NOTES AND CORRESPONDENCE

Comparison of Rain Gauge Measurements in the Mid-Atlantic Region

ALI TOKAY

Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, Baltimore, and NASA Goddard Space Flight Center, Greenbelt, Maryland

PAUL G. BASHOR

Computer Sciences Corporation, NASA Wallops Flight Facility, Wallops Island, Virginia

VICTORIA L. MCDOWELL

Department of Earth and Environmental Sciences, Susquehanna University, Selinsgrove, Pennsylvania

(Manuscript received 10 December 2008, in final form 1 September 2009)

ABSTRACT

A comparative study of daily and monthly rainfall between research and operational gauges was conducted at the mid-Atlantic region. Fifty *research* tipping-bucket gauges were deployed to 20 sites where each site had dual or triple gauges. The gauges were in place to validate the National Aeronautics and Space Administration's newly developed polarimetric radar rainfall estimate. For logistic purposes, these research gauges were collocated with *operational* gauges and were operated over a year at each site. Therefore, this is an experimental study, which involves a mixture of one to five sites of seven operational gauge networks.

A very good to excellent agreement between the two collocated research gauges at daily time scale raised the authors' confidence to consider them as a reference before comparing with the operational gauges. Among operational networks, the National Weather Service's (NWS) Automated Surface Observing Systems (ASOS) weighing bucket and the Climate Reference Network and Forest Services tipping-bucket gauges demonstrated high performance for both daily and monthly rainfall, while the Federal Aviation Administration's Automated Weather Observing Systems (AWOS) tipping-bucket gauges performed poorly. Among the other networks, the ASOS tipping-bucket and Cooperative observer program's stick gauges seemed to be reliable for monthly rainfall, but not always for daily rainfall. The Virginia Agricultural Experimental Station (VAES) tipping-bucket gauges, on the other hand, had a mixture of high and low performance for daily and monthly rainfall. Unlike other gauge networks, VAES gauges were in place for long-term research applications.

1. Introduction

Rain gauges have historically been employed to determine the precipitation climatology of a city or a region. The climatological variables including frequency, norms, anomalies, and extremes are derived relying upon the entire record of the gauge measurements. Rain gauge measurements have also been incorporated in mapping monthly mean global precipitation (Legates and Willmott 1990) and give invaluable input for weather and flood forecasting models. Forecasting models require gauge measurements at an hourly or higher temporal scale, while daily gauge measurements are typically sufficient for climate applications.

Rain gauges are a valuable asset to validate radar rainfall products. For direct validation, gauges are ideally distributed near the radar site just outside the radar's surface clutter. This configuration reduces the differences in the sample volumes and the time–height ambiguity between the two measurements (e.g., Matrosov et al. 2002). For statistical validation, one seeks the agreement on the distributions of rainfall between the radar and

Corresponding author address: Ali Tokay, NASA Goddard Space Flight Center, Code 613.1, Greenbelt, MD 20771.
E-mail: ali.tokay-1@nasa.gov

DOI: 10.1175/2009JHM1137.1

© 2010 American Meteorological Society

gauge measurements (Amitai 2000). Rain gauge measurements are also incorporated in radar rainfall estimation. The probability matching method (PMM), for instance, employs the coincident measurements of radar reflectivity Z and gauge rain rate R in developing a Z–R lookup table (Rosenfeld et al. 1993, 1994). The National Aeronautics and Space Administration (NASA) Tropical Rainfall Measuring Mission (TRMM) satellite validation program employs PMM in constructing radar rainfall mapping in their primary validation sites (Wolff et al. 2005). The gauge measurements are also used to adjust the radar rainfall estimation. The National Oceanic and Atmospheric Administration (NOAA) stage III product is the gauge adjusted hourly radar rainfall at 4-km resolution (Xie et al. 2005).

Similar to radar rainfall products, rain gauges are a valuable asset to validate the satellite rainfall estimates. The daily rainfall from TRMM Multisatellite Precipitation Analysis (TMPA), for instance was compared to the gauge measurements at atolls, buoys, and dense gauge sites (Huffman et al. 2007). Interestingly, the TMPA also employs monthly gauge totals in constructing 3-hourly rainfall maps at $0.25° \times 0.25°$ resolution. The rain gauges that were incorporated in the radar or satellite rainfall products are mainly operated by operational agencies such as the National Weather Service (NWS). The institution that merges the radar or satellite rainfall products with the gauges is responsible for the gauge quality control. The gauges that were used in validating radar or satellite product are mostly operated by a research organization, which is also responsible for the quality control of the gauges.

The rain gauge network in east-central Florida, for instance, was constructed to provide an independent validation for the Weather Surveillance Radar-88 Doppler (WSR-88D) and S-band polarimetric radar in Melbourne, Florida, during the TRMM field campaign, Texas Florida Underflights (TEFLUN-B; Wolff et al. 2005). The same gauge network was also used to study the spatial variability of rainfall (Habib et al. 2001a; Datta et al. 2003; Habib and Krajewski 2002). There are other well-maintained research quality rain gauge networks across the United States and elsewhere. The gauge networks that were developed for hydrological applications covers the subbasins of a watershed (e.g., Habib et al. 2004, 2008), while multipurpose mesonets cover a region as in west Texas (Schroeder et al. 2005) or a state as in Oklahoma (Brock et al. 1995).

The rain gauge network that was used in this study was originally deployed to validate NASA's newly developed S-band polarimetric radar (NPOL). Through collaborative efforts with the NWS Wakefield Office and other federal and state agencies, the gauges were collocated with various operational networks. In this study, we take this opportunity and compare the operational and research gauges by examining their daily and monthly rain totals. The operational gauges are used in constructing radar and satellite rainfall products and, therefore, it is important to assess the accuracy of their measurements. Indeed, this is the main motivation of this study. Since dual or triple research gauges were deployed in each site, the agreement between the research gauges was first done prior to considering them as a reference to evaluate the performance of the operational gauges. We refer to *research gauges* as the instruments in the field for a particular application and expected to be removed once the research goals had been achieved.

The rain gauge networks including site and type of gauges are stated in section 2. Section 3 describes the rainfall statistics used to evaluate the gauges. The performance of research gauges is presented in section 4, while section 5 is dedicated to the performance of operational gauges. The concluding remarks are given in the last section.

2. Rain gauge network

The NASA TRMM satellite validation office (TSVO) rain gauge network consisted of 50 Met One Inc. tipping-bucket gauges at 20 sites. Half of the sites had dual gauges and the other half had triple gauges. The gauge sites were distributed between 15 and 150 km of the NPOL except one at the radar site: Oyster, Virginia. The gauge sites were mostly in coastal areas on both sides of the Chesapeake Bay, in Delmarva, eastern Virginia, and northeast North Carolina. The gauges covered a $2.5° \times 2.5°$ area where gauge distances ranged from 7 to 269 km. Sixteen of 20 gauge sites were collocated (less than 5 m) with operational gauges that were operated by federal or state agencies (Fig. 1). This study concentrates on these 16 sites.

The TSVO rain gauges were deployed beginning in February 2004 and all the gauges were sited within 6 months (Fig. 2). The gauges were visited monthly over the 25-month period between April 2004 and May 2006 except for a winter break of 4 months in both years. Regarding data recording, the time of the tip (1-s resolution) was stored in a data logger that was manufactured by MadgeTech. During each site visit, the performance of gauge, logger, and battery was tested and data was downloaded. The logger was enclosed in a waterproof box that was placed inside the gauge cylinder and therefore, the connection cables between the gauge and the logger were not exposed to the weather.

Throughout the operation, six gauge sites were relocated. During late October 2004, the gauge site (G04)

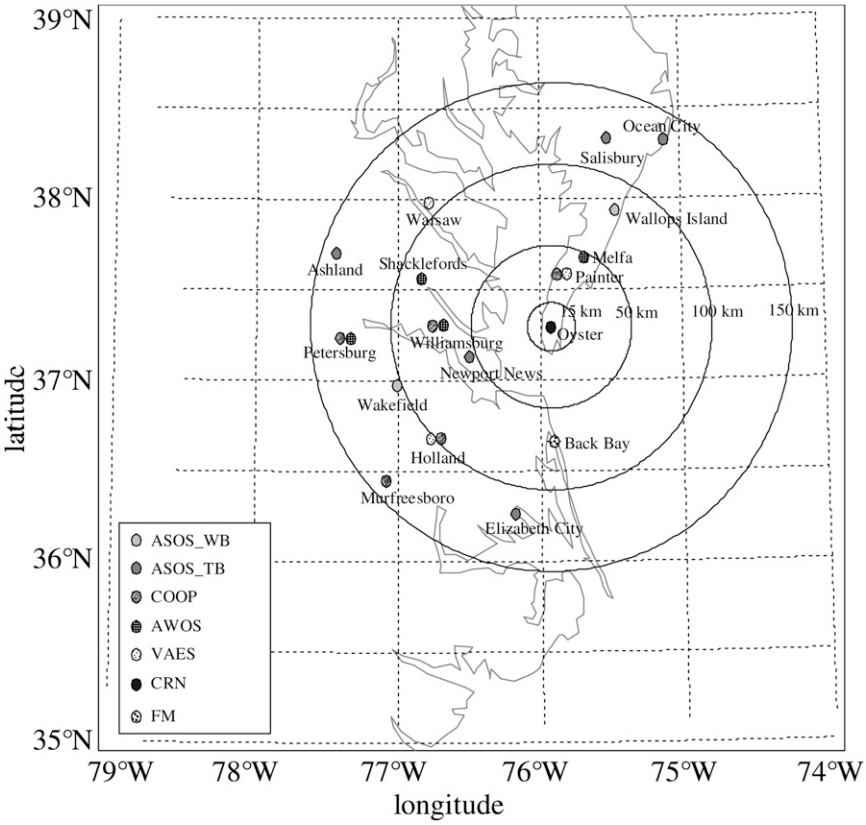

FIG. 1. Map of TSVO gauge sites, which were located next to operational gauges listed at the bottom-left corner. The TSVO gauges were distributed around the NPOL. The radar coverages of 15, 50, 100, and 150 km are also shown.

next to the radar was moved nearby to the U.S. Climate Reference Network (CRN) site. During the May 2005 site visit, two gauge sites (G09 and G10) next to Cooperative Observer Program (COOP) gauges were moved to new sites (G14 and G15) next to Automated Weather Observation Systems (AWOS) gauges. Similarly, two other sites (G08 and G11) that were also next to COOP gauges were moved to new sites (G17 and G19) adjacent to the Virginia Agricultural Experiment Station (VAES) gauges. The relocated gauges stayed in the same town where the distances between the COOP and the AWOS or VAES sites ranged from 432 m to 12 km. The COOP site (G20) in Murfreesboro, North Carolina, was closed and a new site (G21) was opened next to the AWOS site in Melfa, Virginia. At the end, the TSVO gauges were operated at least one year next to or very near seven ASOS, five COOP, four AWOS, three VAES, one CRN, and one Fire Management (FM) gauge.

Most of the operational gauges were tipping buckets, but the gauges in each network had different manufacturers and data collection systems (Table 1). Considering the ASOS networks, two sites had all weather precipitation accumulation (AWPAG) gauges (G01 and G02), while the remaining five sites were equipped with a heated Frise tipping-bucket (FTB) gauges (G03–G07). The AWPAG is a weighing bucket gauge and has proven to be superior to the heated FTB in mixed and frozen precipitation (White et al. 2004; Greeney et al. 2007). Since late 2003, the AWPAG have been replacing FTB across the United States and for the G01 and G02 sites, the replacements occurred in October 2004. In that regard, we disregarded the G01 and G02 data prior to the gauge replacements. Regarding maintenance, the performance of the gauges was remotely monitored and a site visit was scheduled if a problem was identified. In addition, preventive maintenance was performed every 3 months (D. Jones, NWS Wakefield Office, 2008, personal communication).

The COOP sites (G08–G12) had manual stick gauges (Table 1). The COOP is a voluntary based program that has provided climate and weather data for over a century. The gauge is read daily but the time of reading is not recorded even though the reading occurs at the same hour or within a range of half an hour or an hour depending on the station. The gauge reading is performed even in the presence of precipitation except that the

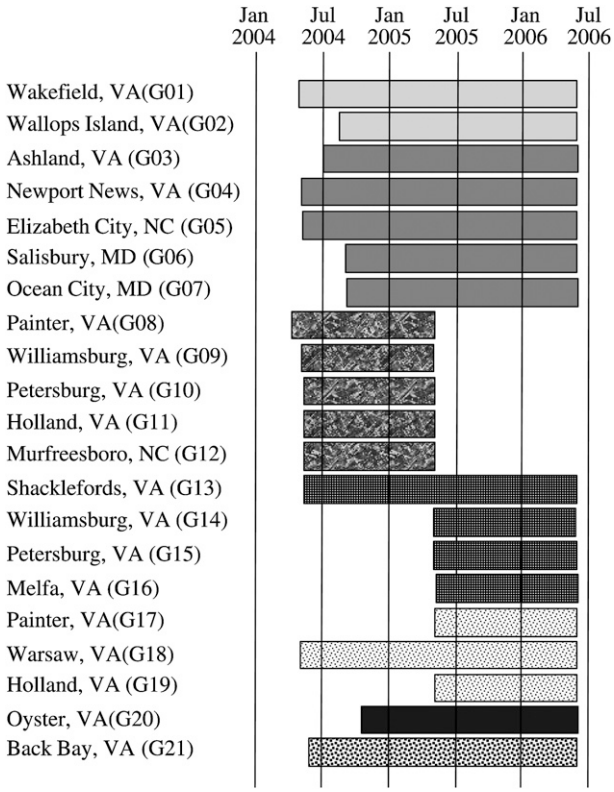

FIG. 2. Operation period of TSVO gauges. The gauge identification numbers are also shown.

gauge is emptied during the next reading cycle. Since the time of reading is not recorded, the daily rainfall can be erroneous if the precipitation occurs at the time of reading. Through close collaboration with the gauge reading personnel at each site, we set the hours shown in Table 2 to determine the daily rainfall.

The AWOS sites (G13–G16) were equipped with tipping-bucket gauges (Table 1). The Federal Aviation Administration (FAA) administrates the AWOS sites under contract with two private companies. The vendor of the tipping-bucket gauges is therefore not uniform across the United States. The maintenance is similar to the ASOS (R. Conlon, Austin Electronics, 2006, personal communication).

The VAES sites (G17–G19) were equipped with tipping-bucket gauges (Table 1) that were attached to a post as part of complete weather station. The VAES was operated by Virginia Technology University and the performance of the gauges was monitored remotely and the sites were visited annually. Although we grouped VAES gauges under operational gauges, they were in the field primarily for research purposes. The VAES gauges started operation between January and March 2004, just a few months before G18 became operational (Fig. 2).

The CRN site (G20) was equipped with the tipping-bucket (Table 1) and Geonor weighing bucket gauges. In this study, we used the tipping-bucket gauge. The quality control of the observations was done at the National Climate Data Center and then placed on the Web site.

The FM site (G21) was equipped with the Forest Services Inc. tipping-bucket gauge (Table 1). The gauge is within National Wildlife Refuge land and local personnel maintain the gauge. The Forest Services technician checks the gauge annually and the gauge is recalibrated every 3 yr (R. Vollick, Fire Management, 2006, personal communication).

Among operational gauges, the CRN gauges were inside a double wooden fence. The AWPAG was equipped with a Tretyakov wind shield (Fig. 3a; Greeney et al. 2007), while the FTB had a vinyl wind shield (Fig. 3b). The rest of the operational and TSVO gauges did not have wind shields. The wind-induced precipitation loss is probably the most notable gauge systematic error. Habib et al. (1999) and Nespor and Sevruk (1999) simulated the wind-induced gauge rainfall loss, while Duchon and Essenberg (2001) examined the differences in rainfall between the pit, shielded, and unshielded gauge measurements. Sieck et al. (2007), with a well-instrumented site, focused on data quality control, and systematic errors including gauge calibration, gauge orifice leveling, and wind effects. The site included various vendors tipping

TABLE 1. Characteristics of the operational gauges.

Organization	Gauge type	Manufacturer/model number	Reporting time/method	Gauge site ID
NWS–ASOS	Weighing bucket	OTT–Germany	Minute/telemetry	G01, G02
NWS/FAA–ASOS	Tipping bucket	Frise	Minute/telemetry	G03–G07
NWS–COOP	Stick gauge	—	Daily/manual reading	G08–G12
FAA–AWOS	Tipping bucket	Texas Electronics/6118–1	Hourly/telemetry	G13–G16
VAES	Tipping bucket	Texas Electronics/525-L	15-min/datalogger/telemetry	G17–G19
CRN	Tipping bucket	Hydrological Services of America/TB-3	15-min/datalogger/satellite transmission	G20
Fire Management	Tipping bucket	Forest Services Inc.	Hourly/datalogger/satellite transmission	G21

TABLE 2. COOP network gauge reading times.

Site ID	Observation time (local time)
G08	1600
G09	1800
G10	0600
G11	0900
G12	0800

buckets, weighing bucket gauges, pit collector gauges, an impact-type disdrometer, and anemometers.

In addition to the systematic errors, rain gauges suffer from random errors. Ciach (2003) presented the random errors utilizing 15 collocated tipping-bucket gauge measurements and pointed out that the averaging of two or more collocated gauge measurements reduce the random errors. Habib et al. (2001b) and Wang et al. (2008), focused on the gauge temporal sampling errors that can be significant at a shorter time (less than 15 min) integration period. These gauge measurement errors should be kept in mind in comparison with the gauge rainfall.

3. Rainfall statistics

A package of rainfall statistics was applied to daily and monthly rainfall to evaluate the performance of the TSVO and operational gauges. The statistical package included the Pearson correlation coefficient, standard deviation, bias, weighted bias, absolute bias, weighted absolute bias, percent bias, and percent absolute bias. The bias in this study means the mean difference between the two variables. The Pearson correlation coefficient, ρ is the ratio of the sample covariance of the two gauge totals (x, y) to the product of the standard deviation and is expressed as

$$\rho = \frac{\mathrm{Cov}(x, y)}{[\mathrm{Var}(x)\mathrm{Var}(y)]^{1/2}}. \tag{1}$$

The standard deviation (SD) of the difference between the gauge totals is the square root of the summation of variance x plus variance y minus 2 times the covariance of the variables and is expressed as

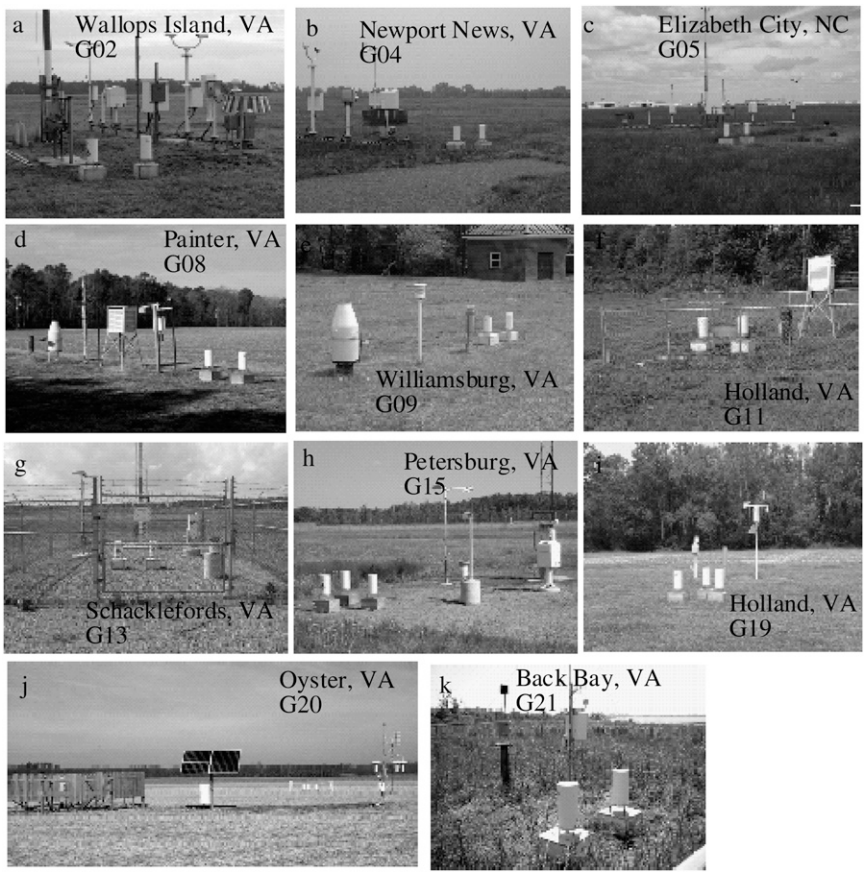

FIG. 3. Picture of the selected sites. (a) ASOS weighing bucket; (b),(c) ASOS tipping-bucket site; (d)–(f) COOP stick; (g),(h) AWOS tipping-bucket site; (i) VAES tipping-bucket site; (j) CRN tipping-bucket site; and (k) FM tipping-bucket site.

$$\mathrm{SD}(x-y)=\sqrt{\mathrm{Var}(x)+\mathrm{Var}(y)-2\mathrm{Cov}(x,y)}. \quad (2)$$

The bias and absolute bias between the two gauge totals are expressed as

$$\mathrm{bias}=\frac{1}{n}\sum_{i=1}^{n}(x_i-y_i), \quad (3)$$

$$\mathrm{absolute_bias}=\frac{1}{n}\sum_{i=1}^{n}|x_i-y_i|, \quad (4)$$

where n represents the rainy days or rainy months. The weighted bias and weighted absolute bias are similar to the biases defined above but weighted with respect to the rain amount:

$$\mathrm{weighted_bias}=\sum_{i=1}^{n}\omega_i(x_i-y_i), \quad (5)$$

$$\mathrm{weighted_absolute_bias}=\sum_{i=1}^{n}\omega_i|x_i-y_i|, \quad (6)$$

where the weighting function is defined as

$$\omega_i=\frac{(x_i+y_i)/2}{\langle x,y\rangle}, \quad (7)$$

where the denominator is the mean gauge total and is expressed as

$$\langle x,y\rangle=\frac{1}{n}\sum_{i=1}^{n}\frac{(x_i+y_i)}{2}. \quad (8)$$

Through weighting function, the rain events that had higher accumulation play a greater role in bias and absolute bias calculations. The percent bias and percent absolute bias, which are normalized quantities, are the ratios of the bias and absolute bias to the mean gauge totals:

$$\mathrm{percent_bias}=\frac{\mathrm{bias}}{\langle x,y\rangle}, \quad (9)$$

$$\mathrm{percent_absolute_bias}=\frac{\mathrm{absolute_bias}}{\langle x,y\rangle}. \quad (10)$$

Although each statistic has its own importance, we selected the percent absolute bias to rate the gauge performance. This is mainly due to the fact that gauge accuracy is often referred to as percent absolute error at a given rain rate interval. For instance, the accuracy of Met One gauges is ±1% at rain rate from 25.4 to 76.2 mm h^{-1}. If the percent bias was less than 5%, the gauge was rated *excellent*. If the percent bias was between 5% and 10% the rating was *very good*, while the rating was *good* when the percent bias was between 10% and 15%. We called the gauge *reasonable* if the percent bias was between 15% and 20%, while a *poor* rating was given when the percent bias was above 20%. This subjective scheme was applied both for daily and monthly statistics.

For daily rainfall, we ran the statistical package in two different conditions. First, we included the rainy days if either TSVO or operational gauge had at least 5 tips (1.2 mm). This tended to eliminate the days that had insignificant rainfall, but not gauge malfunction, which often results in no recorded rainfall. Second, we considered the days where both TSVO and operational gauges had at least five tips. This statistic tends to demonstrate agreement when both gauges are operating. The first criterion was selected to rate the gauge performance. Considering the comparison between the two TSVO gauges, we applied only the second condition since gauge malfunction was mainly due to human error where the gauge was not set back to the operational mode after the site calibration test. Since TSVO gauges are not heated, it is feasible that false tips occur after the accumulated snow melts following the snow event. We carefully eliminated these tips by examining the snowy days in the operational gauge network.

4. Performance of TSVO gauges

Side-by-side comparisons of TSVO gauges at triple gauge sites were mainly done between first (A gauge) and second (B gauge) gauges. This is mainly due to the fact that the third gauge (C gauge) was installed at half of the sites and the installation completed several months after the operation was started. We only used the C gauge in three sites where either the A or B gauge had a data gap due to human error. The comparison of three gauges boosted our confidence on the selection of best gauge before using it as a reference for comparison to the operational gauges.

In addition to side-by-side comparisons, performance of TSVO gauges were evaluated at NASA's Wallops Flight Facility (WFF) where dual TSVO gauges were operated collocated with dual Hydrological Services of America (HAS) gauges. Seventy-five days of accumulative rainfall had less than a 2% different between TSVO and HAS gauges (Fig. 4a). The event rain total statistics, on the other hand, showed less than 5% absolute bias and ±1% percent bias between TSVO and HAS gauges (Fig. 4b).

The comparison of daily rainfall between the two TSVO gauges revealed excellent agreement in 38% of

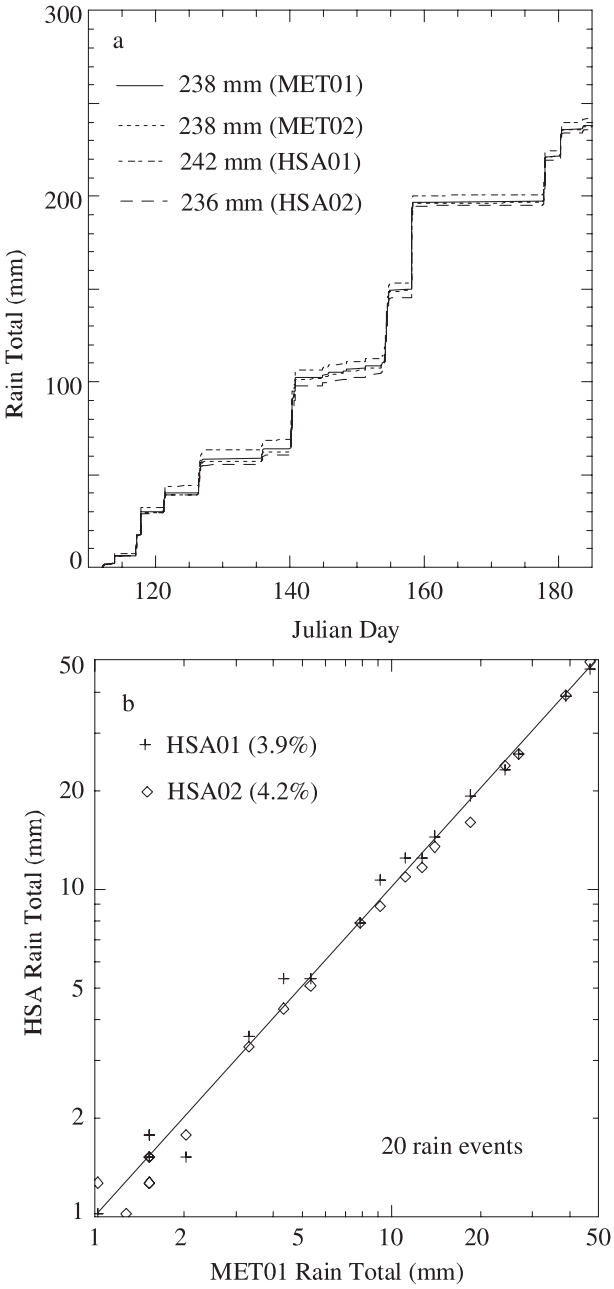

FIG. 4. Comparison of collocated two Met One (MET01, MET02) and two Hydrological Services of America (HSA01, HSA02) tipping-bucket gauge rainfall. (a) Accumulative rainfall for the 75 days. (b) Comparison of event rain totals.

the sites. The remaining sites had very good agreement where the percent absolute bias was between 5% and 9% (Fig. 5h). The correlations were higher than 0.99 (Fig. 5a) and the SD was lower than 2 mm in all sites except G06 where the B gauge had higher totals (Fig. 5b). The higher rain totals of B gauge in G06 resulted in relatively higher biases than other sites (Figs. 5c–f). The weighted bias and weighted absolute bias at G06 were about the same and were above 4 mm, while the rest of the sites had weighted biases less than 3 mm. The absolute value of percent bias was over 7% in G06 and G13 and was lower than 5% in other sites (Fig. 5g). These statistics are considered a benchmark when the TSVO gauge is compared to the operational gauge for daily rainfall. These statistics also demonstrate the error margins of a reliable gauge when a single gauge operates at a site. Although we stress the importance of dual or triple gauges at a site (Ciach and Krajewski 1999; Tokay et al. 2003), the use of a single gauge site may still be acceptable for research application.

5. Performance of operational gauges

The comparison of daily rainfall between the TSVO and operational gauges revealed a range from very good to poor agreement. The agreement between the TSVO and operational network was mostly better for monthly rainfall as expected. There was a clear distinction in gauge performance between the different operational networks. Although we tried to identify the reasoning of the good–poor performance of the gauge, this was not always feasible among several contributing factors including the gauge type, data collection system, the presence or absence of a wind fence, and maintenance.

The performance of operational gauges is evaluated by taking TSVO gauges as a reference and by relying on the statistics between the TSVO and operational gauges. The statistics under the term of *bias* is therefore considered to be an *error* for the operational gauges. The diagrams of daily and monthly rainfall comparisons around the one-to-one line may provide additional information through visual evidence on the over- and underestimation of a gauge and outliers. These diagrams are not shown but included in the discussion when it is relevant.

ASOS weighing bucket (ASOS_WB) gauges at G01 and G02 had very good agreement with the TSVO gauges in daily rainfall (Fig. 5h), while the agreement was excellent in monthly totals (Fig. 6h). The daily totals had only a few outliers from the one-to-one line, resulting in minimal differences in rain statistics when either or both gauges were reporting rainfall. The statistics between the TSVO gauges and between TSVO and ASOS_WB gauges were also very close for correlation, SD, and absolute biases (Figs. 5a,b,e,f). This boosted our confidence on the high performance of these operational gauges. Looking at the monthly statistics, correlations were above 0.99 (Fig. 6a), and standard deviations were under 4 mm (Fig. 6b). The monthly absolute biases and the weighted absolute biases were less than 2 and 3 mm

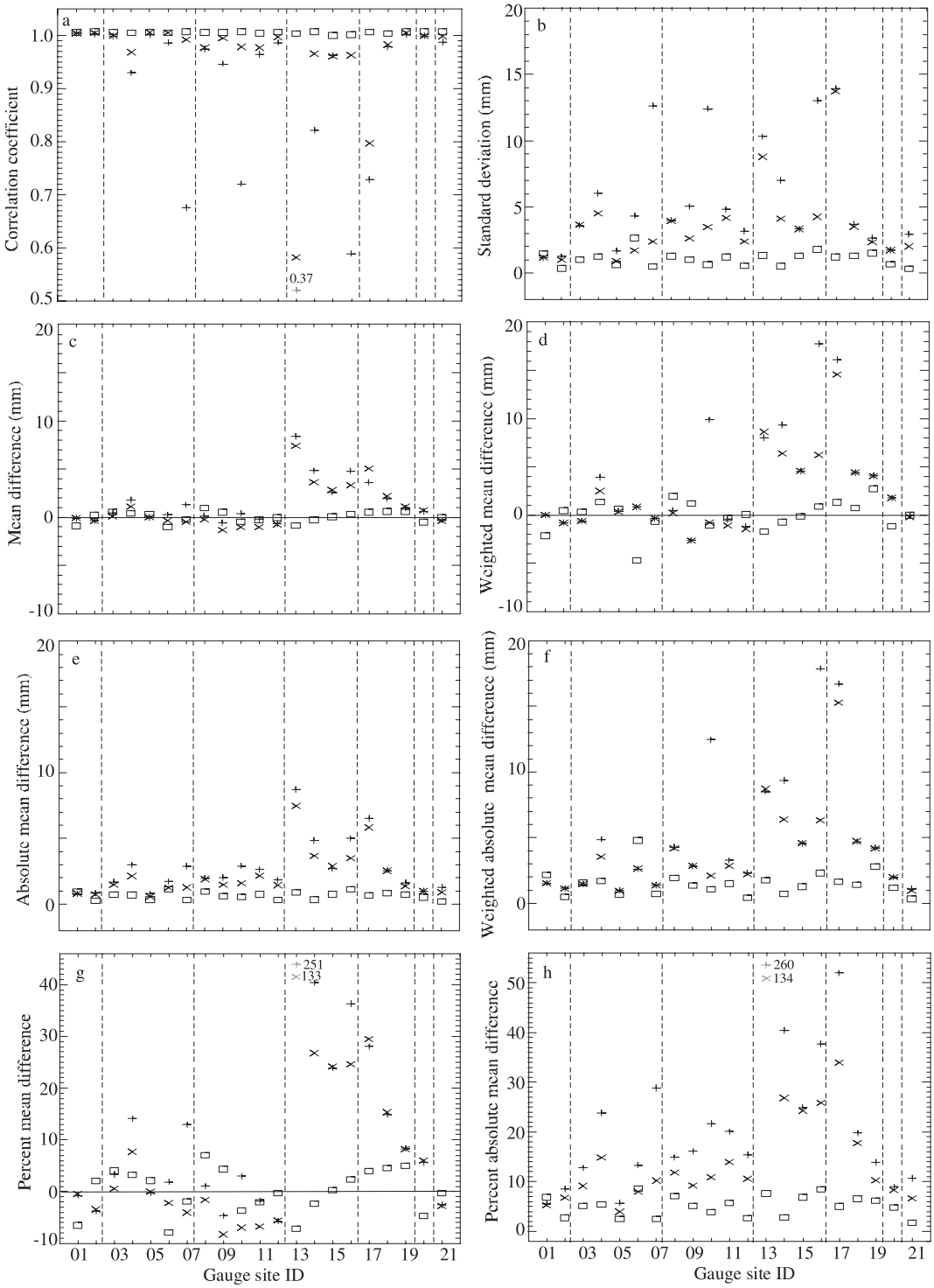

FIG. 5. Daily rainfall statistics between the two TSVO gauges (▭), between the TSVO and operational gauges when either of them recording (+) and when both of them recording (×). The vertical dashed bars separate different types or models of operational gauges.

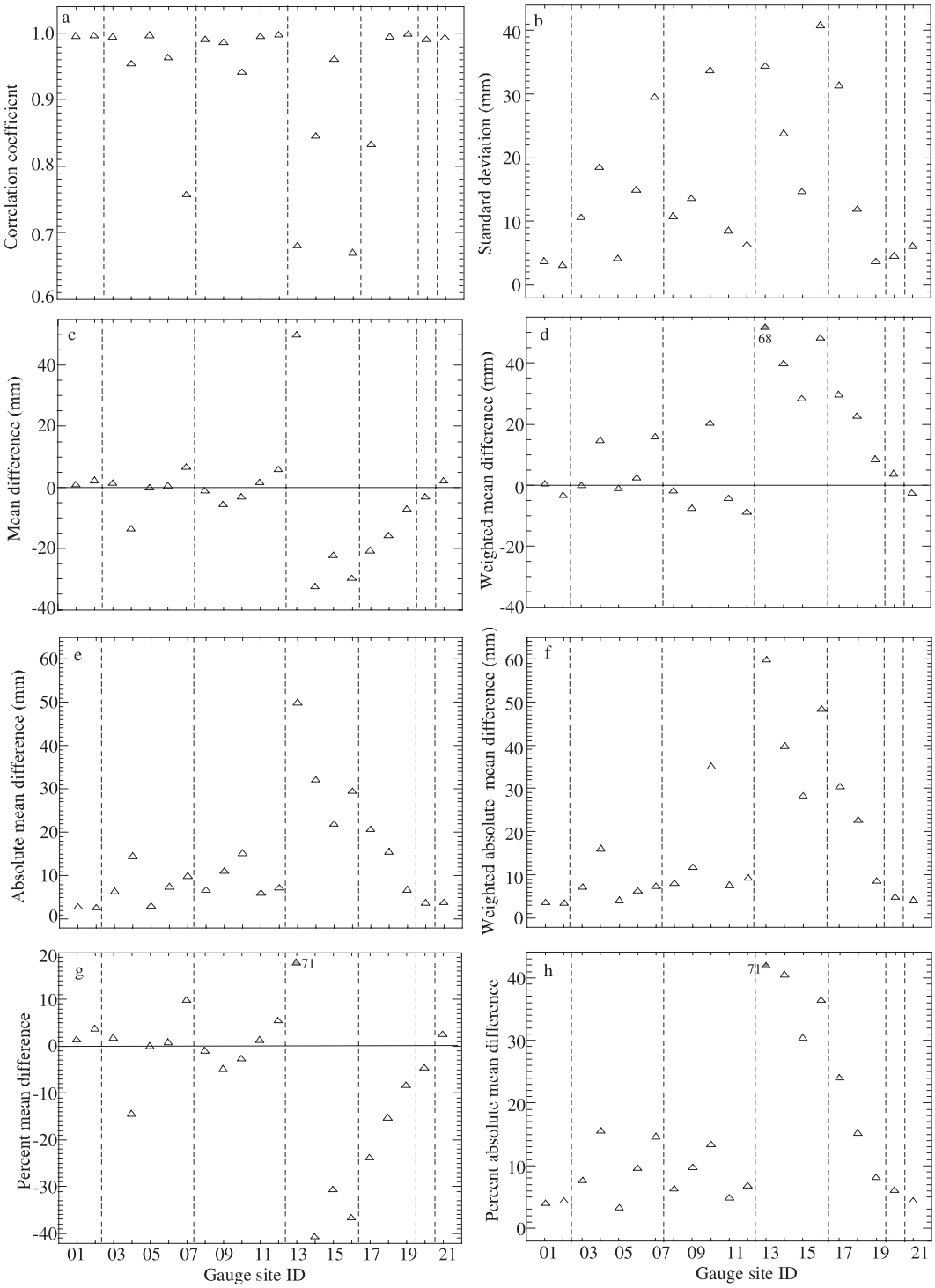

FIG. 6. Monthly rainfall statistics between the TSVO and operational gauges. The vertical dashed bars separate different types or models of operational gauges.

(Figs. 6e,f), respectively, as being the lowest among all operational gauges. These statistics are considered a benchmark when the TSVO gauge is compared with the other operational gauges for monthly rainfall.

ASOS tipping-bucket (ASOS_TB) gauges were rated very good at G05, good at G03 and G06, and poor at G04 and G07 for daily rainfall (Fig. 5h). The scatters around the one-to-one lines were particularly significant at the

G04 and G07 sites. There was also a significant number of days where ASOS_TB either reported false rainfall or did not report any rainfall in all five sites. With the exception of G05, the number of rainy days reduced by about 20% when the daily statistics were recalculated for the days where both gauges reported rainfall. The daily statistics then improved noticeably such that the agreements ranged from excellent (G05) to good (G04 and G07; Fig. 5h). The false and absence of a rain report was indeed one of the major problematic issues of the operational gauges. The correlations at G07 improved from 0.67 to 0.98 and the SD decreased from 12 to 2 mm when both gauges were reporting rainfall (Figs. 5a,b). For the monthly rainfall, the gauge totals followed the one-to-one line in G05, resulting in excellent agreement. In the presence of one to two outliers and deviations from the one-to-one line, G03 and G06 had a very good agreement with the TSVO gauges while the agreement was rated good in G04 and G07 (Fig. 6h). The false or nonreported rainfall influenced the monthly totals as well. The correlation at G07 was the second lowest among all operational gauges (Fig. 6a), and the SD at the same site was the fifth highest (Fig. 6b).

It was not uncommon that an ASOS_TB gauge reported false rainfall or missed rainy days consecutively. G07, for instance, failed to report 8-, 5-, and 50-mm rain totals in 3 consecutive days. G07 also failed to report a 130-mm rain total on another day. Since the daily totals between the TSVO and ASOS_TB gauges were in an agreement before and after these nonreporting days, we attributed the problem to the data transfer system rather than the gauge. Nevertheless, there were other days where the rain totals between the gauges were quite different. Since all the ASOS gauges have the same maintenance and the data has been retrieved through the same mechanism, ASOS_WB gauges seemed to be superior to the ASOS_TB gauges. We should also add that the age of the gauge can be a factor in its performance. While ASOS_WB gauges were new, ASOS_TB gauges had been in the field at least several years before TSVO gauges were installed at the sites.

Performance of COOP gauges ranged from good to poor for daily totals (Fig. 5h). The percent absolute bias had a span of less than 7% between the five sites. The range of percent absolute daily bias was even less than 5% when both gauges were reporting rainfall. This second run of statistics improved the agreement between the gauges noticeably. The correlation in G10, for example, was raised from 0.71 to 0.97 (Fig. 5a). The SD and weighted absolute daily bias at the same site were both lowered from 12 mm to 3 and 2 mm, respectively (Figs. 5b,f). The number of rainy days was 20% less when both gauges were reporting rainfall. As with ASOS_TB sites, the false and missed rainy days contributed significantly to the statistics. Once they were excluded, the percent bias was negative in all sites indicating that the COOP gauges reported higher rainfall (Fig. 5g). The 7%–16% decrease in percent absolute bias between the daily and monthly statistics is a unique feature of the COOP gauges. The COOP gauges were rated from excellent to good based on monthly totals (Fig. 6h). Despite the statistics being based on a large sample of rainy days, not reporting a major rainy day can still affect the statistics, particularly in monthly statistics. The COOP at G10, for example, missed 122-mm total of rainy days, resulting in an outlier in monthly totals and this was the only site that was rated good. The SD and weighted absolute monthly bias were also high at this site (Figs. 6b,f).

The false and unreported daily rainfall in COOP sites is attributed to the observers reading errors. As noted above, the time of the daily measurement is not recorded and this is indeed a major issue if it rains at the time of observation. These factors cause both under- and overestimation of the COOP daily totals. This is why the monthly statistics were substantially better than the daily statistics. Aside from that, the COOP sites results in comparable statistics with the ASOS_TB gauges even though the gauges and data collection methods were very different between the two networks.

AWOS gauges had a poor performance in all sites for daily and monthly rainfall. Among the four AWOS sites, G13 had the poorest performance having an extremely high percent bias and a percent absolute bias in both daily and monthly rainfall (Figs. 5g,h and 6g,h). The G13 gauge did not report or underestimated rainfall in most of the days. It also had the lowest correlation and highest daily bias among all gauges (Figs. 5c,e). Even in the absence of unreported rainy days, the performance of all AWOS gauges was poor because of severe underestimation of rainfall. Given the similarities between ASOS and AWOS sites maintenance, the malfunction of the gauge is believed to be the main reason for the poor performance of the AWOS sites.

Performance of VAES gauges had a wide variety between the three sites. In the presence of false and unreported rainfall and numerous days of underestimation, G17 had a poor performance for daily and monthly rainfall (Figs. 5h and 6h). In contrast, G19, which had a few outliers from the one-to-one line, had good and very good ratings (Figs. 5h and 6h). On the other hand, G18 had a reasonable performance (Figs. 5h and 6h).

The CRN gauge had a very good performance for both daily and monthly rainfall (Figs. 5h and 6h). There were only a very few outliers from the one-to-one line for daily and monthly rainfall and these outliers were at

light rainfall. Given the fact that the CRN sites are well constructed with double wooden fence and the gauges were tested through specialized field campaigns at two NWS research sites (Baker et al. 2005), very good performance of the CRN gauges was not surprising. This particular gauge was activated just 5 months before TSVO gauges were installed at the site.

The FM gauge was rated good for daily rainfall and was excellent for monthly rainfall (Figs. 5h and 6h). False and unreported rainfall was the main reason for the relatively low performance in daily rainfall. In comparison to the CRN site, the maintenance of the FM is less frequent and the gauge had been in the field for 9 yr, so the high performance was rather surprising.

6. Conclusions

As noted in the introduction, the TSVO gauges were deployed to validate the NPOL rainfall estimation. The gauges were collocated with operational gauges, satisfying the open space requirement in the absence of any obstruction. This study was motivated by the validation of satellite rainfall products where the gauges are a key component. The satellite rainfall products are mostly validated over a well-instrumented ground validation site such as east-central Florida or a field campaign site such as the Amazon basin of Brazil, both of which are associated with the TRMM program and are known as physical validation sites. The satellite products also seek validation over areas where there are inconsistencies between precipitation retrieval algorithms. The coastal areas like the mid-Atlantic region is particularly challenging for microwave sensor based precipitation retrieval algorithms. Over the continental United States, the available gauges and rainfall products such as stage III are the only options to validate the space-borne rain estimate in these circumstances (e.g., Yilmaz et al. 2005).

For the decision making to include a particular operational gauge, the user relies on the predetermined threshold of measurement accuracy for a given time scale, even though the user may have an interest in the details of problematic issues between the gauge and data transfer systems. To determine the gauge accuracy, a long-term field campaign with the presence of collocated gauges is desirable. The NWS performs such a field campaign at its two centers (Sterling, Virginia, and Johnstown, Pennsylvania) where various types of gauges have been tested. A similar test facility has been constructed at NASA WFF where more disdrometers than the rain gauges were operated collocated. At WFF, the comparison between two TSVO and two HAS gauges revealed an excellent agreement based on 20 rain events. The agreement between the two TSVO gauges ranged from very good to excellent at operational gauge sites. These findings encouraged us to use TSVO gauges as a reference to determine the performance of operational gauges.

Considering percent absolute bias as a measure for the gauge accuracy, both ASOS_WB, one of the ASOS_TB, and a CRN gauge were qualified as an acceptable gauge for daily rainfall if an agreement threshold of 10% was considered. Since the FM gauge is just above this threshold, it can also be added to the qualified gauges for satellite precipitation validation. For a monthly rainfall, if a 5% threshold is considered, both ASOS_WB, one of the ASOS_TB, one of the COOP gauges, and a FM gauge qualified, while the CRN gauge may also be acceptable since its percent absolute bias was just above 5%. While these subjectively determined thresholds might be considered strict, they indicate that ASOS_WB, CRN, and FM gauges seem to be highly qualified for any application. Since we had only one or two sites for these types of gauges, this statement should be accepted with caution.

The relaxation of the daily rainfall thresholds to 15% increases the number of acceptable gauge sites to 10, while the relaxation of the monthly rainfall to 10% increases the number of acceptable gauge sites to 13. For daily rainfall, two more ASOS_TB, two COOP (one in the gray zone for just being above the threshold), and one VAES gauge qualify, while three more ASOS_TB, three more COOP, and one VAES gauge pass the monthly rainfall threshold. An addition of another 5% to the monthly rainfall threshold results in all ASOS_TB and COOP gauges qualifying as acceptable. For daily rainfall, two out of five ASOS_TB and one COOP still do not qualify at a 20% threshold level. Among other networks, the second VAES gauge also qualified for monthly and daily rainfall with these more relaxed thresholds. Despite the limited number of sites in this study, we conclude that ASOS_TB and COOP gauges can be used for monthly rainfall if the threshold is not strict, while it is risky to use them for daily rainfall. AWOS gauges, on the other hand, should be avoided for any validation and research applications. Recall that the type of gauges might differ at other AWOS sites, so our conclusions should be considered with caution. Since VAES gauges were research oriented, our diverse findings on their performance may provide feedback to the researchers who use these gauges.

This study should be considered as experimental since we neither examined a particular gauge network in depth nor studied all the gauge networks that are used in constructing particular radar or satellite rainfall product. Rather, we investigated the performance of operational gauges that were distributed in approximately 550 km^2. Most of the gauge networks that were used in this study report at hourly and higher temporal resolution and it is

indeed an interest to extend this study to hourly comparisons for a future study. This study alone is not sufficient to evaluate the performance of a particular gauge network even though there is a message for a user to use caution for certain gauge networks. In any case, there is a merit to conduct the experimental studies like this whenever and wherever it is possible and report the findings, so the users will have the knowledge of a particular gauge type or network. This will also help an organization to select a gauge type for a field campaign or long-term operation.

Acknowledgments. We would like to acknowledge the National Climate Data Center, the North Carolina State Climate Office, the Climate Reference Network, the National Weather Service Wakefield Office, Fire Management, the Virginia Agricultural Experimental Station, the Cooperative Observer program, the Automated Surface Observation System, and the Automated Weather Observation System for providing data and feedback on various aspects of data collection and quality and gauge maintenance. The rain gauge project was iniatiated through efforts of Robert F. Adler, former TRMM project scientist, while Richard J. Lawrence of the NASA Goddard Space Flight Center and John Gerlach of the NASA Wallops Flight Facility played a leadership role during the operation. This study was initiated during the third author's summer internship at the NASA Goddard Space Flight Center. We thank the Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, for funding and logistics of the summer student program. We appreciate invaluable comments from Emad Habib of the University of Louisiana, Lafayette, David A. Marks and Jianxin Wang of Science Systems Applications Inc., and Koray K. Yilmaz of the Earth System Science Interdisciplinary Center. The comments from three anonymous reviewers were also helpful. This study was funded through NNX07AF45G under Ramesh Kakar, TRMM program scientist.

REFERENCES

Amitai, E., 2000: Systematic variation of observed radar reflectivity-rainfall rate relations in the tropics. *J. Appl. Meteor.,* **39,** 2198–2208.

Baker, B. C., L. Larson, E. May, H. Bogin, and B. Collins, 2005: Final report: Operational testing of various precipitation sensors in support of the United States Climate Reference network (USCRN). NOAA Tech. Note NCDC USCRN-05-02, 69 pp.

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.,* **12,** 5–19.

Ciach, G. J., 2003: Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Oceanic Technol.,* **20,** 752–759.

——, and W. F. Krajewski, 1999: On the estimation of radar rainfall error varience. *Adv. Water Resour.,* **22,** 585–595.

Datta, S., W. L. Jones, B. Roy, and A. Tokay, 2003: Spatial variability of surface rainfall as observed from TRMM field campaign data. *J. Appl. Meteor.,* **42,** 598–610.

Duchon, C. E., and G. R. Essenberg, 2001: Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. *Water Resour. Res.,* **37,** 3223–3263.

Greeney, C. M., J. V. Flore Jr., J. M. Dover, and M. L. Salyards, cited 2007: Winter test of production all-weather precipitation accumulation gauge for ASOS 2005-2006. [Available online at http://ams.confex.com/ams/87ANNUAL/techprogram/paper_116371.htm.]

Habib, E., and W. F. Krajewski, 2002: Uncertainty analysis of the TRMM ground validation radar-rainfall products: Application to the TEFLUN-B field campaign. *J. Appl. Meteor.,* **41,** 558–572.

——, ——, V. Nešpor, and A. Kruger, 1999: Numerical simulation studies of rain gauge data correction due to wind effect. *J. Geophys. Res.,* **104** (D16), 19 723–19 734.

——, ——, and G. J. Ciach, 2001a: Estimation of rainfall interstation correlation. *J. Hydrometeor.,* **2,** 621–629.

——, ——, and A. Kruger, 2001b: Sampling errors of tipping-bucket rain gauge measurements. *J. Hydrol. Eng.,* **6,** 159–166.

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

——, C. G. Malakpet, A. Tokay, and P. A. Kucera, 2008: Sensitivity of streamflow simulations to temporal variability and estimation of *Z–R* relationships. *J. Hydrol. Eng.,* **13,** 1177–1186.

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.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected global precipitation. *Int. J. Climatol.,* **10,** 111–127.

Matrosov, S. Y., K. A. Clark, B. E. Martner, and A. Tokay, 2002: X-band polarimetric measurements of rainfall. *J. Appl. Meteor.,* **41,** 941–952.

Nespor, V., and B. Sevruk, 1999: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *J. Atmos. Oceanic Technol.,* **16,** 450–464.

Rosenfeld, D., D. B. Wolff, and D. Atlas, 1993: General probability-matched relations between radar reflectivity and rain rate. *J. Appl. Meteor.,* **32,** 50–72.

——, ——, and E. Amitai, 1994: The window probability matching method for rainfall measurements with radar. *J. Appl. Meteor.,* **33,** 682–693.

Schroeder, J. L., W. S. Burgett, K. B. Haynie, I. Sonmez, G. D. Skwira, A. L. Doggett, and J. W. Lipe, 2005: The West Texas Mesonet: A technical overview. *J. Atmos. Oceanic Technol.,* **22,** 211–222.

Sieck, L. C., S. J. Burges, and M. Steiner, 2007: Challenges in obtaining reliable measurements of point rainfall. *Water Resour. Res.,* **43,** W04420, doi:10.1029/2005WRR004519.

Tokay, A., D. B. Wolff, K. R. Wolff, and P. Bashor, 2003: Rain gauge and disdrometer measurements during the Keys Area Microphysics Project (KAMP). *J. Atmos. Oceanic Technol.,* **20,** 1460–1477.

Wang, J., B. L. Fisher, and D. B. Wolff, 2008: Estimating rain rates from tipping bucket rain gauge measurements. *J. Atmos. Oceanic Technol.,* **25,** 43–56.

White, S. G., L. J. Winans, and J. V. Fiore, 2004: Development of the all-weather precipitation accumulation gauge for ASOS. Preprints, *Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface,* Seattle, WA, Amer. Meteor. Soc., 7.3. [Available online at http://ams.confex.com/ams/pdfpapers/69273.pdf.]

Wolff, D. B., D. A. Marks, E. Amitai, D. S. Silberstein, B. L. Fisher, A. Tokay, J. Wang, and J. L. Pippitt, 2005: Ground validation for the Tropical Rainfall Measuring Mission (TRMM). *J. Atmos. Oceanic Technol.,* **22,** 365–380.

Xie, H., X. Zhou, E. R. Vivoni, J. M. H. Hendrickx, and E. E. Small, 2005: GIS-based NEXRAD stage III precipitation database: Automated approaches for data processing and visualization. *Comput. Geosci.,* **31,** 65–76.

Yilmaz, K. K., T. S. Hogue, K. 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.