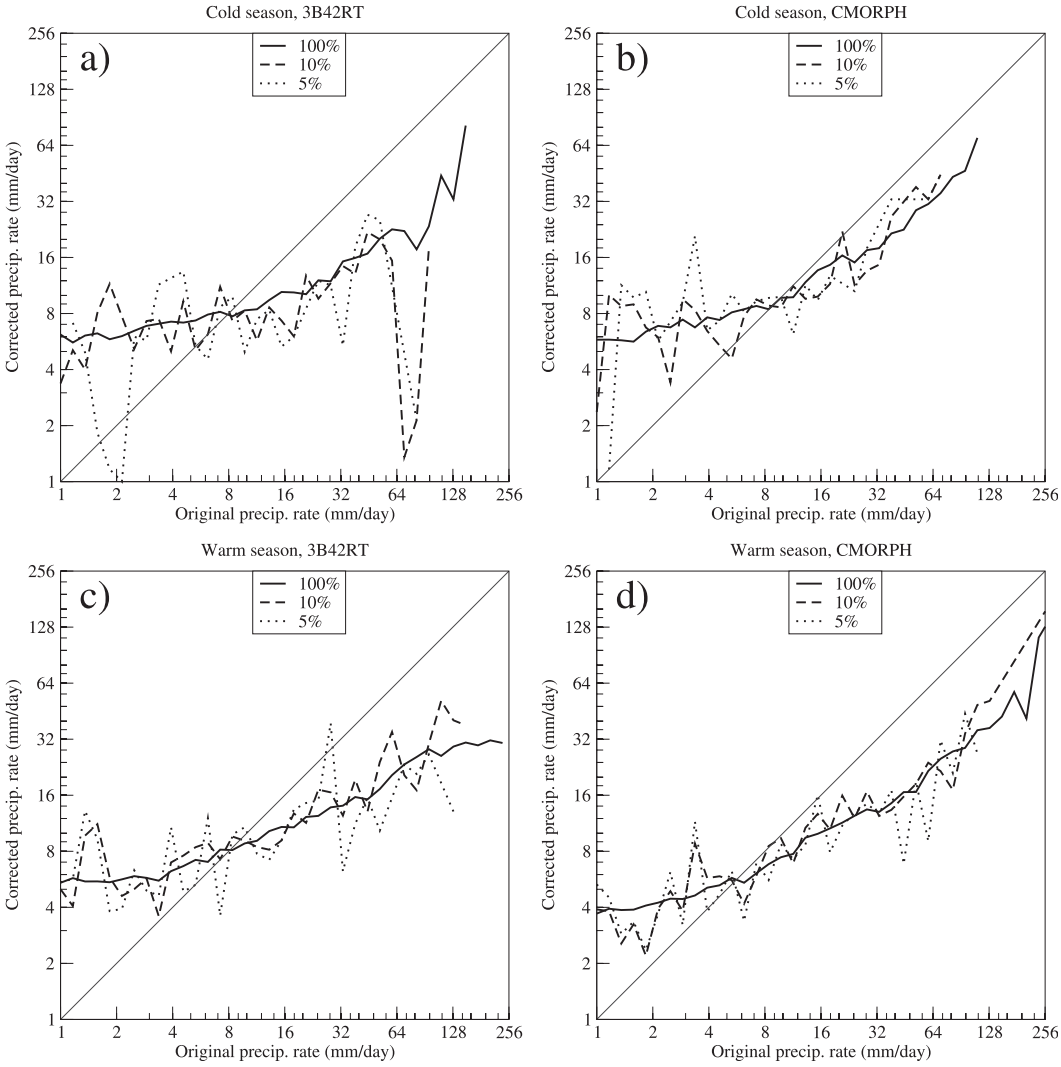

FIG. 3. Correction curves for a sample location in the central United States (42°N, 96°W) during the (a),(b) cold season and (c),(d) warm season from the three gauge densities for 3B42RT and CMORPH, respectively. The diagonal line indicates no correction.

enjoyed 75% and 72% reductions in overestimates for the two summers, respectively. With 10% and 5% gauge densities, the degree of improvement became less pronounced, but was still around 60% or more for 600 gauges, and not far below 50% for 300 gauges. For the two winters, the reduction of biases in CMORPH is rather limited, with amplitudes around 10%–20% for all the three gauge densities, as our scheme cannot improve on underestimates caused by missed detection of precipitation events. The dramatic percentage numbers in 3B42RT for the two winters are rather misleading: the seasonal-scale biases in 3B42RT in the original data are already very low (less than 30% of CMORPH's; also see Fig. 2), so any slight improvement would result in large percentage changes.

b. *Correction curves*

As indicated in section 2, after the training was done, we obtained a lookup table at every grid point mapping a value of satellite estimates to a most likely value of gauge measurements. We denote such a mapping a "correction curve." By examining a correction curve, we can have a better understanding how the correction scheme works. As a sample, Fig. 3 shows the different correction curves at a location in the central United States (42°N, 96°W) corresponding to cold and warm seasons and to the 3 gauge densities, for both 3B42RT and CMORPH.

All the correction curves showed a similar feature: the amplitude of strong events (>16 mm day^{-1}) will be

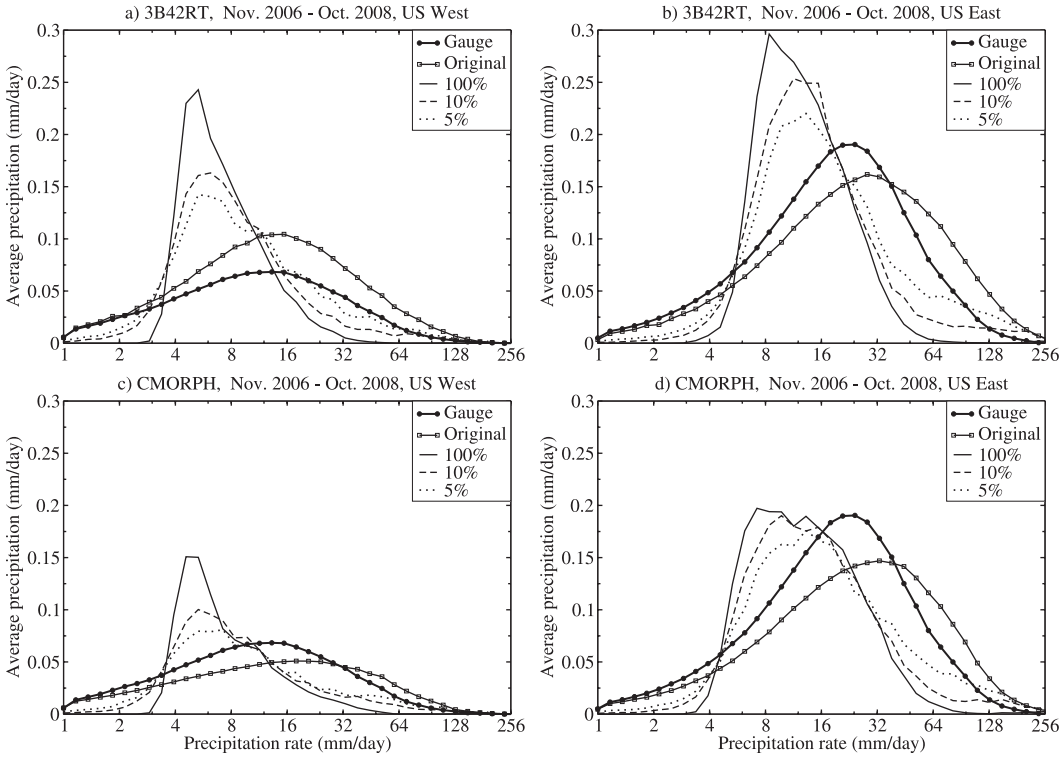

FIG. 4. Daily precipitation amount (mm) as a function of precipitation rate (mm day^{-1}) for (a),(b) 3B42RT and (c),(d) CMORPH over the (a),(c) western and (b),(d) eastern CONUS, respectively, for the original and corrected satellite-based estimates, compared with the gauge data. The CONUS was separated into western and eastern regions along the 100th meridian. The logarithmic scale was used to bin the precipitation rates.

reduced, while that of light events (<8 mm day^{-1}) will be increased. This is more obvious for 3B42RT (Figs. 3a,c). For both 3B42RT and CMORPH, light events less than 4 mm day^{-1} will be nearly uniformly boosted to over 4 mm day^{-1}. This indicates the satellite estimates tend to underestimate light rain by several times during the training period.

The correction curves for the same season and same dataset but from different gauge densities are largely similar. This confirms the results shown in Figs. 1 and 2 that sparse gauge distributions can do a reasonable job as well. As expected, the correction curves obtained with 100% gauges are fairly smooth, while those with 10% and 5% have considerable fluctuations, because of the reduced number of data samples available during the training period, especially during the cold season. At about 70 mm day^{-1}, 3B42RT showed a heavy dip for either 10% or 5% gauge density, probably caused by a few erroneous events when 3B42RT detected strong precipitation whereas the gauges reported little precipitation. The reduced number of samples made the impact of such events outstanding, and such a dip will cause significant overcorrection around this rate.

From the nonlinearity of the correction curves we can see one of the advantages of this scheme: it is entirely self-adaptive. The scheme automatically generates correction curves from the training data themselves, without any predefined error models or tunable parameters. Moreover, as new data become available in post–real time, the training can be repeated to assimilate the new information and update the correction curves.

c. *Impact on daily events*

To examine the impact of the correction scheme on event-scale precipitation values, we computed and compared the intensity distribution of the daily precipitation for original and corrected satellite estimates, with that of the gauge data, for the correction period, in Fig. 4. Before the corrections, 3B42RT has about 1/3 more precipitation events than the gauge data over the western CONUS (Fig. 4a), but has slightly less light precipitation and more strong precipitation events (>32 mm day^{-1}) over the eastern CONUS (Fig. 4b). CMORPH has less light and more strong precipitation events over both the western and eastern part of the continent (Figs. 4c,d). After the corrections, both the amount of strong precipitation and light precipitation are severely reduced

for both dataset, especially for 3B42RT, and for both datasets over the western CONUS, and intermediate rain rates were amplified. Consequently, the intensity distributions became narrower, indicating the precipitation fields were more uniform with the corrections, especially with the 100% gauge density (Fig. 4).

This behavior is undesirable for many applications. The correction scheme improves on the amount of long-term and large-scale estimates of precipitation amount, at the price of shrinking the dynamical range of event-scale estimates, and reducing the number of extreme events in particular. This will adversely affect applications that are more sensitive to extreme events, such as floods and landslides.

Such overcorrections are most likely caused by the inconsistencies of the error characteristics during the training period and between the training and the correction period. As seen from Fig. 2, during the training period both 3B42RT and CMORPH have higher overestimates during warm seasons. The most notable case is the first two years for 3B42RT (Fig. 2a) during which the number of strong events are obviously overestimated. Our correction scheme will ''learn'' to shrink the amplitude of these events much more to reduce the total biases. When this is applied to the correction period, 3B42RT itself is already much improved, therefore the information acquired during the worse performance years will lead to overcorrections for the last two years.

The reduction in the amount of light rain can also be explained. As shown in Tian et al. (2009), most of the underestimates during winters are caused by missed detection of precipitation events, chiefly light in intensity ($<$10 mm day^{-1}). The current scheme cannot reinstate those missed events. To reduce the underestimates, it has to increase the amplitude of the already detected light events, moving them into the range of more intermediate intensity. This is more the case for CMORPH, as it has higher underestimates during the last years (Fig. 2b).

One possible approach to correct the intensity distribution is using the PDF-matching method, which has been widely applied for satellite data intercalibration (e.g., Turk and Miller 2005; Huffman et al. 2007) as well as satellite–gauge data merging (Janowiak et al. 2009). However, this is just a practical solution to a problem caused by other sources—primarily the consistency of the data between the training and correction periods (see section 5), and this method will introduce additional uncertainties in the corrected data. In theory, if the error characteristics in the satellite-based estimates are consistent, the Bayesian scheme can both reduce the biases and match the PDFs of the gauge data. The PDF mismatch is one of the manifestations of the data inconsistency, not the failure of the Bayesian scheme itself.

5. Summary and discussion

In this study, we proposed and tested a new approach for reducing the biases in the satellite-based estimates in real time. The major advantage of this approach over other existing methods is its applicability when real-time gauge data are not available. It takes advantage of the availability and relative abundance of post-real-time gauge measurements, to train an algorithm based on the Bayesian inference. The training period yields spatially and temporally varying statistical relationships between post-real-time satellite-based estimates and gauge measurements. Then such statistical relationships are used to correct real-time satellite estimates when gauge data are unavailable, thus eliminating the acute need of real-time gauge reports as in other existing schemes. This scheme is also entirely self-adaptive: it adjusts itself to different training datasets without the need of any predefined, tunable parameters.

We tested the scheme with two satellite-based products, TMPA 3B42RT and CMORPH, using CPC's unified gauge analysis for training and validation, for a period of 6 yr. The first 4 yr were used as the training period, during which the gauge data were used to guide the Bayesian algorithm, while the following 2 yr were used as the correction period when the gauge data were not used except for evaluation. The results showed that the scheme worked particularly well for summer, having nearly completely reduced the high overestimates over most of the central CONUS for both 3B42RT and CMORPH (Figs. 1 and 2).

A subset of the gauge data were also used to simulate sparse networks and to investigate their effectiveness. Even with only 600 or 300 gauges randomly distributed over CONUS for the training period, the bias reductions ranged from 60% to 80% for the former and from 47% to 63% for the latter. This suggests that our scheme can be used over other areas of the globe, where sparser networks are more common. However, when a network gets too sparse (e.g., 300 gauges), the scheme will leave ''holes'' of uncorrected regions (Fig. 1).

There are also some undesirable but explainable effects resulted from this correction procedure with the two sample datasets, including the narrowing of the rain intensity distribution particularly (Fig. 4). The drawbacks in the corrected data are not an indicator of the method's limitation; rather, they are the manifestation of the training data quality that affects the effectiveness of the method. Specifically, the lack of consistent and abundant training data is the leading cause. The

underlying assumption for our scheme is the consistency of the error characteristics in the satellite-based estimates, including their seasonal and regional dependence in particular. While largely the case, there are considerable inconsistencies between the training period and the correction period within either 3B42RT or CMORPH. As a result, on the daily time scale, the intensity distributions of precipitation were modified by overcorrection of both strong (>20 mm day^{-1}) and light (<4 mm day^{-1}) events, producing precipitation fields with reduced spatial and temporal variability (Fig. 4). From this perspective, we would argue that, if the satellite retrievals could not remove errors, it would be much better to keep the errors consistent. Additionally, the limited availability of the training data leads to less reliable correction for strong events. Therefore, these drawbacks in the results should not discount the validity of the proposed method, which is theoretically more rigorous than other existing solutions with minimal assumptions about the data and their error characteristics. Other empirical approaches, such as linear or nonlinear regression, Kalman filtering, or optimal interpolation, are just special cases of the Bayesian approach (Wikle and Berliner 2007).

The current practical constraints, namely, inconsistencies in error characteristics and lack of sufficient training data, make it too early to apply this method in operational use. However, we envision the proposed approach will be revisited when both the quality and quantity of the training data are improved in the future.

We also want to note that this method is designed only to reduce the systematic errors (biases), not the random errors. The latter are difficult to quantify or correct, especially when the uncertainties in the ground reference data themselves are not known (Barnston 1991; Krajewski et al. 2000). However, the random errors will be affected simply because of the rescaling of the satellite estimates, but this effect can not be quantified in the current study. In addition, the gauge analyses themselves are not accurate presentations of the true area rainfall either (e.g., Villarini and Krajewski 2008); the uncertainties in the gauge analyses contribute directly to the total uncertainties between the gauge analyses and the satellite data (Barnston 1991). With the errors in the gauge analyses unknown in the present study, our working assumption is that the errors in the gauge analyses are much smaller than those in the satellite-based estimates, and reduction in the biases relative to the gauge data is an incremental improvement to the satellite data. This assumption is shared by the other existing gauge-correction schemes and largely true at least over CONUS because of the dense gauge networks (Tian et al. 2009).

Finally, the effectiveness of this scheme shown in this study is by no means to discount the value of real-time gauge measurements. Obviously this scheme works best for error corrections on seasonal scales or longer, and it cannot improve performance metrics such as probability of detection or false-alarm rate. Only through real-time gauge data, preferably from dense networks, intelligently merged with satellite-based estimates, can such performance metrics be improved. Therefore, our scheme can complement, rather than replace, existing techniques to improve real-time precipitation estimates from multiple sensors.

Acknowledgments. This research is supported by the Air Force Weather Agency MIPR F2BBAJ6033GB01. The authors wish to thank Dr. Kenneth Harrison for reviewing our manuscript, and Mingyue Chen, Mathew Sapiano, Dan Braithwaite, Hiroko Kato Beaudoing, and Pingping Xie for helpful discussions and for assistance with data access.

REFERENCES

Barnston, A. G., 1991: An empirical method of estimating rain gauge and radar rainfall measurement bias and resolution. *J. Appl. Meteor.,* **30,** 282–296.

Boushaki, F. I., K. L. Hsu, S. Sorooshian, G. H. Park, S. Mahani, and W. Shi, 2009: Bias adjustment of satellite precipitation estimation using ground-based measurement: A case study evaluation over the southwestern United States. *J. Hydrometeor.,* **10,** 1231–1242.

Chen, M., W. Shi, P. Xie, V. B. S. Silva, V. E. Kousky, R. Wayne Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.,* **113,** D04110, doi:10.1029/2007JD009132.

Ebert, E. E., J. E. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.,* **88,** 47–64.

Gottschalck, J., J. Meng, M. Rodell, and P. Houser, 2005: Analysis of multiple precipitation products and preliminary assessment of their impact on Global Land Data Assimilation System land surface states. *J. Hydrometeor.,* **6,** 573–598.

Habib, E., A. Henschke, and R. F. Adler, 2009: Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmos. Res.,* **94,** 373–388.

Higgins, R. W., W. Shi, and E. Yarosh, 2000: Improved United States precipitation quality control system and analysis. *NCEP/Climate Prediction Center Atlas 7,* National Centers for Environmental Prediction, 40 pp.

Hong, Y., K. Hsu, X. Gao, and S. Sorooshian, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteor.,* **43,** 1834–1853.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.,* **8,** 38–55.

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

Janowiak, J., R. Joyce, and P. Xie, 2009: CMORPH improvements: A Kalman filter approach to blend various satellite rainfall estimate inputs and rain gauge data integration. *Geophysical Research Abstracts,* Vol. 11, Abstract EGU2009-9810. [Available online at http://meetingorganizer.copernicus.org/EGU2009/EGU2009-9810.pdf.]

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

Krajewski, W., G. Ciach, J. McCollum, and C. Bacotiu, 2000: Initial validation of the global precipitation climatology project monthly rainfall over the United States. *J. Appl. Meteor.,* **39,** 1071–1086.

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

Smith, T. M., P. A. Arkin, J. J. Bates, and G. J. Huffman, 2006: Estimating bias of satellite-based precipitation estimates. *J. Hydrometeor.,* **7,** 841–856.

Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.,* **81,** 2035–2046.

Stokstad, E., 1999: Scarcity of rain, stream gages threatens forecasts. *Science,* **285,** 1199–1200.

Tian, Y., C. D. Peters-Lidard, B. J. Choudhury, and M. Garcia, 2007: Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *J. Hydrometeor.,* **8,** 1165–1183.

——, and Coauthors, 2009: Component analysis of errors in satellite-based precipitation estimates. *J. Geophys. Res.,* **114,** D24101, doi:10.1029/2009JD011949.

——, C. D. Peters-Lidard, R. F. Adler, T. Kubota, and T. Ushio, 2010: Evaluation of GSMaP precipitation estimates over the contiguous United States. *J. Hydrometeor.,* **11,** 566–574.

Turk, F. J., and S. D. Miller, 2005: Toward improving estimates of remotely-sensed precipitation with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens.,* **43,** 1059–1069.

Villarini, G., and W. F. Krajewski, 2008: Empirically-based modeling of spatial sampling uncertainties associated with rainfall measurements by rain gauges. *Adv. Water Resour.,* **31,** 1015–1023.

Wikle, C. K., and L. M. Berliner, 2007: A Bayesian tutorial for data assimilation. *Physica D,* **230,** 1–16.

Xie, P., and Coauthors, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.,* **8,** 607–626.

Xiong, A., P. Xie, J.-Y. Liang, Y. Shen, R. J. Joyce, J. E. Janowiak, and P. A. Arkin, 2008: Merging gauge observations and satellite estimates of daily precipitation over China. *Proc. Fourth Int. Precipitation Working Group Workshop,* Beijing, China, WMO, 358–363.

Yilmaz, K. K., T. S. Hogue, K.-L. Hsu, S. Sorooshian, H. V. Gupta, and T. Wagener, 2005: Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeor.,* **6,** 497–517.